# ZULEICA ANTONIA DE CAMARGO

ANÁLISE DA QUALIDADE VOCAL DE
UM GRUPO DE INDIVÍDUOS DISFÔNICOS:
uma abordagem interpretativa e integrada de dados
de natureza acústica, perceptiva e eletroglotográfica

Doutorado em Lingüística Aplicada e Estudos da Linguagem

Pontifícia Universidade Católica de São Paulo São Paulo 2002

#### **ZULEICA ANTONIA DE CAMARGO**

# ANÁLISE DA QUALIDADE VOCAL DE UM GRUPO DE INDIVÍDUOS DISFÔNICOS:

uma abordagem interpretativa e integrada de dados de natureza acústica, perceptiva e eletroglotográfica

Tese apresentada à Banca Examinadora da Pontifícia Universidade de São Paulo, como exigência parcial para obtenção do título de Doutor em Lingüística Aplicada e Estudos da Linguagem sob a orientação da Prof.ª Doutora Sandra Madureira

Doutorado em Lingüística Aplicada e Estudos da Linguagem

Pontifícia Universidade Católica de São Paulo São Paulo 2002

Camargo Zuleica Antonia

Análise da qualidade vocal de um grupo de indivíduos disfônicos: umaabordagem interpretativa e integrada de dados de natureza acústica, perceptiva e eletroglotográfica / Camargo Zuleica Antonia.—São Paulo, 2002. xxvi, 312 f.

Tese (Doutorado) – Pontifícia Universidade Católica de São Paulo. Programa de Estudos Pós-Graduados em Lingüística Aplicada e Estudos da Linguagem.

Título em inglês: Voice quality analysis of a group dysphonic subjects: an interpretative and integrated account of acoustic, perceptual and eletroglottographic data.

1. Voz. 2. Disfonia. 3. Acústica. 4. Percepção auditiva. 5. Fonação.

Autorizo,	exclusivamente	para fins	acadêmicos	e científicos,	a reprodução	total ou	parcial
desta tese	por processos de	e fotocopia	adoras ou ele	etrônicos.			

Assinatura:São Paulo, 30 de Agosto de 2002.
---

# PONTIFÍCIA UNIVERSIDADE CATÓLICA DE SÃO PAULO PROGRAMA DE ESTUDOS PÓS-GRADUADOS EM LINGÜÍSTICA APLICADA E ESTUDOS DA LINGUAGEM

Coordenadora do Curso de Pós-Graduação: Beth Brait

# Zuleica Antonia de Camargo

# ANÁLISE DA QUALIDADE VOCAL DE UM GRUPO DE INDIVÍDUOS DISFÔNICOS:

uma abordagem interpretativa e integrada de dados de natureza acústica, eletroglotográfica e perceptiva

Presidente da banca: Prof.ª Dr	a
	BANCA EXAMINADORA
Prof. Dr	
Prof. <sup>a</sup> Dr <sup>a</sup>	
Prof. <sup>a</sup> Dr <sup>a</sup>	
Prof. <sup>a</sup> Dr <sup>a</sup>	
Ap	orovada em//

## Aos meus pais,

Pela valorização das minhas qualidades e, principalmente, aceitação dos meus defeitos. Por representarem a possibilidade de vislumbrar um porto seguro em todos os momentos, desde aqueles de maior alegria, até os de maior impacto, e que não foram poucos, ao longo desta trajetória.

## **Agradecimento Especial**

Prof.<sup>a</sup> Dr<sup>a</sup> Sandra Madureira,

Pelo sopro de inspiração e de criação, com incentivo constante à busca do saber e à humildade no reconhecimento do que ainda é desconhecido e intocado.

Espero que as marcas de seus conhecimentos, dedicação e harmonia, presentes ao longo do processo de orientação, permeiem as entrelinhas do material impresso nas páginas do trabalho apresentado.

#### **Agradecimentos**

Nas páginas desta tese, além dos conhecimentos científicos que carregam a inerente busca pela retrospectiva histórica, pelo primor do método e pela articulação de conteúdo, registra-se a história de participações especiais:

Prof. Dr Maurílio Nunes Vieira, pela disponibilização de *softwares* de sua autoria e participação nas diversas etapas de avaliação da pesquisa em questão, além de constantes e enriquecedoras discussões teóricas.

Prof. Mário Augusto de Souza Fontes, pela convivência agradável durante o período de elaboração deste trabalho, por sua dedicação ao acompanhamento dos procedimentos de coleta e análise de dados, bem como pela disponibilidade para a discussão de aspectos teóricos relevantes ao projeto.

Prof. Dr Domingos Hiroshi Tsuji, pelo incentivo à pesquisa e cuidadosa avaliação da população estudada.

Dr Luiz Antonio da Silva Freire, pela amizade, apoio em diversas etapas e disponibilidade na digitalização e impressão das imagens.

Dra Irene Queiroz Marchesan e Dr Jaime Luiz Zorzi pela amizade, confiança, trabalho partilhado e apoio na aquisição de equipamentos para desenvolvimento da pesquisa.

Prof.ª Drª Léslie Piccolotto Ferreira, por propiciar o primeiro contato com os profissionais do campo da Fonética da PUC/SP, pelo trabalho e amizade partilhados ao longo dos últimos anos e por seu exemplo de dedicação e constante luta pelo engrandecimento da Fonoaudiologia.

Prof.<sup>a</sup> Dr<sup>a</sup> Iara Bittante, pela amizade, apoio e confiança em mim depositada, bem como colaboração pela colaboração direta ao projeto.

Dr<sup>a</sup> Silvia Pinho pela amizade, pelos ensinamentos transmitidos e constante incentivo ao crescimento profissional.

Prof.<sup>a</sup> Dr<sup>a</sup> Regina Yu, pelas demonstrações de amizade e valorização das diversas etapas de elaboração deste trabalho.

Ernesto Luís Foschi, Laboratório de Rádio da COMFIL, pela dedicação e cuidado ao longo de muitas horas de registro e edição das amostras estudadas.

Prof<sup>a</sup> Yara Castro, pelos ensinamentos preciosos e assessoria estatística.

Fonoaudióloga Monique Donata Tonnini pelo trabalho conjunto, pela amizade e, especialmente, pela companhia, incentivo e dedicação nos momentos de maior dificuldade.

Fonoaudiólogas Caroline Sarkovas, Fabiana Gregio e Fabíola Ariadne Pereira, por trabalharem comigo de forma especial e dedicada ao longo dos últimos anos.

Maurício Chacur, pela colaboração sempre presente e valioso apoio em informática.

Dra Rossana Campos Pereira, pela dedicação na seleção da casuística e seu envolvimento especial com a população atendida.

Profissionais participantes das etapas de análise perceptivo-auditiva, pela paciência e dedicação.

Pesquisadores de outros serviços por propiciarem a oportunidade de discussão do projeto: Johan Sundberg *KTH* (*Stockholm University - Sweden*), Janet M. Beck (*University of Edinburgh - Scotland*) e Prof. Dr Leonardo Fuks (Departamento de Música – Universidade Federal do Rio de Janeiro)

Prof.ª Dr.ª Edmée Brandi, pelo exemplo de dedicação e amor aos estudos da "voz falada", pelo incentivo expresso na forma de calorosas discussões teóricas mescladas a iguais doses de amizade e solidariedade.

Indivíduos participantes da pesquisa, pela inestimável colaboração.

A todos que propiciaram as condições necessárias para o desenvolvimento deste projeto:

Chefias da clínica de Otorrinolaringologia e setor de Fonoaudiologia do Hospital do Servidor Público Municipal (HSPM), Dr Lupércio O. do Valle e Fgas Delmira de Fraga e Karmann e Walquiria Kobayashi.

Diretoria da Divisão Médica e Superintendência do Hospital do Servidor Público Municipal (HSPM), especialmente ao Dr José Francisco Ferraz Ribeiro de Figueiredo pelo reconhecimento da importância de meu afastamento da instituição para cumprir as atividades do curso de doutorado.

Às profissionais do setor de Fonoaudiologia que propiciaram minha ausência ao assumirem o atendimento a mim destinado.

Às bibliotecárias do HSPM, pela dedicação e apoio.

Aos médicos da clínica de Otorrinolaringologia Antonio César B. Augusto, Fátima Regina Abreu Alves, Romualdo Suzano Loureiro Thiago, pela ajuda na seleção do grupo estudado e pelo incentivo ao desenvolvimento deste trabalho.

Dr Marco Gazel pela colaboração e atenção na etapa de coleta de dados.

Fonoaudióloga Rita Mor, pela amizade e carinho constantes, além da colaboração em várias etapas de adequação dos procedimentos e coleta de dados.

Prof.ª Dr.ª Marta A. Andrada e Silva, pela convivência, incentivo e disponibilidade em cooperar, principalmente com uma palavra amiga.

Fonoaudiólogas especialistas em voz pelo CEFAC- São Paulo (Márcia Marçal, Cláudia Pacheco, Maria Valéria e visitantes Camila e Ana Lúcia), pela colaboração em fases

prévias de testes para adequação dos procedimentos.

Cristina Obo, pela prontidão e habilidade em propiciar tempo disponível para o desenvolvimento deste trabalho.

Acácio da Paz Silva, Jefferson dos Santos, Leila H. Marreiro e Maurício Alexandre S. Campos, do Laboratório de Linguagem e Informática da COMFIL (PUC/SP), pelo auxílio constante e agradável convivência ao longo dos últimos anos.

Eduardo V. Kawanishi, pelo empenho manifestado na edição e impressão do trabalho e por sua dedicação às atividades do LIAAC.

Rodney Franco de Lima e Marcelo Fonseca Ferreira do Laboratório de Informática da Psicologia da PUC/SP, pela paciência e dedicação no apoio à execução de gráficos.

Maurício de Cunto, pela constante assessoria envolvendo instrumental de análise.

Graziela Correa da Costa, pela amizade e preciosa colaboração na localização de bibliografia.

Martin Rothenberg e Donald Verrette, da *Glotal Enterprises*, pela pronta colaboração ao uso do eletroglotógrafo.

Chefia do Departamento de Lingüística da COMFIL, Prof.ª Dr.ª Maria Aparecida Caltabiano, pelo apoio e compreensão.

Colegas do Grupo de Estudos Sobre a Fala, pelas atividades compartilhadas, com calorosas discussões que certamente contribuíram para o trabalho apresentado.

Companheiras de pós-graduação, Ana Márcia Vasconcelos, Prof.ª Beatriz Castro de Andrade Mendes, Prof.ª Luisa Barzaghi Ficker e Prof.ª Paulina Artimonte Rocca, pelas sugestões ao projeto, bem como apoio e incentivo em todas as etapas.

Prof.ª Dr.ª Aglael Gama Rossi, pela agradável convivência no LIAAC, pelo estímulo e colaboração nas discussões teóricas.

Alunos de todos os tempos, pela oportunidade de articular conteúdos e de enfrentar novos desafios na busca às constantes questões que emergem das atividades didáticas.

Atuais e ex-orientandos dos cursos de graduação e especialização por acolherem as minhas propostas, com dedicação e envolvimento.

Comissão de Ensino e Pesquisa (CEPE) da Pontifícia Universidade Católica de São Paulo, pelo subsídio financeiro à pesquisa na modalidade de capacitação docente.

Àqueles que demonstraram seu incentivo na forma de presença carinhosa e atenciosa, compreendendo muitas de minhas ausências dos círculos de amigos e familiares:

Amigos do HSPM e da vida (Dinorah Queiroz Lima de Vita, Flávia Villin Denunci, Filomena Rita Palermo Perez, Maria Salete S. Nurmberger, Delmira de Fraga e Karmann, Estela Gomes, Monica Fragoso, Susana P. P. Giannini, Maria Carolina P. P de Barros, Walquiria Kobayashi e Edienete Lúcia Viviani).

Querido grupo GH (Ana Cláudia Barrionuevo, Regina Kubota, Sandra Gedanke e Simone Favaretto).

Amigos Monique D. Tonini, Marcos Loreto e Zeca.

Alunas e amigas especiais Flávia Gobbi e Sabrina Araújo.

Colegas de profissão e amigos Tércia M. F. Ferreira e Vicente Assencio.

Midori Hanayama, pela presença amiga em muitas horas.

Ao simpático e acolhedor casal Pedro e Zélia.

Gú (Augusto), Dadá, Gui, Nardinho, Gu (Gustavo), Dedé e Márcio.

Finalmente, meu agradecimento especial aos numerosos membros da família Miguel Dallora, por terem transmitido a mim a importância de valores como o respeito ao ser humano, solidariedade, determinação e, especialmente, a curiosidade e o fascínio pela imersão nos domínios do desconhecido.

Figura 1.	Lista de Figuras Representação da onda Lx e as informações das fases do ciclo vibratório (fases de fechamento, abertura e aberta)	87
Figura 2.	Ilustração dos traçados de eletroglotografia (a) com flutuação de linha de base (b) e filtrada para a mesma emissão de vogal [a] destacada de contexto de fala encadeada ("a arara")	114
Figura 3.	Ilustração de procedimento voltado à extração dos valores de <i>VOT</i> , com plosão indicada por seta de traçado contínuo e início da sonorização da vogal com seta de traçado pontilhado nas representações de forma da onda acústica (tela superior) e espectrograma de banda larga (tela inferior) para a emissão da sílaba [ka]	119
Figura 4.	Ilustração de procedimento voltado à extração dos valores de <i>VOT</i> , com presença de ruído transiente indicada por seta com traçado contínuo na representação da onda acústica (tela superior) e início da sonorização da vogal por seta com traçado pontilhado na representação da onda Lx (tela inferior) para a emissão da sílaba [ka]	120
Figura 5.	Procedimento de extração de medidas f0 e coeficiente de contato (CC) da onda Lx a partir das estimativas de duração da fase fechada do ciclo vibratório (tela superior) e da duração total (tela inferior)	129
Figura 6.	Espectros de longo termo para duas leituras com estímulos compostos somente com sons sonoros (linha pontilhada) e com composição mista de sons surdos e sonoros (linha contínua) para os falantes A e B, participantes da etapa de teste dos efeitos do contexto nos aspectos de longo termo da emissão	133
Figura 7.	Espectros de longo termo (ELT) de três leituras de texto realizadas por i1, i2, i3, i4 e iR	135
Figura 8.	Distribuição das médias de concentração de energia (em dB) por faixas de freqüência (0-1 kHz, 1-3, 3-5 kHz e 5-11 kHz) dos espectros de longo termo para três leituras efetuadas pelos indivíduos do grupo estudado (i1 a i4) e referência (iR)	138
Figura 9.	Espectrogramas de banda estreita com limites superiores de freqüência estabelecidos em 5000 Hz (a) e 11000 Hz (b) da mesma emissão da vogal [a:] por i1, digitalizada na freqüência de amostragem de 22050 Hz	139
Figura 10.	Espectrograma de banda estreita com limite superior de freqüência estabelecido em 11000 Hz da emissão [aaa] por i1, digitalizada na freqüência de amostragem de 22050 Hz	140

Figura 11.	Espectrogramas de banda estreita com limites superiores de freqüência estabelecidos em 2500 Hz (a) e 11000 Hz (b) da mesma emissão da vogal [a:] por i2, digitalizada na freqüência de amostragem de 22050 Hz	141
Figura 12.	Espectrogramas de banda estreita com limites superiores de freqüência estabelecidos em 5000 Hz (a) e 22000 Hz (b) da emissão [aaa] por i2, digitalizada respectivamente nas freqüências de amostragem de 22050 Hz e de 44100 Hz	142
Figura 13.	Espectrogramas de banda estreita com limites superiores de freqüência estabelecidos em 2500 Hz (a) e 11000 Hz (b) da mesma emissão da vogal [a:] por i3, digitalizadas respectivamente nas freqüências de amostragem de 11025 e 22050 Hz	143
Figura 14.	Espectrogramas de banda estreita com limites superiores de freqüência estabelecidos em 5000 Hz (a) e 22000 Hz (b) da emissão [aaa] por i3, digitalizada respectivamente nas freqüências de amostragem de 22050 Hz e de 44100 Hz	144
Figura 15.	Espectrogramas de banda estreita com limites superiores de freqüência estabelecidos em 11000 Hz (a) e 22000 Hz (b) da mesma emissão da vogal [a:] por i4, digitalizada nas freqüências de amostragem de 22050 Hz e 44100 Hz	145
Figura 16.	Espectrogramas de banda estreita com limites superiores de freqüência em 5000 Hz (a) e 22000 Hz (b) da emissão [aaa] por i4, digitalizada respectivamente nas freqüências de amostragem de 22050 Hz e de 44100 Hz	146
Figura 17.	Espectrogramas de banda estreita com limites superiores de freqüência estabelecidos em 5000 Hz (a) e 11000 Hz (b) da mesma emissão da vogal [a:] por iR, digitalizada na freqüência de amostragem de 22050 Hz.	147
Figura 18.	Espectrograma de banda estreita com limite superior de freqüência em 11000 Hz da emissão [aaa] por iR, digitalizada na freqüência de amostragem de 22050 Hz	148
Figura 19.	Espectrogramas de banda larga com limites superiores de freqüência estabelecidos em 5000 Hz (a) e 11000 Hz (b) da mesma emissão da vogal [a:] por i1, digitalizada na freqüência de amostragem de 22050 Hz.	149
Figura 20.	Espectrograma de banda larga com limite superior de freqüência estabelecido em 11000 Hz da mesma emissão da vogal central de [aaa] por i1, digitalizada na freqüência de amostragem de 22050 Hz	150

Figura 21.	Espectrogramas de banda larga com limites superiores de freqüência estabelecidos em 5000 (a) Hz e 11000 Hz (b) da mesma emissão de "a arara" por i1, digitalizada na freqüência de amostragem de 22050 Hz	151
Figura 22.	Espectrograma de banda larga com limite superior de frequência estabelecido em 11000 Hz (b) da emissão da vogal [a:] por i2 digitalizada na frequência de amostragem de 22050 Hz	152
Figura 23.	Espectrograma de banda larga (a) com limite superior de freqüência estabelecido em 11000 Hz e envelope espectral (b) da mesma emissão da vogal central de [aaa] por i2, digitalizada na freqüência de amostragem de 22050 Hz	153
Figura 24.	Espectrogramas de banda larga com limites superiores de freqüência estabelecidos em 11000 (a) Hz e 22000 Hz (b) da mesma emissão de "a arara" por i2, digitalizada nas freqüências de amostragem de 22050 Hz e 44100 Hz	154
Figura 25.	Espectrogramas de banda larga com limites superiores de freqüência estabelecidos em 11000 Hz (a) e 22000 Hz (b) da mesma emissão da vogal [a:] por i3, digitalizada nas freqüências de amostragem de 22050 Hz e 44100 Hz	155
Figura 26.	Espectrogramas de banda larga com limites superiores de freqüência estabelecidos em 11000 (a) Hz e 22000 Hz (b) e traçado de forma da onda (c) da mesma emissão da vogal central de [aaa] por i3, digitalizada nas freqüências de amostragem de 22050 Hz e 44100 Hz	156
Figura 27.	Espectrogramas de banda larga com limites superiores de freqüência estabelecidos em 11000 (a) Hz e 22000 Hz (b) da mesma emissão de "a arara" por i3, digitalizada nas freqüências de amostragem de 22050 Hz e 44100 Hz	157
Figura 28.	Espectrogramas de banda larga com limites superiores de freqüência estabelecidos em 11000 Hz (a) e 22000 Hz (b) da mesma emissão da vogal [a:] por i4, digitalizada nas freqüências de amostragem de 22050 Hz e 44100 Hz.	158
Figura 29.	Espectrogramas de banda larga com limites superiores de freqüência estabelecidos em 11000 (a) Hz e 22000 Hz (b) da mesma emissão da vogal central de [aaa] por i4, digitalizada nas freqüências de amostragem de 22050 Hz e 44100 Hz	159
Figura 30.	Espectrograma de banda larga com limite superior de frequência estabelecido em 11000 (a) Hz da mesma emissão de "a arara" por i4, digitalizada na frequência de amostragem de 22050 Hz	160

Figura 31.	Espectrograma de banda larga com limite superior de freqüência estabelecido em 11000 Hz da emissão da vogal [a:] por iR, digitalizada na freqüência de amostragem de 22050 Hz	160
Figura 32.	Espectrogramas de banda larga com limites superiores de freqüência estabelecidos em 11000 (a) Hz e 22000 Hz (b) da mesma emissão da vogal central de [aaa] por iR, digitalizada nas freqüências de amostragem de 22050 Hz e 44100 Hz	161
Figura 33.	Espectrogramas de banda larga com limites superiores de freqüência estabelecidos em 11000 (a) e 22000 Hz da mesma emissão de "a arara" por iR, digitalizada nas freqüências de amostragem de 22050 Hz e 44100 Hz	162
Figura 34.	Distribuição das médias de medidas de <i>VOT</i> em segundos de três emissões para as consoantes plosivas surdas nas sílabas [p], [t] e [k] e sonoras nas sílabas [b] [d] e [g] dos indivíduos do grupo estudado (i1 a i4) e do indivíduo referência (iR).	164
Figura 35.	Distribuição das médias dos valores de VOT em segundos de consoantes plosivas surdas [p], [t] e [k] nos respectivos vocábulos para, estar e aplicada e sonoras [b], [d] e [g] nos respectivos vocábulos humildade e lugares emitidos pelo grupo estudado (i1 a i4) e indivíduo referência (iR)	164
Figura 36.	Médias de frequência (Hz) e amplitude (dB) do primeiro (H1) e segundo (H2) harmônicos de três emissões da vogal [a] em posição tônica da emissão de "a arara" em três leituras dos indivíduos do grupo estudado (i1 a i4) e referência (iR)	167
Figura 37.	Traçados cepstrais com informações de qüefrência (em ms) e nível de amplitude para a emissão de "a arara" por i1, i2, i3, i4 e iR	167
Figura 38.	Médias de freqüência (Hz) e amplitude (dB) dos quatro primeiros formantes (F1, F2, F3 e F4) da emissão da vogal [a] em posição tônica em "a arara" para três leituras pelos indivíduos do grupo estudado (i1 a i4) e do indivíduo referência (iR)	170
Figura 39.	Julgamentos de grau geral de alteração vocal e de agradabilidade atribuídos às amostras de fala dos indivíduos do grupo estudado (i1 a i4) e indivíduo referência (iR) nas etapas 1A e 1B de julgamento da qualidade vocal	172
Figura 40.	Total de acertos em julgamentos de inteligibilidade para as três emissões das sílabas [pa], [pε], [bε], [tar], [da], [ka] e [ga] dos indivíduos do grupo estudado (i1 a i4) e indivíduo referência (iR)	175

Figura 41.	Traçados da onda Lx (a) para a emissão de vogal [a] central de [aaa] e DEGG (b) para a emissão vogal [a] tônica na emissão "a arara" (b) de i1	17
Figura 42.	Traçados da onda Lx (a) e DEGG (b) para a emissão vogal [a] tônica em "a arara" e traçado da onda Lx (c) para emissão em freqüência aguda de i2	17
Figura 43.	Traçados da onda Lx (a) para emissões de vogal [a] em "a arara" e (b) vogal [a] central da seqüência [aaa] (c) de i3	18
Figura 44.	Traçado da onda Lx para a emissão de [i] em posição tônica, representativa da maior amplitude detectada nas amostras de i4	18
Figura 45.	Traçados da onda Lx (a) para as emissões de vogal [a] em posição tônica em "a arara" e traçado DEGG (b) para a emissão vogal [a] em posição tônica em "a arara" de iR	18
Figura 46.	Distribuição de medidas de f0 (Hz), Jitter (%), Shimmer (%), coeficiente de contato (%) e índice de velocidade para três emissões da vogal central das seqüências de emissões [aaa] e duas emissões da vogal tônica das emissões "a arara" para dois indivíduos do grupo estudado (i1 e i2) e indivíduo referência (iR)	18
Figura 47.	Distribuição de medidas eletroglotográficas de f0 (Hz) e quociente de contato (%) e acústica de tempo de subida (s) da vogal central das seqüências de emissões [aaa] e da vogal tônica das emissões "a arara" para os indivíduos do grupo estudado (i1 a i4) e indivíduo referência (iR)	18
Figura 48.	Imagens de laringe durante respiração (a) e fonação (b), proveniente de exame telescópico de laringe de i1. Em (b) associação de videoestroboscopia	19
Figura 49.	Imagens de videoquimografia da laringe durante a fonação de i1	19
Figura 50.	Imagens de laringe durante respiração (a) e fonação (b), proveniente de exame nasolaringoscópico de laringe de i2. Em (b) associação de videoestroboscopia	19
Figura 51.	Imagens de laringe durante respiração (a) e fonação (b) (c), proveniente de exame telescópico de laringe de i3. Em (b) (c) associação de videoestroboscopia	19
Figura 52.	Imagens de videoquimografia da laringe durante a fonação de i3	19
Figura 53.	Imagens de laringe durante respiração (a) e fonação (b), proveniente de exame telescópico de laringe de i4. Em (b) associação de videoestroboscopia.	19

Figura 54.	Imagens de videoquimografia da laringe durante a fonação de i4	194
Figura 55.	Distribuição das medidas de VOT em segundos para as três emissões de consoante plosiva bilabial surda [p] em para dos indivíduos do grupo estudado (i1 a i4) e do indivíduo referência (iR)	255
Figura 56.	Distribuição das medidas de VOT em segundos para as três emissões de consoante plosiva bilabial surda [p] em esperto dos indivíduos do grupo estudado (i1 a i4) e do indivíduo referência (iR)	255
Figura 57.	Distribuição das medidas de VOT em segundos para as três emissões de consoante plosiva bilabial sonora [b] em bela dos indivíduos do grupo estudado (i1 a i4) e do indivíduo referência (iR)	256
Figura 58.	Distribuição das medidas de VOT em segundos para as três emissões de consoante plosiva alveolar surda [t] em estar dos indivíduos do grupo estudado (i1 a i4) e do indivíduo referência (iR)	256
Figura 59.	Distribuição das medidas de VOT em segundos para as três emissões de consoante plosiva alveolar sonora [d] em humildade dos indivíduos do grupo estudado (i1 a i4) e do indivíduo referência (iR)	257
Figura 60.	Distribuição das medidas de VOT em segundos para as três emissões de consoante plosiva velar surda [k] em aplicada dos indivíduos do grupo estudado (i1 a i4) e do indivíduo referência (iR)	257
Figura 61.	Distribuição das medidas de VOT em segundos para as três emissões de consoante plosiva velar sonora [g] em lugares dos indivíduos do grupo estudado (i1 a i4) e do indivíduo referência	
Figura 62.	(iR)	257 259

# Lista de Quadros

QUADRO 1.	PROPOSTA DE AVALIAÇÃO DA QUALIDADE VOCAL COM MOTIVAÇÃO FONÉTICA	77
QUADRO 2.	DESCRIÇÃO DA IDADE, NATURALIDADE, GRAU DE ESCOLARIDADE E PROFISSÃO DOS INDIVÍDUOS PORTADORES DE ALTERAÇÕES VOCAIS (i1 A i4) E DO INDIVÍDUO REFERÊNCIA (iR)	96
QUADRO 3.	DESCRIÇÃO DE CAUSA E DURAÇÃO (EM MESES) DO QUADRO CLÍNICO APRESENTADO PELOS INDIVÍDUOS PORTADORES DE ALTERAÇÕES VOCAIS (i1 A i4)	97
QUADRO 4.	JULGAMENTOS DE QUALIDADE VOCAL ATRIBUÍDOS ÀS AMOSTRAS DE FALA DOS INDIVÍDUOS DO GRUPO ESTUDADO (I1 A I4) E INDIVÍDUO REFERÊNCIA (IR) NAS ETAPAS 1A E 1B	173
QUADRO 5.	PERFIL DA QUALIDADE VOCAL SEGUNDO JULGAMENTOS ATRIBUÍDOS ÀS AMOSTRAS DE FALA DOS NDIVÍDUOS DO GRUPO ESTUDADO (i1 a i4) E INDIVÍDUO REFERÊNCIA (iR) NAS FTAPAS 2A F 2B	174

## Lista de Tabelas

Tabela 1.	Valores de <i>VOT</i> expressos em segundos (s) para emissões de consoantes plosivas surdas e sonoras dos indivíduos do grupo estudado (i1 a i4) e do indivíduo referência (iR)	163
Tabela 2.	Valores de freqüência (em Hz) e amplitude (em dB) dos dois primeiros harmônicos (H1 e H2) da emissão da vogal [a] em posição tônica em "a arara" em três leituras pelos indivíduos do grupo estudado (i1 a i4) e do indivíduo referência (iR)	166
Tabela 3.	Valores de qüefrência (em ms) e do nível de amplitude do principal pico e do valor de f0 (em Hz) correspondente à análise cepstrum da emissão da vogal [a] em posição tônica em "a arara" em três leituras pelos indivíduos do grupo estudado (i1 a i4) e indivíduo referência (iR)	169
Tabela 4.	Valores de freqüência (Hz) e amplitude (dB) dos quatro primeiros formantes (F1, F2, F3 e F4) da emissão da vogal [a] em posição tônica em "a arara" em três leituras pelos indivíduos do grupo estudado (i1 a i4) e do indivíduo referência (iR)	170
Tabela 5.	Medidas de f0 (Hz), jitter (%), shimmer (%), coeficiente de contato (%) e índice de velocidade para três emissões da vogal central de três seqüências de emissões [aaa] e duas da vogal tônica em "a arara" para dois indivíduos do grupo estudado (i1 e i2) e indivíduo referência (iR)	184
Tabela 6.	Medidas eletroglotográficas de f0 (Hz), quociente de contato (CC em %) e acústica de tempo de subida de contato (Tsub em ms) para as três emissões da vogal central da seqüência [aaa] dos indivíduos do grupo estudado (i1 a i4) e do indivíduo referência (iR)	186
Tabela 7.	Medidas de <i>VOT</i> (em segundos) para as três emissões das sílabas tônicas dos sete vocábulos ( <b>pa</b> ra, es <b>per</b> to, <b>be</b> la, es <b>tar</b> , humil <b>da</b> de, apli <b>ca</b> da, lu <b>ga</b> res) dos indivíduos do grupo estudado (i1 a i4) e do indivíduo referência (iR) extraídas por inspeção simultânea das ondas Lx e acústica	188

#### Resumo

Este estudo aborda a qualidade vocal com base na teoria acústica da produção da fala (Fant, 1970) e no modelo fonético de descrição da qualidade vocal (Laver, 1980) com o objetivo de investigar os correlatos acústicos do sinal de fala de um grupo de indivíduos disfônicos por comprometimento de mecanismos glóticos, mais especificamente, paralisia unilateral de prega vocal ou fibrose de região glótica, enfocando a sua correspondência às esferas auditiva e fisiológica. Sua concepção decorre da necessidade de enfoque da qualidade vocal que contemple a prática clínica fonoaudiológica no âmbito das disfonias, em que a variabilidade de achados não é revelada em muitas das abordagens tradicionais de investigação do sinal sonoro. O estudo pauta-se na análise dos registros acústicos e eletroglotográficos de emissões da vogal [a] e da fala encadeada (texto padronizado), além de dados de exame laringológico de quatro falantes disfônicos do sexo feminino, na faixa etária de 51 a 72 anos e de um indivíduo referência do mesmo sexo e faixa etária, sem alterações vocais. São conjugados métodos de análise acústica (curto e longo termo), de avaliação perceptivo-auditiva (qualidade vocal e inteligibilidade de fala), com complementação de investigação fisiológica (eletroglotografia e exame laringológico). Os achados indicam a relevância dos aspectos acústicos de longo termo (espectros de longo termo) em correspondência àqueles de curto termo (estrutura harmônica, formântica e tempo de início de sonorização - VOT). Do ponto de vista perceptivo, é ressaltada a validade do uso de roteiro de avaliação da qualidade vocal com motivação fonética. Na esfera fisiológica, os dados eletroglotográficos encontram correspondência aos achados laringológicos, ambos relacionados aos indicativos acústicos. Nesse âmbito, as medidas de VOT apresentam correlação com a inteligibilidade de fala (sílabas), além de revelarem que ajustes de longo termo da qualidade vocal influenciam padrões de emissão no plano segmental, novamente correspondidos a aspectos da fisiologia laríngea (conformação da onda glótica e presença de onda mucosa). Os resultados ressaltam a validade da abordagem interpretativa e integrada da qualidade vocal na situação de disfonias, em face das correspondências entre os planos perceptivo e fisiológico, explicitadas pelo detalhamento da inspeção do sinal acústico de fala, ressaltando a colaboração entre os campos da Fonoaudiologia e das Ciências da Fala, entre elas, especialmente a Fonética.

#### **Abstract**

This study concerns the description of voice quality based on the acoustic theory of speech production (Fant, 1970) and the phonetic description of voice quality (Laver, 1980) with the objective of investigating the acoustic correlates of the speech signal concerning a group of subjects who presents dysphonia caused by glottal mechanism impairments, more specifically unilateral vocal fold paralysis and vocal fold fibrosis. It departs from the necessity of considering voice qualities in the clinical setting as part of the rehabilitation practice, regarding dysphonia. It differs from other theoretical accounts in that individual variability is taken into account. It is based on acoustic and eletroglottographic analysis of the central vowel [a] and data from connected speech (standardized text) as well as data from laringoscopic examination of four female subjects, aged from 51 to 72, and a subject, with no history of vocal impairments, taken as reference. Acoustic analysis (short and long term), perceptual evaluation (voice qualities and speech intelligibility) and physiological assessment (eletroglottography and laringoscopic examination) have been combined as procedures. Results indicate the relevance of long-term events (long-term average spectrum) and their correspondence to those of short-term (harmonic structure, formant structure and voice onset time -VOT). From the perceptual point of view, the usefulness of the phonetically motivated vocal profile has been validated. As far as the physiological aspects are concerned, data from eletroglottography has showed correspondence to the laringoscopic data, and both have been related to acoustic indices. Furthermore, VOT measures indicate there is correlation between timing and speech intelligibility (syllables) and besides that, they have shown that long-term settings have been found to influence segmental qualities. These findings are found to be in accordance with glottal physiology aspects (glottal wave configuration and mucosal wave). The results favor the validity of the interpretative and integrated account of dysphonic voice qualities in which the acoustic arena has played an intermediary role between the physiological and perceptive arenas. In our work, the areas of Speech Therapy and Speech Sciences, among them, Phonetics, are thought as connected for a shared purpose.

# Sumário

Dedicatória	
Agradecimentos	
Listas	
Resumo	
Abstract	
1.Introdução	
2. Uma incursão a modelos teóricos pos	SSÍVEIS DE APLICAÇÃO AO CAMPO DE
ESTUDOS DAS DISFONIAS	
2.1.Teoria acústica da produção da fa	la
2.1.1. Bases físico-acústicas da fa	la
2.1.2. Perspectiva teórica	
2.1.3. Justificativas para adoção d	a teoria acústica da produção da
fala no estudo das disfonias	S
2.2. Modelo de descrição fonética da	qualidade vocal
2.2.1. Perspectiva Teórica	
2.2.2. Justificativas para adoção de	o modelo de descrição fonética da
qualidade vocal no estudo d	las disfonias
2.2.3. Proposta de avaliação da qu	ualidade vocal – VPAS
3. Revisão crítica de estudos das d	ISFONIAS: CORRESPONDÊNCIA ENTRE
PARÂMETROS ACÚSTICOS, PERCEPTIVOS E	FISIOLÓGICOS
3.1.Histórico	
3.2.Perspectivas recentes	
4.CONCEPÇÃO DE ABORDAGEM INTERPRETAT	IVA E INTEGRADA DO SINAL VOCAL NA
VIGÊNCIA DE DISFONIAS	
4.1. Análise acústica de longo termo	
4.2. Roteiro de avaliação da Qualidad	e vocal com motivação fonética
4.3. Eletroglotografia (EGG)	
4.3.1. Histórico e princípios básico	S

4.3.2. Procedimentos de análise	87
5. MÉTODOS DE COLETA DE DADOS	94
5.1. Composição do grupo estudado e seleção de indivíduo referência	95
5.2. Concepção geral das etapas de coleta de amostras	98
5.2.1. Coleta das amostras de fala por meios acústico e	
eletroglotográfico (EGG)	99
5.2.2. Exame otorrinolaringológico	107
6. MÉTODOS DE EDIÇÃO, TRATAMENTO E ANÁLISE DOS DADOS DE ORDEM ACÚSTICA,	
PERCEPTIVA E FISIOLÓGICA	111
6.1. Edição e tratamento das amostras de fala	111
6.2. Procedimentos de análise dos dados	115
6.2.1. Análise acústica	115
6.2.2. Avaliação perceptivo-auditiva	121
6.3. Avaliação fisiológica	127
6.3.1. Eletroglotografia (EGG)	127
6.3.2. Exame otorrinolaringológico	130
6.4. Experimento voltado à investigação dos efeitos da composição do	
estímulo de fala na análise acústica de longo termo	131
7. ANÁLISE DOS DADOS DAS ESFERAS ACÚSTICA, PERCEPTIVA E FISIOLÓGICA EM UM	
GRUPO DE INDIVÍDUOS DISFÔNICOS	135
7.1. Análise acústica	135
7.1.1. Dimensão acústica de longo termo	135
7.1.2. Dimensão acústica de curto termo	139
7.2. Análise perceptivo-auditiva das amostras de fala	172
7.2.1. Avaliação da qualidade vocal com motivação fonética	172
7.2.2. Inteligibilidade de fala (sílabas)	175
7.3. Análise da dimensão fisiológica das amostras de fala	177
7.3.1 Eletroglotografia (EGG)	177

7.3.2. Complementação da descrição por meios fisiológicos – dados	
de exame otorrinolaringológico (telelaringoestroboscopia e	
videoquimografia	189
8. Da produção à percepção sonora: determinação da qualidade vocal de	
INDIVÍDUOS DISFÔNICOS SOB O OLHAR DA PERSPECTIVA FONÉTICA	195
9. Considerações finais	231
10. Anexos	233
11. Referências Bibliográficas	261
Bibliografia consultada	

# 1.INTRODUÇÃO

Nas últimas décadas, a análise do sinal de fala tem recebido muita atenção por parte de várias áreas de conhecimento, tais como a Fonoaudiologia, a Lingüística, a Medicina, a Engenharia, a Física, as Ciências da Computação, a Música e a Semiótica. Cada um desses campos oferece contribuições significativas para o avanço dos conhecimentos acerca da natureza, função e uso da voz nas situações de comunicação e expressão oral.

Nesse universo, o estímulo sonoro passa a ser objeto de estudos de naturezas e finalidades diversas, podendo ser enfocado segundo variados referenciais teóricos. Tais investigações carregam em comum o fato de reconhecerem a complexidade dos processos envolvidos na sonorização do fluxo aéreo pelo aparelho fonador. O interesse despertado em todos os pesquisadores referidos deriva da necessidade de se conhecer a particularidade das ações que resultam num sinal sonoro que carrega informações lingüísticas, paralingüísticas e extralingüísticas.

O mesmo fascínio e motivação impulsionaram um expressivo desenvolvimento tecnológico nas últimas décadas, o qual permitiu investigações mais detalhadas, tanto do comportamento das estruturas do aparelho fonador, quanto das propriedades perceptivas e físicas do som resultante e também levou à criação de sistemas de síntese de fala e reconhecimento de voz.

O avanço na área da Tecnologia da Fala, propiciou a adaptação de recursos de apoio variados aos indivíduos portadores de deficiências, tais como privações sensoriais (audição e visão), demais dificuldades de fala que demandam o uso de meios alternativos de comunicação. Esses estão praticamente incorporados em nosso dia-a-dia, na forma de sistemas de atendimento automático e *softwares* que respondem aos comandos da voz, entre outras aplicações disponíveis da tecnologia de telecomunicação e automação. Algumas dessas possibilidades de aplicação foram previstas por Fant (1962a), enquanto perspectivas para o desenvolvimento subseqüente às primeiras descrições acústicas da fala, das quais foi um dos contribuidores.

Nesse campo, os esforços iniciais foram desenvolvidos para o aprimoramento da análise do sinal sonoro emitido pelo falante, os quais remeteram a tentativas de decompor, de maneira mais detalhada, as qualidades do som que o ouvido capta e processa de forma tão complexa.

Tal desenvolvimento instrumental e teórico foi qualificado por Fant (1962b) como a tentativa de reprodução da totalidade das funções do ouvido humano, enquanto analisador sonoro altamente efetivo e refinado. As ações no sentido do desenvolvimento da instrumentalização foram implementadas pelo projeto Fala Visível, do qual resultou o desenvolvimento do primeiro espectrógrafo sonoro (Koenig et al, 1946). Da seqüência de importantes contribuições de trabalhos como Flanagan (1955, 1957, 1958) e Fant (1956) emergiu a descrição da teoria acústica da produção da fala (Fant, 1970).

Impulsionados pela referida capacidade de proceder a avaliação auditiva refinada dos sons produzidos pelo aparelho fonador humano, os pesquisadores, envolvidos no campo de estudos da qualidade vocal do ponto de vista clínico, desenvolveram propostas de avaliação baseadas na percepção auditiva, com a finalidade de detectar e qualificar as distorções impostas pelas situações ditas desviantes em termos de produção ou recepção de fala. Conseqüentemente, os avanços residiram inicialmente nas possibilidades de se registrar impressões perceptivo-auditivas. Progressivamente, como forma de validação das impressões auditivas registradas, passaram a incorporar aquelas informações de ordem acústica, e finalmente, a complementação por descrições fisiológicas do referido processo.

Transpondo tais avanços para a situação clínica junto a indivíduos portadores de disfonia, deparamo-nos com um panorama no qual manifestações muito semelhantes nas estruturas do aparelho fonador podem promover produtos de qualidade sonora variáveis. Classicamente, o fonoaudiólogo procura dimensionar o impacto de mudanças do estímulo sonoro para os indivíduos disfônicos, porém há pouco investimento no sentido de desvendar as informações que o próprio sinal acústico carrega, principalmente quando se consideram os efeitos desagradáveis gerados em termos de sensação auditiva.

O investimento em uma reflexão sobre a dimensão da alteração do sinal vocal, entretanto, faz-se necessário, e requer uma mudança de conceitualização do que é descrito como alteração. A alteração do sinal vocal não deve ser encarada simplesmente como um desvio em relação à situação idealizada de qualidade vocal, mas considerada em relação ao trabalho individual de implementação de ajustes que os falantes realizam na tentativa de superar os comprometimentos.

Nesse sentido, a motivação para composição do presente estudo emergiu da tentativa de compreensão da individualidade das manifestações sonoras em casos de disfonia. Para realizá-lo, contemplamos formas de análise que a Tecnologia da Fala nos oferece, mais especificamente a análise espectrográfica e eletroglotográfica, e refletimos sobre certos desdobramentos do conhecimento teórico construído pelas chamadas Ciências da Fala, entre elas, especialmente, a Fonética.

É a partir do olhar da Fonética que consideramos os distúrbios vocais. Nosso trabalho pode assim ser caracterizado como concernente ao campo de estudos da Fonética Clínica, numa tentativa de adotar a fundamentação do referido campo de estudos para interpretar as manifestações clínicas na produção sonora, visando o refinamento de nossa atuação junto a indivíduos disfônicos.

O grande desafio que impera nesse campo de estudos, quando a qualidade vocal é enfocada, reside na tentativa de corresponder as esferas relevantes da estrutura acústica do sinal ao atributo perceptivo-auditivo. As limitações residem no fato de que a qualidade vocal engloba várias dimensões do sinal, de forma que os achados de correlação entre um determinado parâmetro acústico e um respectivo julgamento perceptivo podem não sinalizar uma relação estreita. Flanagan (1957), um dos precursores de estudos da acústica da fala e suas relações com a percepção, apontava a dificuldade de se fazer corresponder uma medida acústica à qualidade auditiva, diante do desconhecimento de outros fatores que poderiam colaborar para o efeito sonoro detectado. Nesse sentido, argumentou que a percepção da fala corresponderia a um julgamento mais absoluto do estímulo acústico, do que a uma discriminação diferencial.

A exploração das várias vertentes da qualidade vocal e de seus desvios tem nos conduzido, enquanto clínicos e pesquisadores, a novos desafios na

tentativa de integrar as informações captadas pelo ouvido (esfera perceptiva) aos ajustes das estruturas do aparelho fonador durante a emissão sonora (esfera fisiológica), correspondência esta possível de ser interpretada a partir do nível acústico, considerado em seu relevante papel integrador entre os níveis auditivo e fisiológico (Kent, Read, 1992).

Diante desse universo de possibilidades e limitações, em que a exploração do sinal acústico da fala emerge como um campo promissor de estudos e, ao mesmo tempo, como objeto complexo em termos de registro, processamento e avaliação, este estudo debruça-se sobre emissões de indivíduos portadores de alterações vocais, numa tentativa de descrever a natureza dos fenômenos subjacentes à disfonia, que possam trazer uma real contribuição aos processos de diagnóstico e de reabilitação. Para tanto, a disfonia deixa de ser considerada como um desvio numérico em relação a uma situação supostamente irreal: a qualidade vocal ideal (parâmetro de normalidade) e passa a ser enfocada enquanto produção sonora decorrente de alterações orgânicas e/ou funcionais em toda a extensão do aparelho fonador, cobrindo, inclusive, os eventos de ordem subglótica e estendendo-se pelas ocorrências glóticas e supraglóticas.

Em termos de mecanismos glóticos foram consideradas as alterações de controle de fluxo aéreo por falhas de fechamento e de rigidez, além de prejuízos de mobilidade das prega vocais por desequilíbrio de tensão e massa (Dejonckere, Wienecke, 1994; Isshiki, 1999a). Segundo Isshiki (1999a), o fechamento imperfeito da glote estaria presente na fonação de indivíduos com disfonia, independente de sua causa. A falha de fechamento causaria a modificação na dinâmica do fluxo de ar pelo aparelho fonador, tornando-o turbulento em lugar de permanecer laminar. As alterações no processo vibratório decorreriam do impacto da pressão de ar anteriormente descrita em superfície alterada, tais como aquela decorrente de paralisia ou fibrose de prega vocal, condições eleitas para este estudo.

Vale ressaltar que a caracterização da produção vocal na ausência de disfonia prevê irregularidades ou inconstâncias inerentes à fala humana, vinculadas à sua qualidade de sinal aproximadamente periódico, numa alusão às possibilidades de variação do comportamento vibratório de pregas vocais. Além disso, há que se apontar a vasta gama de mobilizações de estruturas supraglóticas e sua complexa combinação

na caracterização das qualidades auditivas tradicionalmente abordadas enquanto mecanismos de ressonância e articulação na avaliação vocal clínica realizada por parte dos fonoaudiólogos.

A observação de grupos de indivíduos portadores de disfonias traz à tona muitas especificidades, que decorrem dos variados ajustes implementados por esses portadores a fim de superar alterações de ordem anatômica e funcional para cumprir a meta final de sonorização da coluna aérea expiratória, demonstrando a fascinante condição de plasticidade do aparelho fonador humano.

Este estudo propõe-se a investigar a qualidade vocal em seus aspectos fonéticos (acústicos, perceptivos e fisiológicos) como forma de traçar e delinear contribuições para futuras incursões no campo das refinadas e complexas relações entre percepção e produção do sinal vocal e, especialmente, que resultem em algum refinamento de nossa atuação junto a indivíduos que nos procuram por apresentarem a queixa que se traduz em alguma limitação do ponto de vista de seu desempenho enquanto falante: a disfonia. Outra possível contribuição de um trabalho desta natureza refere-se ao respaldo teórico para nortear o mapeamento das informações, em termos de inspeção das derivações acústica e eletrogolotográfica das amostras de fala.

Nesse sentido, não almejamos alcançar generalizações sobre a produção sonora na situação de disfonia. Pretendemos, em contraposição, ilustrar as várias facetas da qualidade vocal presentes num grupo de indivíduos em atendimento clínico por disfonia de fator desencadeante comum, de forma a explorar os fatores que permeiam a interface da complexa relação entre a acústica, a percepção e a fisiologia da produção vocal num grupo de falantes, propiciando à Fonoaudiologia um ensaio teórico para a reflexão das manifestações da ordem da vocalização humana.

Não intentamos aqui um estudo quantitativo baseado em análise de limitados parâmetros e amplo número de falantes. Tais estudos são necessários, embora, muitos deles caracterizem-se por derivar números e não, propriamente, medidas, quer por desconhecimento da complexidade da análise acústica, quer pelos desafios que essa complexidade impõe ao pesquisador.

Com um estudo de forte embasamento teórico, os achados são correlacionados entre si, como forma de permitir uma argumentação mais consistente de um exercício que é praticado por muitos clínicos: identificar as dimensões auditivas salientes do sinal de fala e interpretá-lo como efeitos das diversas mobilizações no aparelho fonador que afetam as situações de comunicação oral.

Apesar de nosso avanço no conhecimento da atividade de diversos segmentos do aparelho fonador e na sua complexa ação, enquanto compartilhamento com funções vitais (respiração e deglutição) exercidas na região de cabeça e pescoço, carecemos de informações sobre a combinação e a sincronização dos eventos glóticos e supraglóticos. Se assim pudéssemos refletir sobre as mobilizações que se encontram subjacentes à produção sonora, poderíamos progressivamente avançar para a abordagem de seu entrelaçamento e sucessão no tempo e os efeitos que produzem, enquanto produtos sonoros resultantes da combinação de cada uma das ações isoladas, numa apropriação do universo sonoro que permeia a qualidade vocal.

Para corresponder ao desafio apresentado, a abordagem acústica do sinal surge como possibilidade de avanço, porém carrega a demanda de exploração da consistência teórico-acústica da produção da fala. Nesse sentido, o modelo fonte-filtro descrito por Fant (1970) possibilita a decomposição do sinal de fala, revertendo aos vários segmentos do aparelho fonador, que colaboram para a qualidade vocal final, a saber, pregas vocais e trato supraglótico. Tal possibilidade de abordagem pode ser ampliada para a dimensão do julgamento auditivo da qualidade vocal no modelo fonético descrito por Laver (1980), o qual remete justamente ao universo das mobilizações que ocorrem simultaneamente em níveis glótico e supraglótico, para caracterizar a qualidade vocal.

Como forma de contemplar a reflexão sobre as características acústicas, perceptivas e fisiológicas dos fenômenos subjacentes às disfonias e às possibilidades de aplicação dos instrumentos de avaliação vocal no ambiente clínico, o presente estudo apresenta a proposta de enfoque da qualidade vocal por meio de modelos teóricos (modelo fonte-filtro na teoria acústica da produção da fala e modelo fonético de descrição da qualidade vocal) com o objetivo de investigar os correlatos acústicos do sinal de fala de um grupo de indivíduos portadores de disfonia por comprometimento de mecanismos glóticos, mais especificamente, por paralisia

unilateral de prega vocal ou fibrose de região glótica, enfocando a sua correspondência às esferas auditiva e fisiológica.

Estudos desta natureza podem colaborar para ampliar nossa compreensão sobre as ações da laringe e da porção supraglótica do aparelho fonador com respeito à produção vocal e para levantar novos indicativos a serem abordados na reabilitação dos impedimentos vocais. Pouca atenção tem sido dispensada à tentativa de abordagem da situação de limitação da produção vocal enquanto reveladora do real potencial de atividade do aparelho fonador humano e de sua plasticidade.

O grupo de indivíduos portadores de alterações de mecanismos glóticos foi selecionado justamente por representar a dimensão fisiológica, ou mais precisamente da fisiopatologia, exposta no conceito de disfonia empregado. Nesse sentido, caracteriza a possibilidade de estabelecimento de grande variedade de compensações, as quais não se restringem à porção glótica do aparelho fonador, ou seja, ao nível que concentra a principal limitação à produção sonora.

Outra justificativa para a condução de estudo desta natureza centrase no fato de que as medidas instrumentais da fala revelam detalhe fonético que é específico de língua e apontam para a existência de forte associação entre aspectos físicos e lingüísticos na fala (Kohler, 1995). Essa concepção sustenta a necessidade de caracterizarmos um estudo em nossa língua.

Diante das considerações anteriores e do referencial teórico adotado, não pretendemos enfocar a dicotomia entre normalidade e alteração do sinal sonoro, de forma que as manifestações de disfonia não serão discutidas em comparação a graus de desvio numérico de medidas isoladas a partir de um padrão de normalidade.

Vale ressaltar que a referência à fonação na ausência de alterações do aparelho fonador passa a ser o ponto de apoio necessário para que se considere, dentre os ajustes, aqueles que são impostos pela mudança de atividade em decorrência de incompetência glótica, subseqüente à instalação de paralisia unilateral ou de fibrose de prega vocal. Tal consideração é reforçada por existirem variadas condições que podem impor mudança de atividade no contexto de alteração do

aparelho fonador, como o que aqui consideramos (limitação da ação glótica), ou no contexto da expressividade (para indicar atitudes e emoções).

Além da importância da adoção de referencial teórico a respaldar a descrição dos dados de avaliação vocal, este trabalho caracteriza-se por ressaltar a importância da interdisciplinaridade, tanto na rotina de avaliação vocal, quanto na prática de reabilitação, uma vez que os fundamentos abordados são compartilhados pelas várias áreas de conhecimento inicialmente expostas.

Nesse sentido, a contribuição à interface dos vários campos de conhecimento é traduzida nesta apresentação no capítulo subseqüente de fundamentação teórica, com uma incursão a modelos teóricos possíveis de aplicação ao campo de estudos da voz, incluindo o detalhamento sobre aspectos acústicos e perceptivos da qualidade vocal, os quais permeiam os métodos adotados e a discussão dos achados. Na seqüência, encontramos a revisão crítica da investigação das alterações vocais. Tal exposição sustenta a descrição de métodos adotados, incluindo a exposição de propostas de avaliação da qualidade vocal por meios acústicos (análise de longo termo), perceptivo (roteiro de avaliação perceptiva da qualidade vocal com motivação fonética), e fisiológicos (eletroglotografia). Finalmente, são expostos os tópicos referentes à discussão dos achados das referidas esferas de abordagem do sinal sonoro emanado pela boca do falante portador de disfonia de características anteriormente delineadas.

Este estudo integra a Linha de Pesquisa Linguagem e Patologias da Linguagem do Departamento de Lingüística da Faculdade de Comunicação e Filosofia (COMFIL) e do Programa de Estudos Pós Graduados em Lingüística Aplicada e Estudos da Linguagem (LAEL), desenvolvido no Laboratório Integrado de Análise Acústica (LIAAC) da Pontifícia Universidade Católica de São Paulo (PUC/SP).

## 2. UMA INCURSÃO A MODELOS TEÓRICOS POSSÍVEIS DE APLICAÇÃO NO CAMPO DE ESTUDOS DAS DISFONIAS

O presente estudo associa a descrição acústica do sinal de fala à dimensão auditiva da qualidade vocal, complementada por aspectos da fisiologia, ou mais precisamente da fisiopatologia da produção vocal. Diante dessa proposta, tornase fundamental a compreensão da possibilidade de articulação dos fundamentos de tais esferas, contemplada por modelos teóricos, os quais são expostos de maneira mais detalhada no texto, justamente por representarem o referencial que permeia nossa incursão no estudo das disfonias. Nesse contexto, destacam-se a teoria acústica da produção da fala, especialmente a descrição do modelo fonte-filtro para a produção das vogais (Fant, 1970) e o modelo fonético de descrição da qualidade vocal (Laver, 1980). Ao longo da exposição teórica, da apresentação dos resultados e discussão, a articulação de ambas propostas teóricas será explicitada. Além disso, tal abordagem provê argumentos para a reflexão e a discussão a respeito da condição adaptativa do aparelho fonador diante da situação de impedimento de função, condição essa que constitui a base do detalhamento da investigação do sinal de fala em face das manifestações de desvios de qualidade vocal.

## 2.1. Teoria acústica da produção da fala

A abordagem da acústica da produção da fala (Fant, 1970) prevê, numa primeira etapa, a consideração das bases físico-acústicas da fala, as quais são articuladas na proposta teórica em questão, apresentada de forma mais detalhada no segundo tópico desta seção e finalizando com a exposição das justificativas para sua adoção no estudo das disfonias. A divisão adotada no texto procurou favorecer nossa incursão aos aspectos físicos do sinal de fala e aos níveis de atividade no trato vocal, aos quais se remetem.

#### 2.1.1. Bases físico-acústicas da fala

A consideração das bases físico-acústicas da fala envolve o detalhamento do processo de geração sonora pelo aparelho fonador humano, ou seja,

das ações dos órgãos que o constituem e os efeitos que essas ações produzem, em face dos fatores de ordem aerodinâmica, determinados pelas variáveis de fluxo e pressão ao longo do aparelho fonador e dos aspectos da psicofísica da audição. Também fundamenta o estudo das representações lingüísticas dos sons da fala (vogais e consoantes) e a análise das influências do contexto fonético (os fenômenos de coarticulação e os efeitos da prosódia nos segmentos).

As bases físico-acústicas da fala são amplamente abordadas em obras destinadas a contribuir para a formação em Ciências da Fala, entre os quais destacamos Flanagan (1972), Fry (1979), Kent, Read (1992), Borden et al (1994), Ladefoged (1996), Kent (1997), Stevens (1998) e Hayward (2000).

Para a compreensão específica dos conceitos básicos da teoria acústica, desenvolvidos ao longo de séculos de pesquisas em acústica da fala e, especialmente, daqueles que colaboram diretamente para o enfoque das disfonias, destacam-se a geração e a propagação das oscilações no aparelho fonador. Nesse universo, figura o estudo das representações das emissões por meio de ondas sonoras, de sua propagação em diversos meios, inclusive naqueles semelhantes ao aparelho fonador, como é o caso das aproximações ao tubo uniforme com uma das extremidades fechada, além de demais efeitos de transmissão sonora, envolvendo especialmente a noção de ressonância, a qual, por sua vez, depende do comprimento de onda ( $\lambda$ ), da amplitude e da freqüência de oscilações, além de sua propagação. Tais aspectos básicos são vitais para que o profissional fonoaudiólogo aprofunde seu conhecimento a respeito da teoria acústica da produção da fala e estão explicitados nas publicações anteriormente referidas que exploram os conceitos da Fonética Acústica e das Ciências da Fala.

Fant (1962a) salienta a importância do desenvolvimento, ocorrido a partir do final da década de 40 em relação aos estudos da fala, e assinala o considerável avanço da Fonética Acústica, suplantando as abordagens de base articulatória ou fisiológica. Neste ponto cabe menção especial ao fato de que a abordagem acústica do sinal vocal em nosso meio viria a se desenvolver décadas adiante, apesar de alguns estudos clássicos acompanharem a evolução do estudo da acústica da produção de fala (Yanagihara, 1967; Rontal et al, 1975), os quais serão

comentados de forma mais detalhada em tópico específico destinado aos estudos das disfonias.

Como primeiro nível de atividade a diferenciar a produção sonora, grande atenção é dispensada ao detalhado estudo da atividade glótica, que envolve a consideração dos efeitos da pressão subglótica e de sua modulação, pressão essa em parte garantida pela atividade refinada de pregas vocais. Para os sons que apresentam exclusivamente a fonte sonora decorrente de atividade vibratória de pregas vocais, a periodicidade fundamental é estabelecida pela atividade valvular da laringe, que pode ser considerada um sistema mecânico, cuja atividade é regulada por sutis mudanças em termos de tensão, massa e comprimento de pregas vocais, interagindo com a pressão subglótica e resultando nas variadas configurações de laringe à fonação (Lieberman, 1963; Fry, 1979; Isshiki, 1999b). Pela possibilidade de oscilação em intervalos regulares de tempo e graças à flexibilidade dos tecidos, a onda sonora resultante deste nível de atividade pode ser descrita como onda periódica complexa.

Do ponto de vista fonético, a decomposição da onda complexa anteriormente referida resulta num espectro harmônico. O termo harmônico deveria estar restrito aos sons produzidos exclusivamente pela fonte sonora das pregas vocais. A associação entre os termos harmônico e periódico não é totalmente aplicável à produção vocal, uma vez que a variação sempre está presente. Nesse sentido, inicialmente foi proposta a referência aos sons sonoros ressoantes enquanto quase periódicos (Timcke et al, 1958; Stevens, House, 1961; Lieberman, 1963; Hollien et al, 1973; Davis, 1979; Kent, Read, 1992; Gould, Korovin, 1994), termo comum nas referências à fisiologia das pregas vocais e nos aspectos acústicos, apesar de restrições à terminologia propostas por Titze (1995) e reforçadas por Baken, Orlikoff (2000) de que a quase-periodicidade envolve um conceito de aperiodicidade, sendo considerada como mais apropriada a referência à superposição de dois ou mais sinais periódicos com relação de razão baseada em número não-inteiro. Esses dois últimos autores usam distintivamente os termos "aproximadamente periódico" e "quase-periódico", aplicando o segundo a situações de alteração de produção vocal.

Visando aprofundar os conhecimentos de aspectos de fisiologia e fisiopatologia da voz, vale recordar que a mucosa da prega vocal perde sua mobilidade quando tem cicatriz, está estirada ou edemaciada. Dessa forma, os fatores referentes a

aderências ao tecido submucoso, propriedades reológicas e redundância da área de mucosa interferem em tais mobilizações (Isshiki, 1999a). A fonação ocorre somente dentro de alguns parâmetros relacionados: área glótica inicial (Ag0), correspondente à área glótica sem fluxo de ar subglótico, além de outros fatores, tais como a pressão subglótica, a rigidez das pregas vocais e a coaptação entre os lábios superior e inferior das bordas das pregas vocais (Isshiki, 1999a) e a tensão longitudinal das pregas vocais (Kent, 1997; Vieira, 1997).

Para o conhecimento mais detalhado de tais mecanismos, vale recordar que, de maneira geral, o fluxo de ar é classificado em laminar ou turbulento. Quando laminar, suas linhas de escoamento estão paralelas umas às outras em todos os pontos, enquanto no fluxo turbulento essas linhas se curvam e criam uma espécie de vórtice, resultando em ruído, característico de muitos casos de disfonia.

Há um índice capaz de determinar se o fluxo de ar tende a ser laminar ou turbulento: o número de Reynolds também aplicável aos estudos de pregas vocais (Flanagan, 1958; Fant, 1970; Kent, Read, 1992; Kent, 1997; Stevens, 1998; Isshiki, 1999a). Se, por um lado, a atividade decorrente da vibração periódica de pregas vocais gera um som aproximadamente periódico, o mecanismo acima descrito favorece o registro de aperiodicidade da onda complexa de fala, substituindo a tendência original de som com predominância de estrutura harmônica.

Nesse ponto de exposição de conceitos, as considerações sobre a natureza periódica e aperiódica dos sons da fala tornam-se inevitáveis, independente da manifestação de alterações do aparelho fonador. O termo fonte de voz passou a ser adotado na teoria acústica como referência à categoria da fonte, enquanto o indicativo da sonoridade dos sons da fala. O espectro harmônico, enquanto correlato acústico da descrição fisiológica da atividade vibratória de pregas vocais, é a energia que ativa os formantes. Não apenas fontes de voz são formadas no aparelho fonador, também fontes de ruído são geradas por liberação de bloqueios totais ou passagem da corrente de ar por estreitamentos entre os articuladores. Estes tipos de fontes serão abordadas de forma mais direta no próximo item desta seção.

Na seqüência de abordagem da atividade do aparelho fonador na produção sonora, destaca-se a importância das estruturas supraglóticas, cujas propriedades dos tecidos revelam, no geral, as características de ressonância do sistema. Nesse ponto, destaca-se a propensão a alto grau de absorção de energia (Fry, 1979), pelo fato de muitas paredes serem revestidas por tecidos moles. Dessa forma, sua curva de ressonância poderia ser pouco seletiva. Além do predomínio de elementos não rígidos, mais moles e absorventes, o trato vocal não mantém um diâmetro absolutamente constante (Borden et al, 1994), levando a apresentar várias freqüências de ressonância, conhecidas como formantes. Tais respostas de ressonância são basicamente dependentes do volume do ressoador, quanto menor o seu volume, mais aguda será a freqüência de resposta de ressonância e vice-versa.

Dada a complexidade de se modelar o efeito de ressonância no trato vocal, a teoria acústica voltou-se à modelagem do tubo retilíneo com uma das extremidades fechada, que sinaliza as pregas vocais em atividade, e outra aberta, que se refere ao limite final do aparelho fonador (os lábios). Dessa forma, as ressonâncias irão variar de acordo com a ocorrência das constrições em pontos de máxima velocidade ou de máxima pressão do trato vocal, recordando-se que em pontos de velocidade máxima, a pressão é mínima e vice-versa (Fry, 1979; Kent, Read, 1992; Ladefoged, 1996). Como tendência geral para tubos modelados à semelhança do trato vocal, quando as constrições ocorrem em ponto de velocidade máxima, há a tendência à diminuição da freqüência de ressonância, enquanto a sua ocorrência em pontos de pressão máxima tende a elevá-la (Borden et al, 1994).

A aproximação do trato vocal a um tubo cria condições de previsão de suas condições de ressonância. Nessa concepção, para um tubo de extensão l (expressa em cm), as freqüências dos formantes podem ser calculadas pela seguinte fórmula:

na qual Fn equivale ao número do formante investigado: primeiro (F1), segundo (F2), terceiro (F3) e assim sucessivamente. A expressão 2n - 1 é a combinação que gera os múltiplos ímpares do comprimento da onda, vinculada aos estudos que ressaltam a relação de ressonância de quarto de onda, em que o tubo (trato vocal) ressoará ao máximo a freqüência cujo comprimento de onda ( $\lambda$ ) for equivalente a quatro vezes a sua extensão. Freqüências múltiplas ímpares inteiras deste comprimento de onda também são ressoadas, de forma que ocorrerão ressonâncias em 1/4; 3/4; 5/4; 7/4, e

assim por diante, da extensão do tubo: a primeira (F1), segunda (F2), terceira (F3) e quarta (F4) ressonâncias, e assim sucessivamente, conhecidas como as ressonâncias de quarto de onda, representadas pela expressão  $\lambda$ /4 (Kent, Read 1992; Borden et al, 1994; Kent, 1997). A designação c representa a velocidade do som, aproximadamente 35000 cm/s (Kent, Read, 1992; Ladefoged, 1996; Kent, 1997).

Essa fórmula pode ser usada para prever as freqüências dos formantes de uma vogal na qual a língua está em posição média e central [a]. Nessa situação, ao assumirmos a extensão do trato vocal masculino em 17,5 cm (distância presumível das pregas vocais até os lábios) teremos a freqüência de F1 estimada em 500 Hz, F2 em 1500 Hz e F3 em 2500 Hz, com intervalos regulares de 1000 Hz para as demais ressonâncias. O tamanho do tubo adotado representa uma idealização próxima à média para facilitar os cálculos (Ladefoged, 1996).

Vale ressaltar, como consideração prévia dos aspectos das descrições acústicas dos sons da fala, que as vogais passaram a ser estudadas neste enfoque na década de 40 do século passado, inspirados em estudos do século XIX, nos quais a produção da vogal [i] foi relacionada a um ressoador simples e de [a] e [u] a ressoadores duplos, um anterior e um posterior (Borden et al, 1994).

Os cálculos de ressonância para os sons ressoantes encontram respaldo nos conceitos anteriormente abordados e representam estimativas relacionadas ao levantamento de estudos de população de falantes de diversas línguas. Nesse universo, destaca-se a atenção aos estudos das ressonâncias do trato vocal, ou mais precisamente, aos formantes das vogais. Justamente pela possibilidade de definição de traços de ordem segmental. Além disso, devem ser consideradas as variações das dimensões relativas ao trato vocal para homens (17,5 cm), mulheres (15 cm) e crianças (8 a 10cm), ocasionando variações nas referidas freqüências de ressonância, levando os falantes do sexo masculino a apresentar freqüências de ressonância cerca de 17% mais baixas do que os do sexo feminino. Os valores referentes às freqüências de formantes de crianças não foram estimados na época (Fant, 1956).

Ladefoged (1996) oferece ao leitor a possibilidade de compreensão dos complexos mecanismos que levam à ressonância, enquanto evidências físico-

acústicas no aparelho fonador, descrevendo as possibilidades de mobilização com base em várias vogais do inglês e complementadas pela descrição de Fant (1973) para as vogais suecas.

A dimensão da percepção auditiva humana também colaborou e ainda colabora para o crescimento dos estudos do sinal de fala e aparece referida em propostas originais de estudos de sons de fala, como é o caso de Flanagan (1955, 1957) quanto ao julgamento perceptivo de estímulos de vogais sintéticas e manipulação de suas duas primeiras ressonâncias.

Na primeira obra, Flanagan (1955) revelou a necessidade de alta resolução na análise e estimativas dos formantes das vogais, especialmente com relação à sua freqüência. A partir da plotagem das áreas de F1 e F2 foi possível identificar os segmentos vocálicos em questão.

Na seqüência, Flanagan (1957) estudou a precisão necessária para a transmissão de dados relativos à amplitude do segundo formante, novamente com uso de estímulos de vogais sintéticas. A variação de julgamento de amplitude não se mostrou dependente de f0. O estudo contribuiu para as pesquisas em síntese e transmissão de fala vigentes no período.

Na seqüência de retomada histórica, as décadas de 50, 60 e 70 do século passado registraram contribuições importantes no sentido do estudo de aspectos da fonte glótica (Flanagan, 1958) e de ressonâncias do trato por meio da descrição de formantes (Flanagan, 1955, 1957, 1972; Stevens, House, 1961; Fant, 1970; Lindblom, Sundberg, 1971; Fant, 1973), os quais representaram importantes contribuições para as propostas teóricas em questão.

## 2.1.2. Perspectiva teórica

Como resultado da evolução dos estudos no campo do detalhamento da produção sonora pelo aparelho fonador humano, Fant (1970) apresentou referências a teorias de sistemas de ressonância e de fontes sonoras na fala. Dessa forma, a teoria acústica da produção da fala foi descrita no início da década

de 60 e serviu de base para os estudos acústicos subseqüentes, como o registro de uma contribuição para a compreensão da relação entre dados fisiológicos do aparelho fonador e físicos da onda de fala. Em publicação anterior à primeira versão da teoria (Fant, 1956), e em outra posterior (Fant, 1962a), o autor expôs conceitos que seriam articulados nas duas publicações da obra.

A partir da combinação de conteúdos anteriormente abordados com respeito aos eventos acústicos no aparelho fonador e seus prováveis desmembramentos no campo da fisiologia e da percepção, Fant (1970) apresentou a segunda versão da obra, revisando alguns conceitos que foram interpretados de forma mais genérica ao longo do período de dez anos subseqüentes à primeira publicação e salientou que o principal objetivo foi fornecer as bases teóricas de aspectos estáticos da teoria de produção da fala. Destacou que especificações do estudo da dinâmica da fala pertenceriam ao desenvolvimento subseqüente à proposta apresentada.

Em publicação prévia, Fant (1956) descreveu a onda de variação de pressão sonora registrada a alguma distância à frente dos lábios do falante numa equação segundo a qual seria fruto da atividade da fonte, do filtro e da radiação, da seguinte forma:

sons da fala = (fonte) . (função de filtro) . (radiação)

na qual: "fonte" referia-se à transformação da onda de velocidade volumar na glote; a "função de filtro" era expressa em termos da proporção entre a velocidade volumar nos lábios e a velocidade volumar ao longo da glote e a "radiação" era considerada enquanto a taxa de transformação da pressão sonora a uma certa distância dos lábios do falante, pela transformação da velocidade volumar nos lábios. Se tais magnitudes estivessem representadas em dB, as operações de multiplicação presentes na fórmula seriam tão somente adições simples.

Apesar de pequenas variações em função da fonte sonora, a qual deixa de ser glótica no caso da maior parte dos tipos de ruído, tal esquema pode ser adotado de forma genérica para o estudo dos sons da fala.

Torna-se necessário, a partir deste ponto, distinguir duas categorias básicas de sons da fala, aqueles que apresentam apenas a fonte vibratória de pregas vocais como marca de estímulo sonoro, caso dos sons ressoantes, em contrapartida

àqueles cuja fonte básica de energia reside na criação de estreitamentos ou obstáculos à corrente de ar, caso dos sons obstruintes. Dada a importância desta classificação, os dois grupos serão abordados separadamente na continuidade do texto.

# 2.1.2.1. Os sons ressoantes: modelo fonte-filtro para a produção das vogais (Fant, 1970)

A fonte sonora é garantida pela ação de pregas vocais no caso da produção de sons vocálicos e dos demais sons consonantais ressoantes (líquidas, nasais e semivogais). Na publicação original, Fant (1970) representou o trajeto de sua produção esquematicamente por meio de uma rede, na qual S(f) indicaria fonte e T(f) a função de transferência (espectro do trato vocal e de lábios) do filtro do trato vocal. P(f) = S(f). T(f) representaria o som de fala correspondente ou produto. Assumida da maneira exposta, tal expressão é dependente da freqüência, justificando as notações (f) constantes nas fórmulas referentes aos preceitos teóricos do modelo da fonte e do filtro.

Numa releitura dos aspectos anteriormente abordados e presentes em descrições mais recentes do modelo fonte-filtro para produção de vogais, o espectro sonoro irradiado pela boca do falante é tido como o produto do espectro do fluxo de ar glótico ou a energia da fonte, da função de transferência vocal ou do(s) ressoador(es), referidos como filtros de transmissão seletiva de freqüência, e, finalmente, das características de radiação, expressas pela relação entre o fluxo de ar da fonte e a pressão sonora irradiada. Para representar essas transformações, é utilizada a seguinte fórmula:

$$P(f) = U(f).T(f).R(f)$$

em que U(f) corresponde originalmente à velocidade volumar e é usado no mesmo sentido de S(f) descrita no parágrafo anterior, T(f) corresponde à função de transferência e R(f) denota a característica de radiação dos lábios (Kent, Read, 1992; Kent, 1993; Kent, 1997; Stevens, 1998).

Enfocando cada nível de atividade exposto na produção sonora para as vogais e consoantes ressoantes, a fonte representada por S(f) ou U(f) ocorre por uma modulação quase permanente do fluxo aéreo, devido aos movimentos de abertura

e fechamento das pregas vocais (Fant, 1970). A propriedade básica da fonte sonora nas pregas vocais é a periodicidade expressa pela duração de período completo (T0) ou pelo seu valor inverso, a freqüência fundamental (f0), de forma que f0 pode ser assumida como o inverso do período e expressa na seguinte fórmula:

$$f0 = 1/T0$$
.

Esta função de tempo pode ser representada, em termos acústicos, por um espectro harmônico graças ao teorema de Fourier, o qual, quando expresso em dB, é representado com declínio linear, no qual as linhas verticais representam os múltiplos inteiros da freqüência fundamental de vibração (f0), com energia revelando decréscimo em proporção de 12 dB/oitava (Fant, 1970; Fry, 1979; Kent, Read, 1992; Kent, 1993; Ladefoged, 1996; Stevens, 1998).

Na continuidade da descrição da produção sonora pelo aparelho fonador humano, destaca-se que o sinal aproximadamente periódico gerado pela fonte laríngea sofre a função de transferência T(f) do trato vocal, que é definida basicamente pelas ressonâncias do trato (Fant, 1970), também referidas como modos naturais de vibração do trato vocal ou formantes. Esta propriedade seletiva T(f) é independente da fonte sonora e está relacionada aos efeitos de transmissão ao longo do trato vocal, decorrentes da extensão total, bem como diâmetro do tubo e são medidos pelo centro da banda do formante (Kent, Read, 1992; Borden et al, 1994; Ladefoged, 1996). Como qualquer tubo com as características anteriormente descritas, com uma extremidade fechada e outra aberta, o trato vocal na produção de vogais terá ressonâncias sob determinadas condições físicas, as quais dependem da relação entre a freqüência da fonte sonora e a extensão do tubo. Conforme exposto no item anterior, tal relação é conhecida como ressonâncias de quarto de onda ( $\lambda$ /4). Dessa maneira, a ressonância ocorre quando o tubo está ajustado ao comprimento da onda do componente de freqüência gerado (Kent, 1997).

Cada formante é caracterizado por um centro de freqüência, conhecido como freqüência do formante, identificado no espectro por meio do pico e da banda, que reflete a faixa de freqüência efetiva do ressoador (Stevens, House, 1961; Kent, Read, 1992; Kent, 1993; Ladefoged, 1996). A freqüência do pico do formante varia com a configuração do trato vocal e, portanto, depende da função de filtro. Daqui novamente depreendemos o conceito de independência de cada nível de atividade no

aparelho fonador, enquanto responsável por um efeito acústico particular, portanto numa referência à linearidade do sistema.

A largura de faixa ou banda do formante, por sua vez, representa a extensão de freqüências em torno da qual um ressoador responderá para um estímulo (Sundberg, 1987). Esta medida de faixa efetiva do ressoador é geralmente conhecida como largura de banda de meia potência, que expressa a relação entre potência e tensão (Ladefoged, 1996). Em tal relação, quando a potência cai pela metade (meia potência), a tensão declina em cerca de 70,7%. Quando referida em dB, a potência passa a ser calculada em operação logarítmica, na qual 3 dB representa o correlato da meia potência prevista para o índice percentual referido. Dessa forma, quando expressa em dB, a medida é calculada a partir do cômputo das freqüências que estão na faixa de 3 dB abaixo do pico. Se a escala não estiver em dB, deve-se adotar aproximadamente 70% da amplitude como limites para identificar as freqüências extremas da largura de faixa do formante.

Ressoadores com banda estreita respondem a uma faixa de freqüência menor, porém levam mais tempo para atingir o máximo de amplitude e a menor taxa de declínio. Ao contrário, aqueles com uma larga faixa de ressonância chegam ao máximo de amplitude mais rápido (Ladefoged, 1996). A largura da faixa do formante também é conhecida como uma medida da taxa de amortecimento no domínio temporal de forma que quanto mais larga, maior o amortecimento (Fant, 1970, 1973; Kent, Read, 1992; Kent, 1993).

A maior parte da energia da fonte está concentrada em freqüências mais graves. A razão desta ocorrência é o declínio de amplitude que ocorre no espectro da fonte, expressa na proporção de 12 dB/oitava. Grande parte da energia está concentrada nos três primeiros formantes da emissão das vogais, de tal forma que uma extensão de freqüência de 3 kHz geralmente é suficiente para representar a informação acústica primária que reflete ajustes fonéticos relevantes ao trato vocal.

Fant (1973) destacou que as freqüências de F1, F2 e F3 seriam os principais determinantes da qualidade fonética de uma vogal. A análise das variáveis articulatórias também é importante, de forma que apontou como relevantes o local e o

grau de constrição do principal estreitamento entre a língua e o palato, além do grau de constrição e alongamento da passagem labial (Fant, 1973).

Finalmente destaca-se a descrição do efeito de radiação *R(f)* dos lábios, no qual prevalece o efeito físico da difração. Para a sua compreensão, deve-se inicialmente considerar que os sons de alta freqüência tendem a difratar e, portanto, a apresentar melhores condições de propagação, projeção e reflexão, justamente por não apresentarem a tendência de contornar os obstáculos, com o é o caso dos sons de freqüências graves. Dessa forma, na extremidade final do tubo adotado como aproximação ao trato vocal ocorre uma ação de ganho de 6 dB/oitava como tendência a facilitar a propagação do som para o meio externo. Fant (1956) admitiu a possibilidade de se enfocar a radiação enquanto o terceiro fator de atividade ou como mais uma parte da função de filtro, de forma que o fenômeno complexo da difração seria a base para a compreensão da função de filtro passa-alta atribuído à função de radiação em várias publicações posteriores (Kent, Read, 1992; Ladefoged, 1996). Dessa forma, quanto mais alta a freqüência, maior é a resposta do ar circundante externo à ação da vibração do ar dentro do aparelho fonador (Ladefoged, 1996).

Fant (1970) destacou que haveria certo grau de correspondência entre os termos fonéticos fonação e articulação e a respectiva terminologia técnica de fonte e filtro. Esta analogia prevê que a fonação seja concebida, separadamente da articulação, na geração da fala. Nesse ponto, ressaltamos novamente a característica de linearidade apontada para o modelo fonte-filtro para a descrição das vogais.

Lindblom, Sundberg (1971) apresentaram um modelo articulatório para obtenção da informação das freqüências dos formantes em função da postura de lábios, da conformação e posição da língua, da altura da laringe e finalmente da posição da mandíbula. Esta última foi considerada como um novo elemento para a tarefa em questão, destacando sua cooperação com a língua e concluindo que sua altura deveria ser considerada como um traço da configuração final do trato vocal. Neste ponto, procederam uma reinterpretação para o conceito clássico de altura de língua. As conseqüências acústicas dos variados graus de mobilização de tais estruturas foram calculadas e passarão a ser apresentadas em conjunto com as observações dos demais autores no que se refere à freqüência dos formantes.

Mantidas as considerações anteriores, os elementos que interferem de forma mais significativa e recorrente na freqüência de cada um dos quatro primeiros formantes (F1, F2, F3 e F4) passam a ser discutidos. Todas as estruturas do trato vocal exercem influência sobre o padrão de formantes, com graus variados dependendo de sua configuração (Fant, 1970, 1973). A extensão do trato vocal seria um dos determinantes da freqüência de formantes, de forma que o seu alongamento tenderia a abaixar todas as freqüências dos formantes, e a diminuição, a aumentá-las (Stevens, House, 1961; Fant, 1970; Laver, 1980; Kent, Read, 1992). O fato de se assumir que a extensão total do trato (relacionada à idade e sexo) afeta os formantes cria um problema básico na Fonética Acústica, relativo à compreensão de como o falante atinge a mesma classificação fonética para vogais que não partilham das mesmas propriedades acústicas. Kent (1997) salienta que este seria um problema que implicaria a normalização do trato vocal. Algumas soluções propostas apontam para transformações não lineares dos valores de formantes. Um exemplo deste tipo de transformação é a escala Bark (Syrdal, Gopal, 1986).

Considerando que existe uma constrição de língua no trato vocal, delimitando uma cavidade posterior e uma cavidade anterior, existem condições primárias, porém não suficientes, para descrever a configuração do trato vocal como duplo ressoador. Fant (1970) salientou que um uso amplo do modelo foi realizado na literatura e que a proposta geral de que o primeiro formante (F1) estaria relacionado à cavidade posterior e o segundo (F2) à cavidade anterior não seria suficiente para todas as situações. Destacou que um de seus objetivos seria mostrar as limitações que esta regra impõe e acrescentar fundamentos fisiológicos para a interpretação das ressonâncias do trato vocal ao se considerar que várias ações ocorrem simultaneamente e podem influenciar tais ressonâncias.

Fant (1970), além das descrições das bases fonético-acústicas da teoria acústica da produção da fala, baseou-se no modelo de Stevens, House (1961) para a caracterização dos formantes das vogais, considerando o ponto de constrição da língua, o grau de protrusão labial e o diâmetro do trato vocal.

Levando-se em conta a necessidade de detalhamento das ações que interferem nos valores de freqüência dos formantes, destaca-se um estreitamento da articulação geralmente ocasiona declínio de F1. Dessa maneira, F1 apresenta

variação com a altura de língua (Stevens; House, 1961) e a abertura de boca (Stevens, House, 1961; Lindblom, Sundberg, 1971; Laver, 1980; Sundberg, 1987).

A elevação da língua tem uma relação direta com o aumento da cavidade posterior, ou seja, da cavidade faríngea. Conseqüentemente, vogais altas tenderiam a apresentar cavidade faríngea ampla (Fant, 1970, 1973; Borden et al, 1994; Ladefoged, 1996). Os mesmos autores também destacam que F1 seria muito influenciado pela abertura de boca, fato este anteriormente detalhado por Lindblom, Sundberg (1971) que, ao estudarem os efeitos de mandíbula juntamente a quatro conformações diferentes de língua, encontraram a tendência de aumento da freqüência de F1 proporcionalmente ao aumento de abertura mandibular. Tais aspectos foram reforçados por Sundberg (1987) em publicação voltada ao estudo da voz cantada. Além disso, a massa das paredes também exerce papel na variação de F1, representando a ressonância que mais contribui para a intensidade total do som, especialmente quando comparado aos demais formantes (Fant, 1970; Kent, Read, 1992).

A altura de língua também parece influenciar mecanismos relativos à atividade da fonte, especificamente a freqüência fundamental, uma vez que vogais altas tendem a apresentar valores de f0 mais altos, comparativamente às vogais baixas (Kent, 1997).

A segunda ressonância do trato vocal, F2, estaria relacionada ao grau de variação do posicionamento da língua no sentido ântero-posterior (Stevens, House, 1961; Lindblom, Sundberg, 1971) e à cavidade de ressonância anterior à constrição de língua (Fant, 1970, 1973; Ladefoged, 1996), desde que assinalada a tendência à diminuição desse efeito à medida em que a mandíbula abaixa (Lindblom, Sundberg, 1971), quando é referida certa constrição em faringe.

Para Borden et al (1994), F2 é mais afetado pelas mudanças na cavidade oral, especialmente a sua extensão. Seu aumento é proporcional à área de constrição, desde que o centro de constrição esteja mais próximo da glote de que dos lábios (Fant, 1970). Para Lindblom, Sundberg (1971) o grau de constrição, definido no modelo articulatório com relação à ação de mandíbula, modifica basicamente F2.

Ladefoged (1996) descreveu a associação de F3 como uma ressonância do restante do trato vocal. Na medida em que a constrição se move para trás e a extensão da cavidade anterior aumenta, F3 tende a diminuir.

O terceiro formante, F3, está relacionado à passagem da constrição, ou seja, à cavidade atrás da constrição de língua e na frente dela (Fant, 1970, 1973). Lindblom, Sundberg (1971) destacaram que o arredondamento labial afeta todos os formantes em termos de diminuição de freqüência, com efeitos mais marcantes de ajustes de lábios especialmente para F2 e F3. Dessa forma, os três primeiros formantes (F1, F2 e F3) têm suas freqüências rebaixadas com constrição labial e aumentadas com constrição próxima à laringe (Kent, Read, 1992).

Do ponto de vista fonético, F4 não teria despertado interesse inicial para estudos e poderia ser influenciado em maior ou menor grau pelo tubo laríngeo (Fant, 1970). Por outro lado, Sundberg (1974) descreve seus efeitos na qualidade do sinal sonoro e destaca a influência da extensão do trato vocal e da dimensão da laringe em comparação às áreas circunvizinhas, especialmente a hipofaringe. Se a última for ampla, F4 tende a depender exclusivamente do vestíbulo laríngeo, de maneira que se estiver expandido, F4 diminui. Ao contrário, se a abertura do tubo laríngeo não for mais estreita que a dimensão faríngea, mostra-se dependente da extensão do trato vocal e da configuração do tubo laríngeo (Sundberg, 1974, 1987).

Neste tocante, Sundberg (1974, 1987) descreve conformações dos espaços laríngeo e faríngeo para uma ressonância de particular interesse no enfoque de cantores profissionais do sexo masculino, caracterizado por um pico de concentração espectral entre as freqüências de 2,5 a 3 kHz, correspondente a F3, F4 e F5. Estudos posteriores demonstraram inclusive que tal faixa pode variar de acordo com a extensão vocal dos cantores (Sundberg, 1999).

De acordo com a descrição de Sundberg (1974), F4 estaria relacionado articulatoriamente ao formato da laringe e da faringe na mesma altura. Dessa forma, aparece como um formante extra, o qual aumenta a capacidade de transferência dos componentes do sinal ao redor de sua freqüência. O termo formante do cantor é geralmente usado para se referir ao pico de freqüência, embora Sundberg (1987) aponte a imprecisão desse uso.

Lindblom, Sundberg (1971) relataram efeitos de diminuição de freqüência de todos os formantes para a posição de laringe baixa, com efeito mais acentuado em F2.

Com relação aos formantes, cabem comentários adicionais a respeito da vinculação entre amplitudes e freqüências. Se a freqüência de F1 abaixar, os formantes mais altos sofrerão declínio em amplitude e, ao contrário, se aumentar, os formantes mais altos sofrerão aumento de amplitude. Se dois formantes aproximaremse em freqüência, ambos os picos aumentarão em amplitude (Kent, Read, 1992).

Essa relação de dependência é uma característica de ressoadores que estão conectados em série, de forma que o sinal de saída de um ressoador é o estímulo para o seguinte, como apontou Pickett (1980). Tal característica é decorrente do declínio espectral, o qual será influenciado pelo tipo de voz e do esforço vocal no sinal sonoro. Segundo o mesmo autor, uma redução do esforço vocal leva a um decréscimo do nível de harmônicos, ocasionando declínio espectral mais abrupto, estimado em 15 dB/oitava, em contrapartida ao declínio espectral menos acentuado que ocorre ao se aumentar o nível de intensidade de voz por esforço vocal, previsto para 9 dB/oitava, devido ao aumento da pressão subglótica.

O efeito de maior esforço vocal também influencia as respostas de ressonância, de maneira que aparece mais nitidamente em F2 e F3 entre 1 e 3 kHz em comparação a F1. Dessa forma, o som que emerge na terminação final do tubo aproximado ao aparelho fonador (os lábios) tem os mesmos harmônicos que o som da fonte (glótica), entretanto as amplitudes dos harmônicos foram modificadas, alterando, de forma significativa, a qualidade do som (Borden et al, 1994).

Fant (1970) descreveu inicialmente as características de ressonância, simulando uma vogal aberta neutra. Os conceitos de formantes e antiformantes passaram a ser expostos no texto na base de pólos e zeros, que na linguagem das ciências exatas correspondem aos termos ressonâncias e antiressonâncias (Kent, Read, 1992). Os zeros são introduzidos quando existem sistemas de desvios de cavidade na produção da fala ou quando as cavidades atrás da fonte

passam ser consideradas, como é o caso da acoplagem do sistema subglótico em caso de fechamento incompleto de pregas vocais.

A primeira situação pode ser representada pela produção dos sons nasais e das consoantes líquidas e semivogais. A segunda ocorre em decorrência do fechamento incompleto de pregas vocais. Dessa maneira, na produção dos sons nasais, o ressoador oral se fecha, enquanto o nasal se abre para o meio externo e representa a parte final do tubo análogo ao do trato vocal. Nas consoantes líquidas (laterais, vibrantes e flapes) ocorre a divisão da cavidade oral pelo bloqueio central da língua. Nas semivogais a configuração articulatória muda, de forma a promover modificação gradual na estrutura formântica, que revela continuamente similaridades com os padrões dos sons vocálicos correspondentes. A expressão matemática seria semelhante àquela das vogais, porém haveria maior quantidade de zeros (Ladefoged, 1996).

## 2.1.2.2. Sons obstruintes (consoantes plosivas, fricativas e africadas)

Do ponto de vista acústico, a produção sonora depende do suprimento de energia e, portanto, deve-se primeiramente estabelecer a sua natureza ao descrever o sinal de fala. Nesse sentido, o modelo descrito por Fant (1970) contempla as variações da característica acústica em função das mudanças ocorridas no aparelho fonador durante a produção de fala. Sintetizando as possibilidades de combinação de fontes de som na produção da fala, poderiam ser citadas: nenhuma (silêncio), somente fonte de voz (sons ressoantes), somente fonte(s) de ruído (sons obstruintes surdos) e, finalmente, fonte de voz mesclada a fontes de ruído (sons obstruintes sonoros). Com exceção da segunda categoria, todas estão previstas na descrição dos sons obstruintes.

A teoria acústica da produção da fala estende sua descrição para os sons obstruintes, caracterizados por uma oscilação acústica sobreposta ao fluxo da respiração, causada por um objeto iniciando fricção, plosão ou interrupção da corrente aérea e podem ocorrer conjuntamente com a fonte sonora de vibração de pregas vocais (Fant, 1970). Tais fontes geram sons de natureza aperiódica no aparelho fonador, mais especificamente na região supraglótica e são representadas pelas fontes

de ruído transiente e contínuo, características respectivamente das consoantes plosivas e fricativas e na sua combinação no caso das africadas.

As fontes de ruído referem-se aos eventos acústicos primários no trato vocal, responsáveis pela geração de sons cochichados, aspirados, fricativos e plosivos. Ao descrever tais mudanças na corrente de ar impostas pela ação dos articuladores, caracteriza-se a turbulência no aparelho fonador como fonte de vários sons, incluindo a porção de fricção dos fricativos e o ruído transiente das plosivas.

Para os sons fricativos a aproximação dos articuladores força a passagem do fluxo de ar por uma constrição. Os fricativos também podem ser descritos em função das equações semelhantes àquelas abordadas para as vogais, com a diferença de que uma nova função se soma, Z(f), representando os zeros decorrentes da referida constrição em algum ponto do trato. Dessa forma, a função de transferência do ruído das fricativas envolve a adição de zeros. Para a função de radiação, o efeito é semelhante ao descrito para as vogais (Kent, Read, 1992).

No caso dos sons consonantais plosivos, a fonte passa a ser o estouro da liberação da oclusão. A fase de fechamento dos articuladores coincide acusticamente com o registro de silêncio. O momento do estouro guarda características do ruído das fricativas, porém de brevidade no tempo. A transição para a vogal seguinte é outro registro que pode guardar particularidades da emissão. Para os sons africados, há ocorrência de ruído transiente e contínuo (Kent, Read, 1992).

As categorias de sons sonoros, surdos e aspirados são rótulos para descrever parte de um *continuum* na fala, no qual o ruído passa a ser a fonte excitação do trato vocal (Ladefoged, 1973). Para o caso das fontes de ruído descritas, a transferência do trato vocal, do ponto de vista acústico, depende da abordagem da cavidade à frente da constrição e do ponto deste estreitamento no trato vocal. Quando a cavidade à frente da constrição é menor, como na produção de [s], o ruído se concentra em freqüências mais altas, em comparação a [ʃ], onde a cavidade é maior (Shaddle, 1990).

O fenômeno acima referido como aspiração merece maior consideração devido às suas implicações na análise acústica da fala. Para tanto,

retomaremos a descrição das consoantes plosivas e, inicialmente daquelas classificadas como surdas. Nessa situação, há um intervalo entre a posição de pregas vocais abduzidas e a gradativa adução para caracterizar a sonoridade da vogal seguinte. Tal ocorrência é comum a várias línguas, como é o caso da língua inglesa. Para Fant (1973), a ocorrência marcante deste período refere-se à maior abertura entre articuladores, aqui considerados como as próprias pregas vocais, levando inclusive a certo amortecimento de F1 da vogal subseqüente. Acentua a tendência à posição de laringe mais relaxada, discordando da descrição de traços distintivos por Chomsky, Halle (1968) para a qual referiram aumento da pressão subglótica. Sua discordância parece encontrar o respaldo de outros pesquisadores da área, os quais realmente apontam o caráter distintivo da aspiração em algumas línguas, porém ressaltam a condição fisiológica de maior abertura entre as pregas vocais quando comparada à produção dos sons surdos (Ladefoged, 1973, 1996; Borden et al, 1994; Laver, 1994). O som semelhante a um [h] denota o período de desvozeamento que sucede o estouro da oclusão, no qual as pregas vocais não atuam e a energia acústica gerada é fruto da corrente de ar passando pelo espaço entre ambas, maior do que aquele gerado na produção das consoantes surdas (Laver, 1994; Borden et al, 1994; Ladefoged, 1996).

Ladefoged (1973) relatou evidências de que em algumas línguas a duração da aspiração seria proporcional ao grau de abertura de pregas vocais. Laver (1994) destacou a duração do intervalo em que as pregas vocais iniciam a vibração para caracterizar a aspiração, de forma que seria caracterizada quando a vibração do segmento sonoro subseqüente ao plosivo tivesse início apenas 30 a 40 ms após o estouro da oclusão. Ressaltou que alguns autores definem a aspiração como o atraso da sonorização, entretanto este conceito não poderia ser generalizado, pois o primeiro tem aplicação mais estrita. Destacou que independente do contexto em que ocorra, a aspiração preenche a condição de que o tipo fonatório atuante em seu período não contém a sonoridade modal, mas algum tipo de ensurdecimento ou escape de ar (voz cochichada ou soprosa).

Para Borden et al (1994), a aspiração e a fonação podem ser assumidas como ações complementares em línguas como o inglês. Trazem à tona questões inerentes à percepção auditiva, recordando que para ouvintes de línguas com

marcante aspiração em segmentos surdos, os sons correspondentes (surdos) das línguas que não têm aspiração são confundidos com sonoros por falantes nativos.

Nesta confluência de ações vozeadas e não vozeadas para sons plosivos surdos, seguidos de vogal, surge o conceito de um índice acústico da maior importância para revelar o controle no tempo das ações motoras no aparelho fonador (Borden et al, 1994), o *VOT*, mantida a sigla para a designação original de *voice onset time*, numa tradução possível de tempo de início de sonorização ou tempo de ataque de vozeamento. Definido como o intervalo entre o estouro da plosiva, marcado pela liberação da corrente de ar e o início da ação de vozeamento pelas pregas vocais, pode ser classificado como pré-vozeado, simultâneo ou pós-vozeado, de intervalo longo ou curto (Lisker, Abramson, 1964) comumente referido como respectivamente negativo, zero e positivo (Kent, 1997).

## 2.1.3. Justificativas para adoção da teoria acústica da produção da fala no estudo das disfonias

Retomando a descrição detalhada dos sons ressoantes, destacamos a possibilidade de exploração dos vários componentes envolvidos na qualidade da produção vocal, enquanto mecanismos característicos da fonte e do filtro. Dessa forma, a abordagem acústica do sinal de fala, mais especificamente das vogais, permite a incursão a eventos característicos de vários níveis de atividade sonora, desde o controle e modulação do fluxo de ar pelas pregas vocais, possibilitando a especial condição de decomposição do sinal de saída emanado pela boca do falante disfônico, sem a necessidade de registro a partir de meios invasivos ao trato vocal, os quais poderiam comprometer a qualidade do sinal registrado.

A relativa independência entre fonte e filtro poderia, num primeiro momento, ser útil para a população em questão no sentido de permitir a possível identificação de mobilizações de vários segmentos do aparelho fonador. Embora tenhamos adotado essa perspectiva, gostaríamos de ressaltar a importância dos estudos acústicos que exploram a interação entre eventos relativos à fonte e ao filtro, revelando implicações importantes para a fala e o canto e para a reflexão no campo clínico da voz, como especialmente no caso da ações integradas das cavidades

subglóticas juntamente às supraglóticas em casos de fechamento incompleto de pregas vocais ou mesmo na seqüência de fala, nos momentos de produção de sons surdos (Fant, 1986; Klatt, Klatt, 1990; Bickley, Stevens, 1991; Cranen, Boves, 1991; Kent, Read, 1992; Kent, 1993; Båvergård, Fant 1995; Hanson, Stevens, 1995; Stevens, 1998).

Em um dos estudos precursores da teoria exposta, Flanagan (1958) abordou as possibilidades da atividade glótica sofrer influências do fluxo aéreo e das comparações entre efeitos resultantes da ação da fonte e do filtro e suas prováveis influências. Nesse campo, o trato subglótico passa a ser enfocado enquanto possíveis cavidades ressoadoras com freqüências estimadas em torno de 600, 1550 e 2200 Hz para indivíduo adulto do sexo masculino, com larguras de faixa entre 200 e 400 Hz (Klatt, Klatt, 1990; Hanson, Stevens, 1995; Stevens, 1998; Spencer, Titze, 2001). Tais valores representariam as ressonâncias naturais da região subglótica quando na vigência de fechamento glótico completo. Quando as pregas vocais se abrem passam a existir efeitos nas faixas presumíveis para ressonância, que podem inclusive corresponder às áreas de anti-ressonância (Stevens, 1998).

Na realidade, é justamente na interface dos eventos de cada nível de atividade do trato vocal que pretendemos discutir os eventos relevantes do sinal de voz tido como disfônico, com a possibilidade de descrição detalhada dos eventos sonoros, de forma a buscar as relações complexas que caracterizam os mecanismos comprometidos nesta população.

A abordagem da qualidade vocal a partir da presente referência, em que as contribuições dos níveis glótico e supraglótico do aparelho fonador foram exploradas do ponto de vista fonético-acústico, apresenta estreita relação com outro modelo teórico incorporado à pesquisa, o qual reforça a dimensão auditiva da qualidade vocal decorrente das mobilizações nos mesmos planos (glótico e supraglótico) que serão descritas no próximo item em termos do modelo fonético de descrição da qualidade vocal (Laver, 1980).

## 2.2. Modelo de descrição fonética da qualidade vocal

## 2.2.1. Perspectiva teórica

Tradicionalmente, a Fonoaudiologia estabelece uma distinção no enfoque de produção sonora pelo aparelho fonador, de forma que a qualidade vocal passa a ser considerada independentemente da fala. Numa tentativa de resgatar a abordagem integrada do sinal sonoro que emana da boca do falante e de contemplar a necessidade de caracterizar a ação do aparelho fonador e seu desempenho na produção vocal, correspondendo à complexidade do fenômeno estudado, optamos pela adoção da abordagem fonética. Neste referencial, pode ser desvendada a produção de sons pelo aparelho fonador em seus correlatos acústicos, auditivos (perceptivos) e fisiológicos (articulatórios).

Tal nível de correlação foi preconizado inicialmente na exposição do capítulo anterior referente à teoria acústica da produção da fala, a qual constitui a base dos estudos na área de Fonética Acústica. Utilizando-se da análise acústica para estabelecer correspondências entre os níveis auditivo e fisiológico, abordaremos os aspectos de ordem perceptivo-auditiva da qualidade vocal, na qual a descrição fonética pode colaborar para a mudança do panorama de dissociação entre qualidade vocal e fala anteriormente descrito.

Neste aspecto, ressaltamos que a distinção classicamente implementada na área de Fonoaudiologia entre fala e voz não é desejável, visto que a noção de qualidade vocal, segundo o modelo teórico adotado, inclui mobilizações dos níveis glótico e supraglótico do aparelho fonador. Os sons da fala são o resultado da filtragem efetuada pelo nível supraglótico sobre a energia sonora gerada na glote, no caso de sons constituídos apenas de fontes sonoras, ou nas próprias cavidades supraglóticas, no caso daqueles com fontes de ruído.

O aspecto temporal também deve ser abordado para que se definam as abordagens de curto termo, como é o caso dos segmentos (sons) da fala e de longo termo, nas quais as propriedades prosódicas destacam-se, sendo a qualidade vocal um destes aspectos de longo termo da emissão. Diante da concepção exposta, as

variações que ocorrem no fluxo da fala são tão somente um primeiro passo na incursão à abordagem da qualidade vocal.

O modelo fonético de descrição da qualidade vocal proposto por Laver (1980) representa uma importante contribuição no sentido de enfocar a qualidade vocal, enquanto característica individual do falante, como o produto de dois tipos de fatores, denominados intrínsecos e extrínsecos. Os primeiros derivam das características da anatomia própria ao aparelho fonador de determinado falante. Fatores extrínsecos, por outro lado, derivam dos ajustes musculares de longo termo do aparelho fonador intrínseco (Laver, 1980). Tais ajustes extrínsecos de longo termo são conhecidos como *settings* e constituem a unidade analítica do modelo em questão (Abercrombie, 1967; Laver, 1978, 1981, 2000).

A unidade analítica merece maior consideração, justamente por compor o ponto central do modelo de descrição da qualidade vocal. Sua definição é reforçada no âmbito do uso do aparelho fonador, enquanto padrões neuromusculares de controle voluntário, como o resultado de hábitos musculares aprendidos social e idiossincraticamente durante a vida do falante (Laver, 1979, 1980, 2000).

A qualidade vocal passa a ser estudada do ponto de vista de settings laríngeos e supralaríngeos, destacando-se que a noção de setting é muito anterior à proposta do presente modelo. O termo foi introduzido por Honikman (1964). Segundo Laver (1980), o conceito data do século XVII. O setting seria referido como um traço recorrente na posição do trato vocal conseqüente a um ajuste muscular de longa duração.

A proposta de modelo é apresentada com base em discussão sobre a motivação fonética para descrição de qualidade vocal. Em primeiro lugar, destaca a função semiótica da qualidade vocal, baseada na necessidade de descrição do conteúdo (contexto) da interação dos falantes. A qualidade vocal veicularia informações de características físicas, psíquicas e sociais do falante. Nesta concepção, o *setting* passa a ser descrito no nível suprassegmental e não seria considerado propriamente uma unidade lingüística.

A justificativa para o uso de tal unidade analítica encontra respaldo na abordagem de qualidade vocal enquanto uma "abstração cumulativa num período de tempo da qualidade característica de um falante, que lhe é peculiar nas momentâneas e esparsas flutuações de curta duração usadas para fins lingüísticos e paralingüísticos" (Laver, 1980). Ou ainda, segundo Abercrombie (1967), aquelas qualidades presentes mais ou menos o tempo todo num falante, enquanto a qualidade quase permanente, percorrendo a maior parte dos sons emanados pela boca do falante. Nesse sentido, o modelo inclui a noção temporal.

Diante dessa definição, Laver (1978) vislumbrou a possibilidade de analisar os *settings* na qualidade vocal (como tendências quase permanentes, sustentando e colorindo as articulações segmentares momentâneas) e na dinâmica vocal (como tendências restritivas dos aspectos de longo termo, tais como extensão de freqüência, de intensidade e duração).

No conjunto dos fatores de natureza orgânica que interferem na qualidade vocal, destacam-se aqueles de ordem estrutural, enquanto a constituição anatômica, os aspectos mecânicos e a massa do aparelho fonador. Os maiores elementos seriam a extensão do trato vocal, o formato e o volume das três cavidades que o moldam (bucal, faríngea e nasal), o formato de maxila e mandíbula, o tamanho e o formato dos dentes, o volume da língua, a constituição estrutural e a mecânica da laringe, o volume e a força do sistema respiratório (Laver, 1979, 1980, 2000). Tais fatores orgânicos contemplariam tanto elementos da ordem da normalidade, quanto das alterações do sistema (Laver, 1980, 1981, 2000), o que constitui um importante aspecto a apontar a indicação do modelo para o campo clínico.

A influência dos níveis orgânico e fonético sobre a percepção da qualidade vocal pode ser detectada ao longo de variações momentâneas relacionadas aos segmentos. Nesse sentido, a qualidade vocal emerge como uma abstração cumulativa num período de tempo da qualidade característica do falante, a qual é deduzida das flutuações momentâneas, esporádicas da articulação de curto termo (segmentos), usadas pelo falante com finalidade lingüística e paralingüística (Laver, 1980, 2000).

Neste ponto da descrição do modelo, Laver (1980) estabeleceu importante distinção entre os segmentos fonéticos e os *settings*. A diferença básica entre ambos repousaria no domínio do tempo, sendo os primeiros de curta duração e os últimos, ajustes de longo termo. Na Fonética tradicional seria equivalente à menção a articulações secundárias, no que se refere ao *setting*.

Outro ponto que emerge na discussão dos conceitos da Fonética clássica seria a tradição em enfocar a Fonética enquanto segmentação do fluxo contínuo da fala em unidades lingüísticas básicas (Laver, 1979). Dada a motivação lingüística, as diferenças entre os segmentos passam a ser enfatizadas, muito mais do que as similaridades. Como alternativa, o mesmo autor propõe uma abordagem mais ampla, referente tanto às diferenças, quanto às similaridades na performance da fala, focando os segmentos individuais como ações momentâneas sobrepostas a um *setting* de longo termo do aparelho fonador. Nessa concepção, os *settings* representariam as similaridades e os segmentos as diferenças (Laver, 1979).

Laver (1980) adota a visão da língua enquanto um instrumento social, na qual a teoria fonética deve ser a responsável por descrever toda a atividade recorrente, padronizada e sonora que caracteriza a linguagem falada. Nesse sentido, Esling (2000) realiza uma revisão dos aspectos da qualidade vocal entre línguas diversas, reforçando alguns conceitos de Laver (1980) com relação à função lingüística da qualidade vocal.

Tratando da complexa relação entre os elementos apontados (os settings e os segmentos), Laver (1980) descreve os princípios que regem a relação tanto entre os vários settings, quanto aquela referente aos settings e aos segmentos.

Nos princípios que regem a relação entre os *settings*, destacam-se a interdependência e a compatibilidade. O primeiro atua nos níveis acústico e fisiológico, podendo ser fundamentado no funcionamento interdependente do trato vocal laríngeo e supralaríngeo. No caso da acústica, dois *settings* ocorreriam simultaneamente, explorando diferentes partes do trato supralaríngeo. Na interdependência fisiológica, um *setting* particular pode acionar outro mecanismo simultaneamente, justamente pelo fato da musculatura estar interligada.

A compatibilidade opera nos níveis fisiológico, auditivo e acústico e resulta em exclusão de um deles por serem incompatíveis, soma de efeitos por serem compatíveis e diminuição da proeminência daquele mais vulnerável. A compatibilidade rege a relação entre o *setting* e a anatomia individual, na medida em que as variantes anatômicas do aparelho fonador determinam sua facilidade ou dificuldade na realização de determinada mobilização.

Nos princípios que regem a relação entre os settings e os segmentos, além dos referidos princípios de compatibilidade e interdependência, ressalta a susceptibilidade do segmento ao efeito articulatório, auditório e acústico do setting. Nesse âmbito, nenhum setting se aplica normalmente a cada segmento da emissão do falante. Torna-se importante dividir os segmentos entre aqueles mais susceptíveis à influência do setting e aqueles não susceptíveis. Por extensão, os settings podem ser abstraídos da cadeia de segmentos como as propriedades partilhadas dos elementos com a mesma susceptibilidade.

Diante das descrições anteriores, a unidade analítica do modelo exposto até o presente ponto de exposição referido como setting passará a ser referido como ajuste ao longo do presente trabalho.

Neste ponto da revisão teórica e de contraposição aos conceitos da Fonética clássica, vale ressaltar que o próprio autor apontou a possibilidade de os ajustes poderem ter diferentes domínios temporais. Sendo de natureza polissegmental, um ajuste pode (e deve) ser uma propriedade que se estende por mais de um segmento, mas não há limite para sua ocorrência no tempo (Laver, 1978, 1979, 1980).

Laver (1980) prevê a supremacia do segmento em relação ao ajuste, postulando o primeiro como básico e o segundo como derivado. A própria susceptibilidade do segmento ao ajuste poderia entretanto pode ser usada como argumentação no sentido de se defender uma inversão dessas posições. Essa é uma crítica que pode ser feita à concepção do modelo (Albano, 1988).

Partindo para a descrição das qualidades vocais, a exposição dos ajustes possíveis e previsíveis pelo aparelho fonador parte do conceito de um ajuste de referência: o ajuste neutro. Refere-se a uma constelação de ajustes ocorrendo

simultaneamente em diversos locais do trato vocal e não somente à situação de normalidade ou estado de repouso do aparelho fonador. Trata-se de um sistema de referência a partir do qual torna-se possível caracterizar os ajustes que resultam em determinada qualidade vocal.

O ajuste neutro foi caracterizado em diversas publicações (Laver, 1979, 1980, 1994, 2000; Laver et al, 1981) da seguinte maneira:

- lábios: não projetados;
- · laringe: nem abaixada, nem elevada;
- diâmetro do trato supralaríngeo: constante;
- articulações orais anteriores: lâmina da língua;
- base de língua: nem avançada, nem recuada;
- trato vocal sem constrições;
- mandíbula: nem acentuadamente aberta, nem fechada;
- nasalidade: presente apenas nos segmentos característicos;
- vibração de pregas vocais: periódica, com tensão longitudinal e adutora moderada, idem para compressão medial;
- uso da corrente de ar: eficiente, sem ruídos;
- tensão muscular no aparelho fonador: nem alta, nem baixa.

De maneira sintética, as condições que se aplicam à situação de ajuste neutro seriam a de que extensão do trato vocal não deve ser modificada pela ação da musculatura. Além disso, o trato vocal não deve sofrer perturbações em qualquer ponto por ação de lábios, mandíbula, língua ou faringe (Laver, 2000).

Do ponto de vista fonatório, Laver (2000) destaca que a vibração das pregas vocais deve ser regular, periódica, sem ruído audível decorrente de aperiodicidade. Deve também ser eficiente no uso de ar, sem fricção audível, de forma que o grau de atividade de todos os músculos do sistema fonatório deve ser moderado.

As qualidades vocais foram descritas a partir dos ajustes que representam variantes do ajuste neutro em nível supralaríngeo (articulatório) e laríngeo (fonatório). Ao todo, foram descritos 53 tipos de ajustes (Laver, 1980), agrupados em laríngeos, supralaríngeos e de tensão.

Cada um dos ajustes é cuidadosamente abordado no modelo, com as mobilizações do aparelho fonador agrupadas em ajustes supralaríngeos, laríngeos e de tensão. Compuseram esta descrição, aspectos de estruturas como lábios, posição de laringe no pescoço, do trato vocal supralaríngeo (da ponta de língua, da base de língua, dos pilares, dos músculos constritores da faringe, da mandíbula, do esfíncter velofaríngeo), da vibração de pregas vocais e da tensão da musculatura do trato vocal.

A descrição dos ajustes supralaríngeos envolve aqueles relacionados à alteração da dimensão vertical do trato vocal (longitudinais) e de diâmetro (transversais), além daqueles relativos ao esfíncter velofaríngeo, resultando na descrição esquemática abaixo:

Ajustes Supralaríngeos

## Longitudinais

Modificações efetuadas pelo falante no eixo longitudinal do trato vocal

Laringe

Alta

Baixa

Lábios

Protrusão

Labiodentalização

## •Transversais (latitudinais)

Tendências quase permanentes a manter um efeito particular constritivo ou expansivo na área transversal de todo o trato vocal

#### Lábios

Expansão horizontal do espaço interlabial

Expansão vertical

Constrição horizontal

Constrição vertical

Expansão horizontal e expansão vertical

Constrição horizontal e constrição vertical

Expansão horizontal e constrição vertical

## Constrição horizontal e expansão vertical

## •Língua

Ponta/lâmina

Articulação de ponta

Articulação de lâmina

Articulação retroflexa

## Corpo

Dentalizada

Alveolarizada

Palatoalveolarizada

Palatalizada

Velarizada

Uvularizada

Faringealizada

Faringolaringealizada

## Base

Base de língua avançada

Base de língua retraída

## Pilares

De fauces ou de pilares

•Faringe

Faringealizada

Mandíbula

Posição aberta

Posição fechada

Posição protruída

Sistema Velo-faríngeo

Nasal

Denasal

Na descrição dos ajustes fonatórios (laríngeos), ganharam destaque os parâmetros de atividade muscular, destacando aqueles de natureza adutora da glote cartilaginosa, de compressão medial pela ação da musculatura intrínseca adutora e de tensão longitudinal, além de ajustes de tensão generalizada. Deve-se considerar que

as descrições de ordem fisiológica foram baseadas em estudos anteriores à publicação original e, portanto, sofreram vários refinamentos nas duas últimas décadas.

Além disso, as descrições não foram derivadas de técnicas de investigação do aparelho fonador durante a realização dos referidos ajustes, mas foram supostas para aqueles gerados pelo próprio autor e registradas em gravação em áudio. As descrições não se caracterizam como o ponto central do modelo, de forma que os refinamentos em relação à terminologia e ação mais refinada no plano de atividade intrínseca da laringe e, inclusive, de demais segmentos do aparelho fonador, poderiam ser tidos como material a ser submetido a futura revisão e implementação com a expectativa de publicação posterior com validação de dados da esfera fisiológica, acústica e perceptiva, porém tal não ocorreu.

Na época da descrição, o instrumental de análise envolveu programas de computador para identificação de pólos de predição linear, espectrógrafo, tubo de Sondhi e laringógrafo. Uma descrição mais apurada, pautada na tecnologia atual para investigação do aparelho fonador em atividade, certamente nos vislumbraria com o detalhamento de atividade do aparelho fonador e da sonoridade resultante. Apesar de publicação do modelo não ter sido revisada posteriormente, o roteiro de avaliação dele derivado foi extensamente revisado e será apresentado neste trabalho em novo item referente ao delineamento de tendência de avaliação nas disfonias. Mobilizações possíveis neste plano de atividade incluem:

## **Ajustes Laríngeos**

#### Tipos simples

Modal

Falsete

Cochicho

Vocal fry (crepitância)

#### Tipos compostos

Vocal fry cochichado

Voz cochichada

Falsete cochichado

Voz crepitante

Falsete crepitante

Voz cochichada crepitante

Falsete cochichado crepitante

Voz soprosa

Voz áspera

Falsete áspero

Vocal fry áspero

Voz áspera cochichada

Falsete áspero cochichado

Voz áspera crepitante

Falsete áspero crepitante

Voz áspera cochichada crepitante

Falsete áspero cochichado crepitante

Os ajustes de tensão foram, posteriormente, excluídos de roteiros baseados no modelo, justamente por representarem um ponto polêmico na proposta original. Há ajustes de qualidades vocais que se caracterizam pela presença de um componente de tensão localizado. Ao postularmos um ajuste de tensão generalizado, poderíamos eventualmente ter a sobreposição de ajustes localizados e generalizados de tensão ou apenas ajustes generalizados. Isso iria contra a postulação básica do modelo de que os ajustes são atomísticos e não holísticos. Uma região do aparelho fonador na qual tal critério poderia ser exemplificado seria o nível glótico, mais especificamente o fechamento glótico, aspecto este que será abordado no texto relativo à proposta de avaliação de qualidade vocal a ser utilizada no presente trabalho.

A descrição do modelo é acompanhada por um sistema de notação que permite a transcrição dos ajustes e uma gravação, na qual o próprio autor executa as várias mobilizações características dos ajustes descritos, de forma a ilustrar o conteúdo apresentado (Laver, 1980). Na mesma gravação, existem exemplos de vozes de falantes de variados dialetos, complementando a simulação efetuada com muita habilidade pelo autor.

Da exposição acima, emerge um sistema descritivo da qualidade vocal como um modelo fonético geral aplicável à performance de todos os seres humanos de anatomia e fisiologia normais. Entretanto Laver (1980) destaca uma provável aplicação futura para o campo das alterações de qualidade vocal, ao argumentar que apesar do modelo em si traçar apenas uma vasta gama de ajustes dentro da normalidade, muitos deles podem despertar interesse no caso de alterações em diversos níveis do aparelho fonador.

A importância de descrição de qualidade vocal, seguindo a proposta inicial de Laver (1980), repousa no fato de constituir um dos sistemas de classificação vocal mais ricos e detalhados, propiciando uma descrição altamente representativa das variações de qualidade vocal (Kent, Ball, 2000).

Neste ponto da exposição, torna-se importante expor os demais aspectos que caracterizam o modelo fonético de descrição da qualidade vocal (Laver, 1979, 1980). Enquanto modelo científico, parte de descrições não-impressionísticas, as quais permeiam as discussões e a própria terminologia tradicional e amplamente utilizada em avaliação de qualidade vocal.

O modelo é descritivo no sentido em que oferece fundamentação e sistematização com base na percepção auditiva, estabelecendo correspondência com correlatos específicos em cada um dos três níveis de análise: articulatório, fisiológico e acústico. Laver (1980) reforça que as especificações foram oferecidas por sua total responsabilidade, registrando a tentativa de obter informações com especialistas em cada uma das áreas, particularmente em fisiologia e acústica. Tal aspecto caracteriza um ponto positivo no modelo ao oferecer descrições sistematizadas e fundamentadas dos componentes de qualidade vocal.

O modelo também é replicável, de forma que juízes podem ser treinados a emitir julgamentos consistentes do sistema, preenchendo uma premissa importante quando se consideram as particularidades das análises de ordem perceptiva e a subjetividade dos julgamentos.

A característica integrativa emerge da concepção da análise fonética. Cada qualidade sonora é analisada enquanto o conjunto dos ajustes

independentemente controláveis, os quais compõem o efeito auditivo final. Dessa forma, estabelece-se um sistema de referências, no qual há muitas combinações possíveis, não sendo necessária a descrição de todos os tipos de vozes, mas apenas dos aspectos primários, a partir dos quais se estabelecem as combinações, respeitando os princípios da interdependência e da compatibilidade dos ajustes. A adoção das descrições em âmbito da acústica permitem a integração das informações auditivas, articulatórias e fisiológicas, à semelhança do que pode ser depreendido a partir da exposição dos vários ajustes possíveis, descritos previamente nesta revisão.

Nesse aspecto, destaca-se a última característica do modelo, a sua natureza atomística, a qual, ao contrário do modelo holístico, descreve os aspectos pontuais relativos aos efeitos sonoros decorrentes de elementos identificáveis e controláveis de forma independente no aparelho fonador. Cada qualidade vocal é analisada segundo todos os ajustes possíveis e independentes, cujos efeitos de composição caracterizaram a qualidade vocal final (Laver, 1979). Nesse aspecto, destaca-se a possibilidade de localização dos pontos do aparelho fonador que realmente contribuíram para a qualidade sonora gerada. As implicações clínicas para tal caracterização são inegáveis, uma vez que os falantes portadores de alterações vocais podem apresentar mobilizações altamente individualizadas para compensar as dificuldades presentes, cabendo aqui inclusive uma proposta de identificação de ajustes compensatórios.

## 2.2.2. Justificativas para adoção do modelo no enfoque da qualidade vocal de indivíduos disfônicos

O modelo é cientificamente descrito, integrando informações detalhadas dos correlatos articulatórios, auditivos e acústicos decorrentes dos mais variados ajustes laríngeos e supralaríngeos. Sendo um modelo geral, aplicável à performance do trato vocal dos seres humanos e considerando as diferenças individuais, pode-se fazer uma adaptação para a população em questão, uma vez que o sistema de referência das diversas modalidades analíticas pode servir de base ao estudo dos refinados ajustes presentes nas tentativas de compensação de uma situação limitante de produção vocal, como é o caso da paralisia de prega vocal.

Além disso, sabe-se da refinada ação do trato vocal à fonação e das variadas compensações que podem ocorrer, graças à sua plasticidade. Dessa forma, o sistema de referência dos variados ajustes permite-nos investigar a refinada ação de compensação, fato este justificável inclusive pelas descrições acústicas, nas quais se assume que exista uma relativa interdependência entre os níveis de atividade glótica e supraglótica, ou mais precisamente, laríngeo e supralaríngeo. Além disso, o enfoque tradicional da situação de manifestação de disfonia tem como base a comparação aos padrões estabelecidos para a situação de normalidade, o que é altamente questionável na literatura (Hollien, 2000).

Por outro lado, o enfoque do ajuste neutro afasta a comparação ou referência à normalidade e passa a se referir aos possíveis ajustes do trato vocal, configurando-se como um excelente sistema de referência. Outro ponto a ser destacado como argumento favorável à adoção do modelo é o fato do quadro a ser estudado interferir em alguns mecanismos relacionados ao plano glótico, mantendo intacta a estrutura do restante do trato vocal. Nesse sentido, não fere as bases para as quais o modelo foi descrito, de forma que as suas descrições podem contribuir para entendermos tanto o plano alterado (mecanismos glóticos) quanto aqueles preservados, os quais podem entrar em atividade, no vasto campo da caracterização dos mecanismos compensatórios.

Algumas limitações foram apontadas ao modelo, porém não se verifica na literatura qualquer outra descrição semelhante, principalmente no que se refere aos correlatos acústicos, fisiológicos e auditivos das mais variadas qualidades vocais. Nesse sentido, apresenta-se como um verdadeiro refinamento para descrição da qualidade vocal, inclusive aplicável para as situações de alteração da qualidade vocal.

Considerada dessa forma, a motivação fonética ao se estudar a qualidade vocal passa a ser inevitável, permitindo decompor, a partir do estudo da qualidade fônica, tanto os aspectos de componentes da qualidade fonética, enfocada enquanto conseqüência da influência de segmentos e ajustes, como da qualidade vocal, tida como conseqüência da influência de ajustes e de componentes orgânicos (Laver, 1979). Estudos desta natureza exploram a base da atuação fonoaudiológica,

não exclusivamente para a situação do campo da disfonia, uma vez que a argumentação pode ser estendida para muitas outras manifestações.

Nas publicações subseqüentes (Laver et al, 1981; Laver, 1991, 1994, 2000), o modelo passa a fundamentar o roteiro de avaliação do perfil vocal, conhecido em sua designação original como *VPAS* (*Vocal Profile Analysis Scheme*), contemplando os graus de variação que incluíam a situação de alteração vocal, com sugestões para uso em rotina clínica.

O modelo sofreu algumas críticas, como Fourcin (2000), ao tecer breves comentários a respeito de propostas de avaliação vocal, numa publicação dedicada a eletrolaringografia, refere-se à proposta de Laver (1991) como um modelo pseudo-foneticamente motivado, argumentando que teria confundido voz, no sentido de excitação do trato vocal, com atributos de ajustes do trato vocal. Sua crítica é, entretanto, extremamente superficial, sem consideração sobre as bases teóricas do modelo.

## 2.2.3. Proposta de avaliação da qualidade vocal - VPAS

O roteiro *VPAS* foi desenvolvido com base na habilidade perceptiva auditiva dos fonoaudiólogos em julgar as vozes enquanto combinações de ajustes complementares, cujo método pode ser aprendido. A modalidade publicada em Laver et al (1981) representa a décima versão do protocolo, como resultado do trabalho de uma equipe composta por foneticistas, fonoaudiólogos e cientistas da fala, financiado pelo Conselho Governamental Britânico de Pesquisas Médicas. Procura estabelecer também as correlações da acústica das diversas categorias de ajustes aos fenômenos subjacentes de ordem fisiológica.

Os ajustes presentes na proposta de avaliação incluem aqueles de natureza supralaríngea (longitudinais, transversais e de acoplagem velar) e laríngea ou fonatória. Cada um dos ajustes presentes na escala atual (Laver, 2000) será brevemente descrito na seqüência deste texto, adotando-se uma breve referência aos aspectos auditivos acústicos e fisiológicos decorrentes das variadas configurações assumidas pelo aparelho fonador.

Vale ressaltar a característica integrativa do modelo, de forma que muitas vozes representam os tipos compostos, resultantes da combinação de vários ajustes descritos. Conseqüentemente, a exposição de cada ajuste não esgota as possibilidades de qualidades vocais previstas. Algumas adaptações de ajustes foram efetuadas em função da publicação de Laver (2000) e dos próprios conhecimentos atuais da fisiologia do aparelho fonador, de forma que alguns dos conceitos expostos na publicação de Laver (1980) não serão abordados na presente abordagem de qualidade vocal.

Nos tipos combinados, devemos considerar as limitações de compatibilidade entre os ajustes, com base nos correlatos fisiológicos e acústicos (Laver, 2000).

No grupo dos ajustes supralaríngeos longitudinais destacam-se os mecanismos que promovem mudanças na extensão total do trato vocal, totalizando um conjunto de quatro mobilizações possíveis:

- •Laringe alta: encurta a extensão total do trato vocal, tendo como consequência acústica o aumento da frequência de todos os formantes da emissão e tendência ao aumento de valores de frequência fundamental (f0).
- •Laringe baixa: aumenta a extensão total do trato vocal, gerando acusticamente o declínio da freqüência de todos os formantes, especialmente dos mais baixos e tendência à diminuição dos valores de f0
- •Protrusão labial (contração que vai além da mudança do espaço horizontal entre os lábios): aumenta o eixo longitudinal do trato vocal, diminuindo a freqüência de todos os formantes.
- Labiodentalização (consequência do ato paralingüístico de sorrir): diminui de forma discreta a extensão total do trato vocal, com interferência auditiva mínima. Acusticamente o efeito está relacionado ao fechamento da abertura labial, gerando uma constrição labial que favorece o rebaixamento da frequência dos formantes.

No grupo dos ajustes supralaríngeos latitudinais (transversais) são referidas as tendências de manutenção de um efeito constritivo (ou expansivo) no diâmetro de uma dada localização ao longo do trato vocal. Tais modificações devem-se

à ação de diversos órgãos, definindo vários grupos: labiais, mandibulares, linguais e

faríngeos. São referidos separadamente no roteiro de avaliação:

•Labiais: aumentam ou diminuem o espaço entre os lábios. São descritos todos os

estados possíveis de parâmetros horizontais e verticais, totalizando oito tipos de

combinações entre as ações de contração (C) e expansão (E) nos planos vertical

(V) e horizontal (H). São previstas as possíveis combinações com protrusão e não

protrusão labial descritos no plano longitudinal. Os efeitos acústicos são variáveis

em função das combinações descritas.

•Mandibulares: movimentos verticais (abertura e fechamento), horizontais, laterais e

rotação. Os efeitos acústicos são similares aos ajustes de lábios (latitudinais) e aos

longitudinais, incluindo elevação da freqüência do primeiro formante (F1)

proporcional ao aumento da abertura.

•Linguais: o ponto de referência é a posição de longo termo do centro da língua. A

descrição está concentrada no local e grau da constrição máxima no trato vocal,

não importando se ocorreu na cavidade bucal ou faríngea. São separados neste

ponto ajustes de ponta e lâmina e dorso de língua, numa nomenclatura que retoma

a terminologia fonética no que se refere ao ponto de articulação dos sons

consonantais.

•Componente de anteriorização: acusticamente promove a distância entre a

frequência de F1 e F2.

Dentalizado

Alveolarizado

Palatoalveolarizado

Palatalizado

Componente de posteriorização: tendência geral a aumento da freguência de

F1 e diminuição de F2.

Velarizado

Uvularizado

Faringealizado

Faringolaringealizado

•Faríngeos (orofaringe): ligados à base de língua, a qual (independente da

conformação do corpo) fecha ou abre a orofaringe. Efeitos acústicos na contração de

faringe incluem o aumento de frequência de F1 e a diminuição de F2, ficando os formantes mais próximos.

- Constrição faríngea: fecha a parte média da orofaringe por contração das paredes de faringe ou retração do corpo da língua, ocasionando aumento de F1 e rebaixamento de F2, fazendo com que as duas freqüências aproximem-se.
- Expansão faríngea: abre a parte média da orofaringe, levando geralmente a diminuição da freqüência de F1.

Os ajustes de faringe também afetam a banda dos formantes, na dependência da tensão de suas paredes, quanto a maior tensão, menor a banda do formante, denotando tendência de absorção da energia da onda sonora no caso de ausência de tensão de suas paredes.

Como última categoria dos ajustes supralaríngeos, Laver (1980) expõe os velofaríngeos, para os quais reserva muitas discussões e argumentos, destacando que várias inadequações emergem do conceito de nasalidade, tais como o fato da nasalidade ser um conceito auditivo que congrega uma série de fenômenos auditivamente similares. Nesse sentido, destaca sua condição de ressonância de natureza particular. Reforça que há um fator essencial para induzir a ressonância na cavidade nasal: a razão do diâmetro da abertura horizontal da faringe até a cavidade nasal pela abertura relativa vertical da faringe à boca. Somente quando esta razão excede um nível crítico, a ressonância audível será induzida (Laver, 1979). Divide esta categoria em dois ajustes:

Nasal

Denasal

Finalmente nos ajustes laríngeos (fonatórios), Laver (1980) inicia com a descrição do ajuste laríngeo neutro, referente à voz modal. A partir daí descreve três pares de categorias básicas e suas possíveis combinações:

- •Modal e falsete: ajustes que ocorrem isoladamente e em combinações, porém não entre si.
- •Cochicho e vocal fry: ajustes ocorrem isoladamente e combinados entre si.

•Aspereza e soprosidade: componentes, e não propriamente ajustes, que ocorrem apenas em tipos compostos, combinando-se a algum ajuste.

Dessas possibilidades de combinação surgem os seguintes ajustes:

### Simples

- •Voz modal: descrita no ajuste neutro, com respeito à condição das pregas vocais nos aspectos de adução, tensão e características de vibração.
- •Falsete: as pregas vocais permanecem ligeiramente abertas, e a pressão subglótica é menor do que na voz modal. A pequena abertura glótica acrescenta um componente mais próximo a um cochicho (escape de ar com fricção audível), do que à soprosidade (escape de ar sem ruído audível). Entretanto como a pressão subglótica é mínima, esta característica será leve. Acusticamente apresenta valores de f0 altos, poucos harmônicos, declínio espectral acentuado, da ordem de cerca de 20 dB por oitava.
- •Cochicho: ajuste que envolve baixa tensão adutora e compressão medial de moderada a alta. Tais mecanismos tende a gerar ficção audível. Proposto como tradução aproximada para o termo original *whisper*.
- •Vocal fry (crepitância): as pregas vocais encontram-se aduzidas (volumosas e comprimidas), com tensão adutora alta e longitudinal baixa. As pregas vestibulares podem estar aduzidas, com possibilidade da face inferior de prega vestibular entrar em contato com porção superior da prega vocal. O resultado auditivo é de um trem de pulsos. Os valores de fo são baixos. Na literatura fonética é também referido como laringealização.

### Compostos

- •Voz cochichada (combinação dos ajustes cochicho e modal), no qual há fricção audível causada pelo escape de ar por uma pequena fenda. Na designação original seria a qualidade referida como *whispery voice*.
- •Voz crepitante (combinação dos ajustes *vocal fry* e modal): pulsos irregulares, baixos valores de f0. Na designação original seria a qualidade referida como *creaky voice*. Na literatura fonética é referido como voz laringealizada ou rangida.

Neste grupo também são descritos os componentes que podem modificar alguns ajustes, porém não registram ocorrência isolada:

- •Aspereza: aumento da tensão laríngea, resultando em onda glótica irregular, perturbações de f0 e característica sonora desagradável.
- •Soprosidade: variação do ajuste modal, resultando em modo de vibração insuficiente e acompanhado por leve e discreta fricção audível. A compressão medial é mínima e o esforço muscular é pequeno. Por esta razão, só vai combinar com o ajuste modal e tem pouco efeito auditivo.

Os últimos dois tipos fonatórios compostos descritos no *VPAS* referem-se à combinação dos dois ajustes acima descritos com o ajuste modal, resultando em:

- Voz áspera (combinação do componente aspereza e do ajuste modal): grande tensão adutora e compressão medial comparada ao ajuste modal.
   Para iniciar a fonação, há compensação com aumento da pressão subglótica.
- •Voz soprosa (combinação do componente soprosidade e do ajuste modal): redução da tensão da musculatura do sistema muscular fonatório, levando ao escape de ar maior do que no ajuste de voz soprosa, porém com menos fricção em relação ajuste de voz cochichada.

O leitor automaticamente recorre à questão relacionada à voz rouca ou rouquidão como possível ajuste a também compor o roteiro. A sugestão refere-se à combinação do ajuste cochicho (pelo escape de ar) com o componente de aspereza (irregularidade vibratória), resultando em ajuste de voz áspero-cochichada (Laver, 1979, 1980; Laver et al, 1981), podendo registrar o acréscimo de crepitância (Laver et al, 1982).

A referência aos ajustes de tensão geral não aparecem descritas na versão mais atual do *VPAS* (Laver, 2000), entretanto esteve presente em Laver et al (1981), na qual são referidos os tópicos de tensão supralaríngea e laríngea. Laver (1979) descreve que seria preferível manter a denominação de tensão geral, uma vez que os ajustes de tensão poderiam ser descritos como conjuntos de ajustes locais

simultâneos no aparelho fonador. As críticas ao fator tensão foram devidamente abordadas no item anterior, de forma que alguns ajustes prevêem, em si, certo grau de tensão muscular em contraposição ao ajuste neutro. Outra diferença refere-se à inclusão de aspectos prosódicos, de organização temporal e comentários referentes ao suporte respiratório e diplofonia na primeira versão do roteiro referida.

No anexo 1 segue exposta a versão da proposta de Laver (2000) em versão original e com proposta de tradução baseada em terminologia amplamente utilizada no meio fonoaudiológico.

Na primeira versão publicada (Laver et al, 1981), a aplicação do *VPAS* era dividida em passos. No primeiro, o avaliador estabelecia a distinção entre ajuste neutro e não neutro (subdivido em normal e anormal), conhecido como primeira passada. Na segunda etapa de análise, conhecida como a segunda passada, o avaliador atribuía graus, de forma os ajustes considerados normais foram subdivididos em graus 1, 2 e 3 e anormais em graus 4, 5 e 6. A menção a anormal referia a necessidade de tratamento de alterações da qualidade vocal. Tais denominações não se encontram presentes no mesmo formato na versão mais recente, a qual apresenta a divisão em seis graus de escala, porém sem as denominações normal e anormal no roteiro. Entretanto, os três primeiros graus da escala seriam referência à situação de normalidade e os demais à situação patológica ou extremo paralingüístico de expressão da emoção.

Nesse sentido, pode-se supor uma dicotomia para análise, nas quais ficam praticamente excluídas as possibilidades de manifestação de alteração em grau leve em relação ao ajuste neutro. Tal consideração não será mantida na proposta de avaliação de qualidade vocal adotada no presente estudo.

Os graus da escala dos ajustes propostos no *VPAS* obedecem à orientação de que a divisão básica é efetuada em termos da susceptibilidade do segmento ao efeito do ajuste. Quando as seqüências fonológicas são respeitadas no produto acústico, define-se a situação de normalidade. Quando o ajuste afeta em maior grau os segmentos não susceptíveis, trata-se de uma situação de fala com comprometimento. O fato de não afetar o segmento não exclui a possibilidade de alteração de qualidade vocal.

### Os graus são distribuídos da seguinte maneira:

- •1-3: faixa dos ajustes presentes na fala normal
- 1- pequena diferença em relação ao ajuste neutro
- 2- leve diferença em relação ao ajuste neutro
- 3- moderada diferença em relação ao ajuste neutro
- •4-6 patologia ou extremos paralingüísticos da expressão da emoção
- 4- notável diferença em relação ao ajuste neutro
- 5- severa diferença em relação ao ajuste neutro
- 6- extrema diferença em relação ao ajuste neutro

Algumas observações adicionais referem-se a exceções na aplicação da escala acima exposta. No caso dos ajustes supralaríngeos, Laver (2000) destaca que os faríngeos são divididos em apenas duas categorias, além do neutro, devido à dificuldade de discriminação auditiva de pequenas variações da expansão ou constrição faríngea. Dessa maneira, as possibilidades não neutras incluem constrição em faixa normal contra excessiva (alterada). A denasalidade é julgada da mesma maneira. No caso dos ajustes fonatórios, tanto o falsete como o cochicho (quando representarem os únicos ajustes fonatórios presentes) são tidos como evidência de desordem vocal. A ocorrência isolada de ajuste fonatório *vocal fry* (crepitância) também é considerada como indício de patologia.

O roteiro também comporta inscrições que denotem ajustes que ocorram em menor freqüência em relação aos demais por meio da inscrição "i", sinalizando que o mesmo é de ocorrência intermitente (Laver, 2000).

Como consideração final de exposição do *VPAS*, Laver (2000) ressalta a equivalência de fatores orgânicos e fonéticos na qualidade vocal, de forma que um mesmo ajuste poderia ocorrer por um estilo característico do falante ou por uma conformação anatômica particular. Destaca que na teoria fonética geral, as diferenças entre falantes são geralmente ignoradas, partindo-se da concepção de um aparelho fonador ideal e tratando da qualidade vocal exclusivamente pelo apelo da qualidade fonética. Reforça ainda que os conceitos de qualidade fonética e de qualidade vocal não podem ser totalmente compreendidos sem a consideração da influência dos fatores orgânicos na fala em geral e na qualidade vocal em particular. Os aspectos relativos à aplicação clínica e adaptação de roteiro de avaliação vocal serão retomados quando da descrição de proposta de métodos para o estudo em questão.

# 3. REVISÃO CRÍTICA DE ESTUDOS DAS DISFONIAS: CORRESPONDÊNCIA ENTRE PARÂMETROS ACÚSTICOS, PERCEPTIVOS E FISIOLÓGICOS

Este capítulo destina-se à revisão crítica da investigação das alterações vocais em seus desdobramentos acústicos, perceptivos e fisiológicos, incluindo especificidades da situação eleita para estudo, representada pelas limitações de mecanismos glóticos. Tal perspectiva é delineada em tópico inicial relativo ao histórico de estudos que influenciaram as diversas áreas de conhecimento envolvidas no estudo das disfonias. Na seqüência, são apontadas as tendências vigentes, numa continuidade das abordagens registradas no histórico anterior, na medida em que contemplem aspectos concernentes à concepção teórica deste trabalho. Estruturado desta maneira, este capítulo destina-se a traçar um panorama da evolução dos estudos das alterações vocais, a culminar com a concepção de método adotado no presente estudo.

### 3.1. Histórico

As alterações vocais são enfocadas há muitos anos na literatura, traduzindo várias tendências de abordagem, concentradas em descrições com enfoque inicialmente de base perceptiva, gradualmente incorporando investigações por descrições acústicas e, finalmente, complementadas por avaliações de ordem fisiológica. Em cada uma das esferas apontadas destacaram-se propostas diferenciadas no que se refere ao instrumental de análise, ao tratamento de dados e à sua interpretação, os quais, em muitas situações dificultam a comparação dos dados entre si. Há também que se considerar as questões relativas à língua e portanto a padrões diferenciados para falantes de distintas comunidades, mesmo quando se enfoca o mesmo país.

O histórico de tais modalidades de análise guarda importante relação com o advento de instrumentos que favoreceram o maior detalhamento de informações referentes à produção vocal. Coincidentemente, o período posterior ao desenvolvimento instrumental registra um aumento significativo de publicações sobre o

tema, na tentativa de estabelecer tal validação, bem como de levantamento de indicações e limitações para sua aplicação ao âmbito clínico. Dois adventos marcaram especialmente o campo de avaliação das alterações vocais por meios não invasivos ao aparelho fonador: a concepção do espectrógrafo sonoro (Koenig et al, 1946) e o desenvolvimento da glotografia de alta freqüência (Fabre, 1957).

Quanto aos métodos para observação do aparelho fonador em movimento, destaca-se o da observação laríngea com auxílio de espelho de dentista em pleno século XIX, resultando na laringoscopia indireta e impulsionando técnicas futuras de endoscopia laríngea, as quais resultaram, cerca de um século depois, no desenvolvimento da fibra óptica para tal finalidade (Sawashima, Hirose, 1968) e demais métodos de investigação direta.

No final do século XIX, destaca-se o método de estroboscopia, voltado à observação do padrão vibratório das pregas vocais por meio de um efeito visual, criando a ilusão de movimentação lenta de pregas vocais (Hirano, 1981; Kitzing, 1985; Bless et al, 1987; Faure, Muller, 1992; Hirano, Bless, 1993). Entretanto, houve desenvolvimento anterior de método de observação do padrão vibratório de pregas vocais ciclo-a-ciclo: o aparelho para fotografia de alta velocidade da laringe humana também produzido nos *Bell Telephone Laboratories* (Timcke et al, 1958). Esse aparelho foi pouco utilizado na prática clínica devido aos custos.

Dessa forma, a estroboscopia acabou ganhando popularidade, apesar de não revelar o fenômeno ciclo-a-ciclo. Décadas à frente, foi desenvolvida a videoquimografia como tentativa de adaptação de método para avaliação do fenômeno ciclo-a-ciclo, entretanto em ponto restrito de prega vocal (Švec, Schutte, 1996; Tsuji, Sennes, 1998; Larsson et al, 1999, 2000; Tsuji et al, 1999a, 1999b; Tsuji et al, 2001).

Independentemente da concepção teórica e técnica empregada, existe unanimidade quanto à necessidade de compreensão mais detalhada do sinal que se apresenta irregular ao ouvido, tanto por parte dos clínicos e pesquisadores, como do próprio falante portador de alteração vocal.

Enquanto modalidades de avaliação não invasivas do aparelho fonador, registramos anteriormente o espectrógrafo sonoro e a eletroglotografia. Nesta

seção, nos ocuparemos de alguns detalhes da espectrografia acústica, enquanto a eletroglotografia será abordada no próximo capítulo, enquanto correspondência à composição de método de estudo proposto.

Retomando ao âmbito da espectrografia sonora, apesar das investigações puramente acústicas, mais precisamente no que tange à produção, propagação e recepção sonora, terem longa tradição (presentes inclusive em descrições de filósofos gregos), esse ramo de estudos até a década de 60 do século XX ainda era negligenciado, como demonstrou Lindsay (1966).

No que se refere especificamente à descrição do fenômeno acústico subjacente às alterações vocais, destaca-se a invenção do espectrógrafo sonoro no século XX por Koenig et al (1946), favorecendo uma nova e interessante incursão no campo de produção e percepção de sons produzidos pelo aparelho fonador humano, independentemente de manifestações de alterações. Na seqüência de tal evolução, registram-se as propostas de teorias e modelos que passaram a representar a grande possibilidade de respaldo para as análises, muitas dos quais caracterizam-se como referências até o presente momento (Flanagan, 1955, 1957, 1958, 1972; Fant, 1962b, 1970, 1973; Lindblom, Sundberg, 1971).

No campo das alterações vocais, destacam-se Isshiki et al (1969) quanto à classificação de rouquidão baseada em representações da espectrografia. A publicação abordou conceitos de aperiodidade, f0, formantes, índices de soprosidade e quebra de energia do formante. Apresentaram uma série de casos clínicos, traçando um paralelo entre o uso da espectrografia na área de voz e o uso do audiograma na avaliação auditiva.

A década de 60 registrou também outro importante estudo no que se refere à análise acústica e uso de recursos de síntese para análise do parâmetro de rouquidão, a partir da descrição de mudanças da estrutura harmônica e ruído, culminando numa proposta de classificação de quatro tipos de rouquidão de grande relevância ao campo de estudos das disfonias (Yanagihara, 1967).

Os achados comuns nas emissões de vogais [a],[ɛ], [i], [u] e [ɔ] referiram-se à presença de ruído nos formantes, ao aumento de componentes de alta

freqüência, especialmente acima de 3 kHz e à perda de harmônicos de alta freqüência. Tais achados mostraram-se proporcionais ao grau de rouquidão atribuído às amostras por meio da avaliação perceptiva.

Como resultado das inspeções, foi adotada uma escala, na qual o grau 1 corresponde a componentes harmônicos regulares mesclados a componentes de ruído concentrados nas faixas dos formantes; o grau 2 refere-se a componentes de ruído, predominando sobre os harmônicos, especialmente no segundo formante das vogais  $[\epsilon]$  e [i], com ruído adicional na faixa de freqüências altas (acima de 3 kHz) nas mesmas vogais; o grau 3 corresponde à ocorrência, nas vogais  $[\epsilon]$  e [i], de formantes totalmente substituídos por ruído, principalmente na faixa de freqüências acima de 3 kHz, caracterizada por energia intensa e amplamente distribuída; e, finalmente, o tipo 4 caracterizado por presença de ruído no segundo formante das vogais [a],  $[\epsilon]$  e [i]. Vale salientar que foi um estudo conduzido com vozes masculinas sem alterações, com adição progressiva de ruído por meio de síntese de fala (Yanagihara, 1967).

Outras contribuições na mesma década referem-se às publicações de Lieberman (1963), Moore, Thompson (1965) e Emanuel, Sansone (1969). Rontal et al (1975) e Hirano et al (1978) marcaram sua colaboração na década seguinte, com importantes contribuições para a avaliação de resultados de reabilitação vocal, cirúrgica e tratamento medicamentoso de alterações vocais.

Neste ponto da exposição, salienta-se a tendência inicial dos estudos acústicos voltados à modelagem do aparelho fonador masculino, especialmente pela limitação instrumental da época, a qual dificultava o uso de amostras de fala da população feminina e infantil. As referências posteriores chegaram a traçar hipóteses de possibilidade de diversidade de aspectos da teoria acústica da produção de fala para o aparelho fonador feminino (Titze, 1989). Dessa maneira, o estudo de vozes femininas e infantis representou durante muito tempo um verdadeiro desafio do ponto de vista instrumental e, por um outro aspecto, exigiu cuidadosa atenção, dada a concentração de indivíduos do sexo feminino portadores de alterações vocais (Askenfalt, Hammarberg, 1986), incluindo-se neste item os aspectos de conformação glótica durante a fonação, os quais favoreceriam a ocorrência de fenda glótica durante a fonação de indivíduos do sexo feminino.

Na evolução dos estudos neste campo particular das alterações vocais foi possível detectar a tendência de desenvolvimento de várias propostas de avaliação vocal com base na percepção auditiva de juízes treinados, as quais estão, em sua maioria, centradas nos aspectos laríngeos da produção vocal, provavelmente, porque essas propostas derivaram de trabalhos, desenvolvidos com equipes de otorrinolaringologia, que analisaram casos de alterações glóticas de origem orgânica e funcional que demandavam tratamento medicamentoso, fonoterápico e/ou cirúrgico.

Tais propostas ganharam popularidade, destacando-se a escala *GRBAS* (na correspondência das letras iniciais de *Grade, Rough, Breathy, Asthenic* e *Hyperfunctional*) para avaliação da rouquidão (Isshiki et al, 1969; Hirano, 1981) e a proposta *SVEA, Swedish Voice Evaluation Approach* (Hammarberg et al, 1980; Hammarberg, Gauffin, 1995; Hammarberg, 2000), ambas respaldadas pelo método diferencial semântico (Osgood et al, 1969). Grande mérito deve ser atribuído ao trabalho dos referidos grupos especialmente por alertarem os clínicos e pesquisadores para a necessidade de sistematização de meios de avaliação da qualidade vocal.

Em outra perspectiva, destaca-se a consideração dos aspectos laríngeos e supralaríngeos de qualidade sonora resultante, abordando a ação, ou mais precisamente a constelação de ações no aparelho fonador, com sustentação do modelo fonético da qualidade vocal no enfoque do *VPAS* (Laver et al, 1981; Beck, 1988; Laver, 1994; Laver, 2000). Nesse âmbito, também podemos finalmente referir propostas de avaliação clínica que contemplaram os aspectos supralaríngeos, enquanto mecanismos de articulação e ressonância (Wilson, 1994; Brandi, 1996).

Quanto ao estímulos para avaliação perceptiva, destacaram-se inicialmente o predomínio de emissões de vogais sustentadas, e, gradualmente a concentração de sua indicação para avaliação da eficiência de mecanismos laríngeos (Takahashi, Koike, 1975; Sorensen, Horii, 1984; Askenfalt, Hammarberg, 1986; Titze et al, 1987; Feijoo, Hernandez, 1990; Rydell et al, 1995), reconhecendo-se a necessidade de inclusão das emissões encadeadas, propostas em estudos de longo termo da voz (Lieberman, 1963; Gauffin, Sundberg, 1977).

As controvérsias quanto ao tipo de estímulo a ser utilizado novamente expõem a divergência dos conceitos subjacente à produção do sinal vocal. Conforme exposto, durante várias décadas, figuraram as propostas de vogais sustentadas, numa alusão à sua representatividade de aspectos de controle de atividade laríngea, e pelo fato de não apresentarem variações de f0 e de intensidade dependentes de fatores prosódicos (Iwata, Von Leden, 1970; Horii, 1979).

Em contrapartida, ou mais precisamente pelo mesmo motivo, Askenfalt, Hammarberg (1986) defenderam o uso de emissões encadeadas como representativas da produção vocal, especialmente por contemplarem tais modulações, de caráter altamente informativo de qualidade vocal. Portanto, para as avaliações baseadas nas propostas de avaliação conhecidas *SVEA* e *VPAS*, anteriormente referidas, destacou-se a necessidade de estímulos de fala encadeada. Para a escala *GRBAS*, as propostas iniciais referiam uso exclusivo da emissão sustentada da vogal [a], porém gradativamente também passaram a incorporar a possibilidade de uso em emissões encadeada (Piccirillo et al, 1998).

Tais aspectos sobre as propostas de instrumentos de avaliação perceptiva são detalhados na fundamentação do método de avaliação da qualidade vocal com motivação fonética a ser utilizado na presente pesquisa (capítulo 4.2). Neste tópico restringiremos nossas considerações às correspondências entre as esferas acústica, perceptiva e fisiológica.

### 3.2. Perspectivas recentes

Da evolução de tendências anteriormente delineadas em relação à avaliação perceptiva, emerge a condição multidimensional da qualidade vocal. Gerratt et al (1993), Kreiman, Gerratt (1995) e Gerratt, Kreiman (2000) apresentaram os resultados de seqüência de estudos voltados à validação de escalas perceptivas de avaliação vocal e traçaram várias observações a respeito da percepção da voz humana. Revelaram que as dimensões percebidas da voz não corresponderam, no espaço perceptivo, a escalas contínuas, mas resultaram em conjuntos que não partilhavam das mesmas unidades subjetivas. Em função dos achados, criticaram a maior parte das propostas de avaliação perceptiva, apesar de reconhecerem a

complexidade dos parâmetros envolvidos no julgamento auditiva da qualidade vocal (Gerratt, Kreiman, 2000).

Deste cenário, também emergem concepções distintas nas abordagens temporais da emissão do indivíduo portador de alteração vocal. Se considerarmos que a voz (ou mais precisamente a fala) manifesta-se numa sucessão de eventos temporais altamente integrados e precisos, surgem propostas distintas nos aspectos de curto termo e de longo termo das emissões.

Ao longo da presente exposição, começa a ser delineado um panorama no qual a abordagem das disfonias remete a um ponto divergente de concepção de qualidade vocal subjacente às propostas de avaliação apresentadas. Tal controvérsia ou, em muitas ocasiões, imprecisões, favorecem abordagens distintas e por vezes limitadas da alteração vocal. Na maior parte dos estudos, a qualidade vocal alterada tem como base a referência a um conceito ideal de normalidade, sendo descrita a partir de evoluções em escalas lineares de alteração. Com o advento da análise acústica, tais abordagens foram reforçadas por uma verdadeira proliferação de medidas e índices acústicos voltados à detecção de irregularidades e tentativas de comparação a duvidosos critérios de normalidade, os quais deveriam ser cuidadosamente aplicados ao contexto clínico (Gerratt, Kreiman, 2000; Kreiman, Gerratt, 2000).

Desta exposição pode-se compreender que o sinal acústico do falante disfônico passou a ser graduado em função da alteração ou perturbação do sinal, subjugando a estrutura sonora básica da emissão a um plano secundário. Tal enfoque limitou a possibilidade de se analisar a correspondência dos aspectos auditivos, acústicos e fisiológicos da qualidade vocal do indivíduo portador de disfonia. Diante de tais observações, Gerratt, Kreimann (2000) ressaltaram que o caráter multidimensional da qualidade vocal foi erroneamente confundido com a possibilidade de associação da qualidade vocal a um conjunto de medidas e seus desvios considerados em relação a um grupo adotado como padrão de normalidade, num conceito que faz alusão a um ideal de qualidade vocal.

Em tal perspectiva de abordagem de alterações vocais, algumas correspondências foram traçadas para unir as dimensões auditiva, acústica e fisiológica, porém, em sua maioria, mantendo referência a dicotomia normalidade e

alteração vocal. Nesse universo, figuram várias tentativas de quantificação de parâmetros espectrais, especialmente centradas na dimensão de curto termo da emissão. Registraram-se, inicialmente, tentativas de quantificação do ruído e de sua relação, principalmente ao parâmetro perceptivo de rouquidão, traduzidas em várias estimativas de ruído ou de irregularidade vocal (Kojima et al, 1980; Kitajima, 1981; Yumoto et al, 1982; Yumoto, 1983; Hiraoka et al, 1984; Yumoto et al, 1984; Klingholz, Martin, 1985; Hirano et al, 1986; Kasuya et al, 1986; Hillenbrand, 1987; Cox et al, 1989a, 1989b; Sasaki et al, 1991; Qi, 1992; Deliyski, 1993; Awan, Frenkel, 1994).

Na seqüência, registraram-se as tentativas de correspondências entre medidas de f0 e de perturbações em freqüência (*jitter*) e amplitude (*shimmer*) e parâmetros de avaliação perceptiva de rouquidão, aspereza e soprosidade (Eskenazi et al, 1990; Feijoo, Hernandez, 1990; Toner et al, 1990; Rabinovic et al, 1995; Scherer et al, 1995; Bielamowicz et al, 1996; Dejonckere et al, 1996; Wyuits, 1996). Apesar de algumas referências a correlações positivas detectadas, os autores apontaram limitações quanto ao estabelecimento de relações estreitas entre as dimensões acústica e perceptiva (Eskenazi et al, 1990; Feijoo, Hernandez, 1990).

Em contrapartida, McAllister et al (1996) revelaram as dificuldades de estabelecimento de correlação entre parâmetros auditivos de disfonia e medidas de perturbação em amostras de fala contínua de crianças, ao contrário do que estaria sendo estabelecido para emissão vocal de adultos no período acima delineado, à semelhança de achados de Askenfalt, Hammarberg (1996) e Camargo (1996).

No mesmo período, tratando especificamente da abordagem das medidas de perturbação, ocorria o trabalho de uma equipe centrada nas tentativas de refinamento dos métodos de extração de medidas de perturbação e sua aplicabilidade na detecção de alterações laríngeas. Adotaram a eletroglotografia como forma mais segura de captação de indícios de atividade glótica, resultando em instrumentos de análise adotados no presente estudo (Vieira et al 1996a, 1996b; Vieira, 1997; Vieira et al, 1997).

Vieira et al (1997) salientaram que poucos estudos analisaram comparativamente o *jitter* acústico e *jitter EGG* com o objetivo de avaliar a validade da medida para vozes alteradas. Analisaram os traçados de onda eletroglotográfica e acústica em emissões das vogais [a], [i] e [u] de 15 indivíduos disfônicos para avaliação da possibilidade de extração de *jitter*, propondo método combinado de demarcação do sinal. Procederam classificação inicial da traçado da onda, a fim de delimitar ou restringir a possibilidade de extração de medidas e encontraram maior confiabilidade do método para as emissões da vogal [a].

Tal refinamento metodológico permitiu que as medidas de perturbação, especialmente *jitter EGG* (eletroglotográfico), revelassem o refinamento de atividade laríngea em casos movimentos anormais da cartilagem aritenóide. Neste caso, a extração de valores de *jitter EGG* forneceu indícios para a condução do diagnóstico (Vieira, 1997; Vieira et al, 1997). A continuidade de estudos dessa dimensão poderá prover futuras implementações neste complexo campo de demarcação de sinais com aperiodicidade, oferecendo uma nova perspectiva para a abordagem de medidas acústica no campo das disfonias.

O parâmetro perceptivo de soprosidade vocal também recebeu atenção quanto à abordagem de parâmetros acústicos espectrais como é caso da presença de energia em harmônicos de alta freqüência (Hirano et al, 1986; Klatt, Klatt, 1990; Shoji et al, 1991; Hillenbrand et al, 1994; Naranjo et al, 1994).

Nas análises de curto termo, também destacaram-se as possibilidades de análise pelo procedimento *cepstrum*, portanto análise cepstral, como recurso promissor da abordagem da estrutura harmônica do som (Feijoo, Hernandez, 1990; De Krom, 1993; Dejonckere, Wienecke, 1994; Dejonckere, 1998; Awan, 2002; Heman-Ackah et al, 2002).

Muitos autores abordaram os componentes espectrais das alterações vocais, estabelecendo correlações entre as diversas qualidades vocais e padrões de distribuição de energia acústica (De Krom, 1995; Martin et al, 1995; Dejonckere et al, 1996; Wolfe, Martin, 1997; Fröhlich et al, 1998).

Gauffin, Sundberg (1980) revelaram a importância de medidas como a amplitude do fundamental (L0) no espectro e sua correlação à amplitude vibratória e, conseqüentemente, com a amplitude dos harmônicos. No caso da voz com escape de ar, descreveram L0 maior que a amplitude do primeiro formante (L1). Na voz tensa, ao contrário, L0 aparece menor que L1. Na fonação considerada ótima, seria esperado o equilíbrio entre L0 e L1. Os autores propuseram a abordagem da razão L0/L1 para refletir tais particularidades da ação glótica. Chamaram a atenção para as variações de longo termo da intensidade vocal.

Nessa linha de estudos, registram-se várias propostas de consideração das proporções de amplitude de componentes espectrais, especialmente dos dois primeiros harmônicos (H1 e H2) e dos formantes, especialmente F1 e F3 (Holmberg et al, 1995). Também ocorre a busca por estimativas de variações de amplitude em faixas de freqüências diferenciadas do sinal em amostras de fala encadeada de longa duração. Atualmente, também são aplicadas às alterações vocais e serão abordadas em item específico no próximo capítulo.

Ainda na dimensão de curto termo, o tempo de início de sonorização (VOT) também foi abordado em relação a alterações vocais decorrentes de comprometimentos de mecanismos glóticos (Camargo, 1998; Andrade, 2000; Gregio, 2000) em pesquisas desenvolvidas no LIAAC (PUC/SP). Sua aplicação esteve restrita aos estudos apontados no campo das alterações vocais, entretanto denota possibilidade futura de maior aplicação, dada sua relevância em relação a eventos das esferas respiratória (Hoit et al, 1993), glótica e supraglótica (Gandour, Dardarananda, 1984; Higgins et al, 1998; Ryallls et al, 1999; Madureira et al, 2002). Além disso, descrições referentes ao envelhecimento (Decoster, Debruyne, 1997) e às perspectivas de conjugação de técnicas de fluxo e eletroglotográficas (Higgins et al, 1998) sinalizam possível tendência de incorporação da abordagem do índice de VOT para a situação de disfonias.

Neste âmbito de perspectivas futuras, deve também ser registrada a consideração da plasticidade vocal como parâmetro indicativo da possibilidade de desenvolvimento compensatório para as manifestações da disfonia (Dejonckere, Lebaq, 2001). Tal consideração, por sua vez, nos remete à situação eleita para este estudo, as falhas de mobilização glótica por paralisia unilateral ou fixação de pregas

vocais, enquanto possibilidades de caracterização dos mecanismos compensatórios subjacentes à manifestação. Nesse sentido, Kim et al (1982) ressaltaram a importância de aspectos espectrais relativos à flutuação de f0, componentes harmônicos de alta freqüência e nível de ruído e freqüência de F1. Hammarberg et al (1984) reforçaram os aspectos de longo termo, além da distribuição de f0 e das taxas de perturbação da forma da onda acústica.

Alguns estudos centraram-se na avaliação dos resultados funcionais decorrentes de tratamentos cirúrgicos propostos para cada fase dos avanços referidos na área, destacando-se relatos dos resultados vocais após injeção intracordal de *Teflon* (Fritzell et al, 1974; Hammarberg et al, 1984; D'Antonio et al, 1995; Livesey, Carding, 1995), silicone (Hirano et al, 1995) e, mais recentemente, das cirurgias de arcabouço laríngeo, como é o caso das tireoplastias (D'Antonio et al, 1995).

No âmbito das investigações fisiológicas nesta população, em especial, destacaram-se estudos eletromiográficos (*EMG*) (Lindestad, Peasson, 1994), métodos de diagnóstico por imagem (Ooi et al, 1995), além do enfoque compensatório de atividade de pregas vestibulares (Pinho et al, 1999) revelando o deslocamento de pregas vestibulares em todos os casos estudados, porém sem atividade vibratória. Hirano, Mori (2000) apresentam síntese de vários achados anteriores para casos de paralisia de prega vocal nas esferas auditiva, acústica e fisiológica (aerodinâmica e estroboscópica).

Diante das considerações anteriores e como menção final do tópico de revisão crítica da avaliação das disfonias, ressaltamos a complexidade envolvida nas correlações entre as esferas acústica, perceptiva e fisiológica da qualidade vocal, a qual envolve o respaldo de consistentes correntes teóricas. Nesse campo, Colton, Estill (1981) enfocaram tal nível de correlação e apontaram propostas caracterizadas como métodos perceptivos (espaço multidimensional), distintivos (traços distintivos, como o método diferencial semântico) ou multidimensionais, no sentido de explicação das relações entre os correlatos acima delineados, enfocados pela teoria acústica da produção da fala, particularmente no modelo fonte-filtro para a produção das vogais (Fant, 1970). Salientaram a importância da adoção de modelos teóricos, como forma de aprofundar as descrições entre as três dimensões envolvidas na qualidade vocal (acústica, percepção e fisiologia).

A importância da adoção de modelos para enfoque de qualidade vocal com vistas ao estabelecimento de correlações entre percepção e demais níveis de descrição do sinal sonoro residem em se estimar os fatores que realmente seriam relevantes à percepção auditiva. Nesse âmbito, Zwicker, Fastl (1999) exploraram modelo para estudo de percepção da aspereza, enquanto Bergan, Titze (2001) investigaram o mesmo parâmetro perceptivo em sinais com subharmônicos. Outros estudos sobre os componentes subharmônicos foram desenvolvidos por Fuks et al (1988), Cavalli, Hirson (1999) e Fuks (1999).

Sintetizando parte dos estudos sobre a temática de descrições perceptivas e acústicas da voz, com fundamentação fisiológica, Hammarberg, Gauffin (1995) traçaram um importante paralelo entre correlatos perceptivos, acústicos e fisiológicos de alterações vocais com comprometimento glótico, servindo de base para vários estudos nesse campo de estabelecimento de correlatos entre os três níveis de produção sonora. Apontaram a necessidade de refinamento das amostras de fala como uma possibilidade para o avanço no estabelecimento de correlações mais consistentes.

# 4. CONCEPÇÃO DE ABORDAGEM INTERPRETATIVA E INTEGRADA DO SINAL VOCAL NA VIGÊNCIA DE DISFONIAS

Diante da proposta de investigação das correlações entre parâmetros acústicos, perceptivos e fisiológicos na manifestação de disfonia, este capítulo destina-se à fundamentação dos métodos de avaliação da qualidade vocal incorporados ao presente estudo, trazendo ao campo fonoaudiológico as considerações teóricas e práticas de áreas correlatas, as quais começaram a ser exploradas de forma mais consistente em nosso meio apenas recentemente. Neste tocante, a divisão por seções visa contemplar a exploração de uma das modalidades de avaliação de cada dimensão da qualidade vocal. Nessa concepção, a dimensão acústica foi representada pela análise de longo termo, a perceptiva pelo roteiro de avaliação de qualidade vocal com motivação fonética e, finalmente, a fisiológica pela técnica da eletroglotografia (EGG).

O tópico dedicado à avaliação da qualidade vocal na dimensão perceptiva apresenta a adaptação de roteiro de avaliação empregado na pesquisa, de forma que a abordagem dos fundamentos teóricos do modelo fonético da qualidade vocal (Laver, 1980) é dimensionada para a rotina clínica. Para cumprir tal tarefa, realizamos análise comparativa a outros métodos de avaliação que ganharam tradição em termos de avaliação vocal perceptivo-auditiva.

A opção pela explanação mais detalhada de um dos procedimentos de avaliação referente a cada uma das esferas envolvidas na incursão à qualidade vocal, em situação de alterações de mecanismos glóticos, visa fornecer subsídios à apreciação da estruturação de métodos, bem como de discussão dos resultados obtidos. Além disso, configura-se como uma contribuição para incentivar futuros estudos nessa linha de pesquisas, em que o fonoaudiólogo necessita do conhecimento mais refinado de vários procedimentos explorados de forma rotineira no âmbito das Ciências da Fala e da Fonética Acústica.

### 4.1. Análise acústica de longo termo

A representação típica deste âmbito de análise é o espectro de longo termo, ELT, numa versão abreviada da tradução da sigla *LTAS* (*Long-Term Average Spectrum*). Representa a somatória média de uma série de espectros de curto termo gerados pelo método *FFT* (*Fast Fourier Transform*), no processamento de uma amostra extensa o suficiente para não se restringir a características específicas de uma certa classe de segmentos fônicos. Tais considerações levam em conta os achados de estabilização do traçado do ELT a partir de um determinado intervalo de tempo, referido como situado entre 15 e 30 segundos (Pittam, 1987).

A qualidade vocal passa a ser enfocada de acordo com a variação de amplitude em faixas seletivas de freqüência de amostras de fala com duração superior a 15 segundos, nas quais a influência dos aspectos segmentares seria mínima. Nessa situação, o estímulo de fala analisado pode incluir desde a leitura de um texto até trechos de emissão espontânea. Alguns estudos destacaram particularidades na composição da amostra, com propostas variando desde fala espontânea (Pittam, 1987), canto (Cleveland et al, 2001; White, 2001), ao uso de texto padronizado contendo sons surdos e sonoros (Wendler et al, 1986; Pittam, 1987; Figueiredo, 1993; Ananthapadmanabha, 1995) e planejamento para sua composição exclusiva por trechos sonorizados (Dejonckere, Villarosa, 1986; Barrichelo et al, 2001). Outra possibilidade registrada é a exclusão de trechos não sonorizados e pausas (Frøkjær-Jensen, Prytz, 1976; Hammarberg et al, 1980, 1984; Löfqvist, 1986; Hammarberg, Gauffin, 1995; Mendonza et al, 1996). Dependendo do propósito de estudo, podem ainda ser utilizadas modalidades distintas de tratamento e filtragem de sinal (Pittam, 1994). Para Pittam, Millar (1988) porém o ponto central seria a necessidade de alcançar a estabilização dos parâmetros de longo termo, no sentido de se enfocar a duração das amostras analisadas e não propriamente a composição do estímulo de fala em termos de predominância sonora.

O cuidado na composição da amostra a ser estudada encontra respaldo no fato de que foi inicialmente relatada a maior correlação dos aspectos de longo termo aos ajustes laríngeos da emissão, destacando-se, dessa forma, a

informação primária da vibração de pregas vocais. Deve-se também recordar que o procedimento de FFT utilizado para análise acaba por favorecer a decomposição do sinal, numa clara referência à atividade glótica. As particularidades das modalidades de composição e tratamento da amostra serão discutidas mais adiante neste texto, quando da abordagem das variadas indicações e aplicações ao âmbito da clínica fonoaudiológica, especialmente por não existir consenso e, muito menos, padronização do procedimento de análise de longo termo.

Dessa forma, vários estudos buscaram a correspondência das variações de ajustes laríngeos e supralaríngeos aos traçados e análises numéricas provenientes do ELT. Alguns achados foram divergentes, porém predominou a tendência de maior influência de variações de ajustes laríngeos na conformação geral e medidas do ELT (Nolan, 1983, Kitzing, 1986; Pittam, 1987; Pittam, Millar, 1988; Pittam, 1994; Mendonza et al, 1996). Alguns ajustes supralaríngeos também foram revelados por tal representação espectral, porém em menor proporção na literatura dominante nas décadas de 70 e 80. Tal situação se inverte em décadas posteriores, quando as ressonâncias do trato vocal passam a ser comumente referidas em estudos de diversas populações, cobrindo diferentes faixas etárias e qualidades vocais (Figueiredo, 1993; Hammarberg, Gauffin, 1995; Cleveland et al, 2001; Linville, Rens, 2001; White, 2001). Estudos de longo termo foram realizados, inclusive, em amostras da gravação que acompanham a publicação original do modelo fonético de descrição da qualidade vocal (Nolan, 1983; Ananthapadmanabha, 1995).

Apesar de difusão ainda restrita em nosso meio, especialmente do ponto de vista de aplicação clínica, as análises de longo termo aplicadas às amostras de fala humana remontam ao final da década de 20 do século passado, com a descrição da análise espectral de fala contínua (Pittam, Millar, 1988). Tal desenvolvimento coincidiu com período anterior ao grande desenvolvimento do instrumental da acústica, impulsionado pela concepção do espectrógrafo sonoro, que viria a ocorrer somente em meados da década de 40. Dessa forma, suas referências também são anteriores à descrição da teoria acústica da produção da fala (Fant, 1970).

Os primeiros estudos estiveram relacionados à produção de fala, com vistas ao desenvolvimento de tecnologia para telecomunicações. No curso da

segunda guerra mundial passaram a ser aplicados a novas situações, tais como a investigação dos efeitos de altas altitudes na voz e na fala, além de distorções induzidas pelo uso de máscaras de O<sub>2</sub> e outros gases. A partir de 1950, as publicações aumentaram, iniciando-se na década de 80 o período de maior vigor nas pesquisas voltadas ao estudo das alterações do sinal de fala com possível aplicação clínica.

Pittam; Millar (1988) traçaram a história do desenvolvimento da medida de ELT apresentando uma revisão crítica das tentativas de quantificação dos padrões acústicos representados no ELT. Os parâmetros considerados essenciais referiram-se ao tempo de integração da energia espectral, à faixa de freqüência para as medidas de energia e à resolução de freqüência da medida.

Diante da profusão de estudos, várias aplicações possíveis do ELT foram destacadas, tais como a identificação do falante, a educação e a reabilitação vocal. Neste último aspecto, o controle da eficácia do tratamento de várias situações clínicas, tais como estimativas do grau de severidade das disfonias em casos de paralisias laríngeas, disfonias espasmódicas, TCE (traumatismo crânio-encefálico), efeitos de treinamento vocal ou ainda estudos da voz esofágica (Frøkjær-Jensen, Prytz, 1976; Wedin, Örgren, 1982; Hartmann, Cramon, 1984; Dejonckere, Villarosa, 1986).

A aplicação clínica do ELT é reforçada especialmente por possibilitar a documentação da evolução, sendo mais limitado como meio diagnóstico ou de triagem (Hartmann, Cramon, 1984; Löfqvist, 1986; Pittam, Millar, 1988). Apesar dessa tendência, alguns estudos reforçaram a possibilidade de diferenciação entre vozes normais e com alterações (Frøkjær-Jensen, Prytz, 1976; Hammarberg et al, 1980; Dejonckere, Villarosa, 1986; Hammarberg, Gauffin, 1995), incluindo a indicação da possibilidade de distinção entre indivíduos normais e portadores de várias categorias de alterações laríngeas, inclusive benignas e malignas (Dejonckere, Villarosa, 1986).

Como pode ser depreendido da exposição anterior e das referências apresentadas, as informações do ELT, enquanto ferramentas de avaliação clínica, tiveram inicialmente uso restrito quando comparadas às abordagens de curto termo, tais como medidas de f0, de perturbação e ruído, referidas no capítulo anterior.

Tendência interessante, que deve ser assinalada, refere-se ao

declínio das referências na década de 90, com a maior parte concentrada em final da década de 70 e ao longo dos anos 80. Finalmente voltam a ser referidas no início dos anos 2000. Tal trajetória defere daquela de medidas de curto termo anteriormente referidas, as quais mantiveram-se em franca proliferação. Notadamente a representação acústica de longo termo encontra maior repercussão em área de estudos de Música, Fonética Forense e da voz na interação social.

Nesse sentido, o falante é aqui considerado como capaz de exercer controle sobre certos aspectos de qualidade de longo termo (extrínsecos), enquanto outros (intrínsecos) estariam além da possibilidade de seu domínio volitivo (Nolan, 1983).

No campo do reconhecimento de falantes, o grande centro de interesse do autor, procurou caracterizar aquelas dimensões de qualidade vocal que seriam determinadas pela constituição física do aparelho fonador daquele falante em questão e portanto não apresentariam variação em função dos contrastes de intenção comunicativa. Para tal finalidade, a descrição do traçado do ELT foi considerada representativa das mudanças da qualidade vocal com influência de atividade laríngea.

Alguns estudos desenvolvidos especificamente no campo das alterações vocais apresentam as bases da técnica em questão, caracterizando-se como referências importantes ao enfoque da aplicação clínica do ELT. Nesse âmbito, Frittzell et al (1974) descreveram características vocais após injeção de teflon para paralisias laríngeas, reforçando que o declínio da energia espectral registrado no ELT acima de 3 kHz depende basicamente das características acústicas do som gerado pela vibração de pregas vocais. Numa emissão normal e eficiente, o declínio de energia tende a ser suave se o recurso de pré-ênfase for aplicado. No caso da ocorrência de escape de ar pelo fechamento glótico incompleto, as pregas vocais geram um espectro contendo um fundamental, porém com menos parciais e de menor amplitude. Sob tais condições, o traçado ELT exibe claramente um pico na região do fundamental e um declínio acentuado em freqüências acima do pico. Além disso, é gerado ruído, o qual passa a excitar o segundo formante, de forma que um pico pode ser esperado na região correspondente a F2, com amplitude em relação àquela do fundamental.

Como forma de padronizar as observações relativas à conformação do traçado espectral, mais especificamente a inclinação ou declínio espectral, foram propostas estimativas de energia espectral por faixas de freqüência. Nesse sentido, o parâmetro alfa (α) ganhou popularidade na década de 70, ao refletir a razão entre energia espectral acima de 1-3 kHz e 0-1 kHz (Frøkjær-Jensen, Prytz, 1976; Frøkjær-Jensen, 1979), o qual seria sugestivo do grau de tensão do trato vocal, especialmente da compressão medial.

No desenvolvimento subseqüente, muitos estudos enfocaram tal proporção, com propostas de variações em termos faixas de freqüências abordadas, as quais refletem-se nas várias descrições de variações de energia por faixas de freqüência do sinal.

Na seqüência, Hammarberg (1980) e Hammarberg et al (1980) apresentaram resultados de trabalhos conduzidos em clínica de instituição hospitalar voltados à validação de procedimento para avaliação perceptivo-auditiva, o qual futuramente resultaria na proposta de avaliação perceptiva conhecida como *SVEA*. Como forma de estabelecer as correlações entre os vários descritores de qualidade vocal utilizados na rotina de atendimento, também incorporaram o ELT. Para as qualidades descritas como tensa, sonora e estável a amplitude espectral apresentou-se elevada nas faixas de 0-2 e 2-5 kHz, com declínio abrupto em 5-8 kHz. No caso da voz soprosa, foi registrado aumento entre 5-8 kHz e queda abrupta entre as bandas de 0-2 kHz e 2-5 kHz. Tais dimensões estiveram amplamente relacionadas à conformação do ELT, enquanto correspondência aos respectivos descritores de qualidades hiperfuncional e hipofuncional.

Numa sucessão de estudos em alterações vocais Dejonckere, Cuvelier (1983); Dejonckere (1983; 1984) e Dejonckere, Villarosa (1986) identificaram, dentre várias faixas do sinal estudado, aquelas que separam vozes normais e com alterações e correspondência significativa com os tipos de rouquidão classificados acusticamente por Yanagihara (1967). O parâmetro relevante foi descrito como razão W, representando a razão entre a energia espectral de 0-6 kHz e 6-10 kHz.

Nolan (1983) propôs uma modificação no parâmetro alfa de forma a calcular a razão da energia espectral nas faixas de 0-1,5 kHz e 1,5-3 kHz e mediu o

declínio por oitava na faixa de 0 a 2500 Hz, diferenciando subcategorias de ajustes fonatórios, porém novamente não aqueles da esfera supralaríngea. Tais resultados foram questionados por Pittam (1987) por não terem sido submetidos a tratamento estatístico.

No mesmo ano, Kitzing (1986) enfocou as alterações vocais de outro ponto de vista: aquelas produzidas em órgãos sãos, por meio do processamento de amostras de fala encadeada de 10 fonoaudiólogos experientes, reproduzindo quatro diferentes qualidades reconhecíveis. Dentre várias medidas estimadas, a proporção de energia entre o fundamental f0 e o primeiro formante diferenciou as qualidades vocais enfocadas. Algumas variações também foram encontradas na região de 3 a 4 kHz, as quais poderiam corresponder a ressonâncias do trato vocal, como descrito em estudos de canto, por se referir à concentração do formante do cantor.

Löfqvist (1986) encontrou correlação entre maiores valores de razão entre 0-1/ 1-5 kHz com hipofunção e os menores com hiperfunção, reforçando achados anteriores, inclusive com relação ao aumento de concentração ente 5-8 kHz a componentes de ruído em voz hipofuncional. Despertou a atenção dos pesquisadores com relação à possibilidade de mudanças nos padrões de ELT ao longo do dia.

Tais achados foram semelhantes aos relatados posteriormente por Hammarberg, Gauffin (1995), apenas com ligeira mudança em relação às faixas estudadas, por incluírem as faixas de 0-2 kHz e de 2-5 kHz. Outras mudanças relatadas referem-se à ampliação para freqüências mais agudas (6-10 kHz e 10-16 kHz) possibilitando a distinção entre indivíduos disfônicos e não disfônicos (Shoji et al, 1991).

Figueiredo (1993), em estudo voltado à identificação de falantes, propôs novas medidas, tais como declínios de retas ajustadas às faixas definidas (método dos mínimos quadrados), obtendo o decréscimo da energia espectral por oitava. Cada ELT foi reduzido a seis declínios, nas faixas de 0-0,5 kHz, 0,5-1 kHz, 1-2 kHz, 2-3,5 kHz e 5-8 kHz. Apesar de restringir a informação de longo termo a seis variáveis, tais estimativas foram capazes de separar os falantes a partir de análise estatística para composição de *clusters*.

Por outro lado, Pittam (1994) reafirmou a limitação que o uso de medidas e estimativas pode impor aos aspectos de longo termo da emissão. Neste ponto de exposição e discussão do tema, destaca-se a possibilidade de abordagem da vasta gama de dados oferecidos pelo processamento da amostra de fala por modalidade de longo termo. Independente da indicação e aplicação específica, emerge a necessidade de cuidadoso tratamento da informação obtida, especialmente do conjunto de dados numéricos que refletem as variações de energia das amostras processadas. Nesse aspecto, Nolan (1983) oferece ao leitor a possibilidade de descrição e discussão dos aspectos de longo termo da qualidade vocal, bem como a ilustração de uma variedade de traçados de ELT, a qual facilita a incursão inicial do pesquisador no tema.

Pittam (1994) estabeleceu relações entre os padrões acústicos de longo termo e atributos perceptivos ao enfocar a qualidade vocal no contexto social. Diante de tal abordagem seria possível comparar medidas acústicas com indicadores de percepção auditiva, fato este um tanto criticado por autores que estudam detalhadamente a percepção de qualidade vocal (Kreiman, Gerratt, 1995, 1996, 2000; Gerratt, Kreiman, 2000).

Anteriormente, Pittam (1987) investigou séries de intervalos de freqüência relevantes do ponto de vista perceptivo (intervalo fixo em hertz), os quais poderiam direcionar a modalidade de tratamento estatístico dos dados provenientes do ELT. Vários intervalos situados entre 125 e 500 Hz foram avaliados nas faixas de 0-2 kHz e 0-3 kHz, dada sua relevância lingüística e, portanto, perceptiva. A melhor diferenciação ocorreu para intervalos das escalas 1,5-Bark e 200-Mel, respectivamente correspondentes a 8 valores na extensão de freqüência de 0 a 2152 Hz e a 7 valores na extensão de freqüência de 0 a 2031 Hz, os quais, inclusive, discriminaram ajustes supralaríngeos, exceto a nasalidade. O intervalo fixo de 200 Hz também foi proposto como intervalo útil para tal finalidade de análise de longo termo.

Nesse aspecto, Ananthapadmanabha (1995) não detectou relação entre a conformação do ELT e a qualidade vocal percebida, para amostras manipuladas por meio de procedimentos de síntese. Para as qualidades relacionadas a mobilizações laríngeas, os fatores determinantes de qualidade vocal estiveram relacionados aos níveis de energia dos formantes mais agudos e de f0. Para os ajustes

articulatórios, os formantes representaram a dimensão acústica saliente.

Mendonza et al (1996) identificaram diferenças entre vozes femininas e masculinas pelo ELT em função da presença de ruído próximo e à diminuição da inclinação espectral entre 0-1/1-5 kHz, graças à presença de muito ruído na região do terceiro formante (F3). Nestes casos, relataram qualidade vocal eminentemente soprosa.

Figueiredo (1993), em estudo do âmbito da fonética forense em nosso meio, ressaltou a possibilidade de detecção de correspondências aos ajustes supralaríngeos, desde que alguns cuidados sejam adotados na análise. O cômputo de todos os pontos e o seu devido tratamento com base em correlações permite a investigação da estrutura fina a revelar tais componentes.

Em publicações recentes, detecta-se a tendência à retomada das abordagens de longo termo em variadas situações que envolvem a atuação fonoaudiológica, os quais cobrem desde a produção vocal na infância (White, 2001) até a senescência (Linville, Rens, 2001), com possibilidades de separação do sexo para os ambos os extremos de faixas etárias. Neste aspecto, White (2001) foi capaz de diferenciá-los com base na localização dos picos espectrais, nos quais a região de 5 kHz revelou um pico para meninos e apresentou-se achatada para as meninas. Com abordagem semelhante para diferenciação de vozes masculinas e femininas com base nas previsões de ressonância do trato vocal, Linville, Rens (2001) detectaram menor amplitude dos picos do ELT na população feminina, relacionando-a às mudanças diferenciadas no trato vocal, em que as mulheres tendem a apresentar o alongamento mais acentuado.

As referências estendem-se às investigações comparativas entre situações de canto e fala (Barrichelo et al, 2001; Cleveland et al, 2001). Traçam interessantes considerações a respeito das possibilidades de assimilação de alguns ajustes do canto à situação de fala.

Diante de tantos aspectos impressos nas representações de longo termo da emissão, destacamos a sua validade enquanto possibilidade de enfoque de correlatos acústicos da qualidade vocal, investigada neste estudo sob a perspectiva fonética, cujo roteiro passa a ser detalhado no próximo tópico desta seção.

### 4.2. Roteiro de avaliação da qualidade vocal com motivação fonética

A motivação básica da proposta de avaliação perceptiva da qualidade vocal a ser apresentada refere-se ao modelo fonético de descrição da qualidade vocal (Laver, 1980). Algumas considerações sobre atualizações da proposta *VPAS* são delineadas como forma de compreendermos a adaptação do instrumento de avaliação para a pesquisa em questão.

Neste aspecto particular, vale ressaltar que as observações de achados ou adaptações implementadas ao *VPAS* por Laver et al (1981) e Laver (1991, 1994, 2000) encontram correspondência com os aspectos explorados por outras propostas de escalas voltadas à descrição mais detalhada dos aspectos laríngeos da qualidade vocal. Destacam-se, neste último âmbito, o trabalho do grupo sueco, que resultou na proposta da *SVEA* (Hammarberg et at, 1980; Hammarberg, Gauffin, 1995; Hammarberg, 2000) e da Sociedade Japonesa de Logopedia (Isshiki et al, 1969; Hirano, 1981), com a escala *GRBAS*. Tais aspectos foram considerados em função do impacto da concordância de duas propostas de grande respaldo do ponto de vista de aplicabilidade clínica, ambas fundamentadas no método diferencial semântico (Osgood et al, 1969).

Os aspectos relativos aos ajustes supralaríngeos, por sua vez, não foram destacados em outra escala ou proposta de avaliação formal vocal perceptiva, sendo freqüentemente descritos em roteiros de avaliação como características de ressonância e articulação, como é o caso da proposta *Buffalo Profile System (BVP)* por Wilson (1994), e, em nosso meio, das escalas Brandi para avaliação da voz falada (Brandi, 1996, 2002). Dessa maneira, optou-se por manter as bases de concepção de um modelo com embasamento fonético, considerando-se a representatividade que oferece ao estudo da qualidade vocal por incorporar aspectos relevantes das dimensões realmente representativas da qualidade vocal para os falantes em questão. Além disso, adota a avaliação de emissão encadeada para caracterização do perfil de

qualidade vocal, em detrimento aos procedimentos concentrados em avaliações de vogais sustentadas. As modificações propostas para o presente estudo ocorreram exclusivamente em termos dos ajustes laríngeos e estão centradas nas descrições referentes aos variados mecanismos fisiológicos que levam à ocorrência de escape de ar entre as pregas vocais durante a fonação e aos distintos efeitos auditivos decorrentes. Algumas dificuldades na adaptação também envolveram a tradução ou versão da terminologia para descrição de qualidade vocal, denotando que este representa apenas um primeiro passo na adaptação de tal conteúdo para a nossa realidade.

O primeiro ajuste a sofrer modificações refere-se à terminologia equivalente a sussurro ou cochicho em português (*whisper* no original), relacionado fisiologicamente ao escape de ar, gerando acusticamente sinal de ruído turbulento glótico. Na proposta sueca, para ajuste semelhante, manifestado em grau extremo, encontramos o termo afonia, caracterizado pelo fluxo aéreo turbulento audível. Quando combinado ao ajuste modal, resultaria em algo próximo ao que chamaríamos de voz cochichada ou até mesmo sussurrada em nossa língua, numa tentativa de tradução do termo original *whispery voice*.

Além das discordâncias que os termos sussurro e cochicho podem gerar em termos de julgamentos de qualidade vocal, tradicionalmente, na avaliação clínica vocal, o escape de ar aparece referido como soprosidade (Isshiki et al, 1969; Hammarberg et at, 1980; Hirano, 1981; Wilson, 1994; Hammarberg, Gauffin, 1995; Hammarberg, 2000). Por sua vez, na proposta de *VPAS*, os termos soprosidade e voz soprosa poderiam corresponder respectivamente ao componente *breathiness* e ao ajuste *breathy voice* (Laver, 1980, 2000; Laver et al, 1981).

Diante do exposto, resta a adaptação de terminologia diferenciada para distinguir os ajustes relacionados a *whispery* e *breathy voice*. Nesse sentido, referências de Hammarberg et al (1986) e Beck (1988) colaboraram com a tarefa, no sentido de abordarem justamente tais refinamentos de descrição de atividade glótica. Hammarberg et al (1986) ao exporem pontos da escala *SVEA*, revelam a possibilidade da fator de soprosidade combinado a hiperfunção aproximar-se da designação de *whispery voice* (Laver, 1980). Beck (1988), por sua vez, num estudo voltado à investigação das variações orgânicas de qualidade vocal, salienta que *breathy voice* 

corresponderia a um ajuste praticamente ausente do ponto de vista clínico. Reforça a dificuldade de registro em gravação, justamente por envolver o uso paralingüístico de expressão de intimidade, proximidade entre interlocutores. Tal mobilização guardaria proximidade com a descrição de murmúrio por Ladefoged (1973).

Dessa forma, o termo adotado por Laver (1980) enquanto soprosidade é diverso do uso amplamente difundido em escalas de avaliação clínica. Entretanto a terminologia se propõe a explicitar diferenças no âmbito da percepção e da fisiologia de tais ajustes, de forma que o *whisper* representaria o fluxo turbulento com extremos fisiológicos de tensão laríngea e aumento de pressão subglótica, enquanto o *breathy* representaria um leve escape perceptível com baixos níveis de tensão e de fluxo aéreo. Tais ajustes poderiam ser interpretados como extremos de uma mesma tendência de escape de ar entre as pregas vocais, porém com variações de fluxo e tensão, conforme reforçado por Hammarberg et al (1986).

Diante de tais considerações e, especialmente, em face dos achados clínicos de voz, Beck (1988) sugeriu a remoção do ajuste *breathy voice* do *VPAS* e propôs uma única categoria de ajuste referente à fricção (escape de ar), que no seu grau máximo, a nosso ver, soaria como afonia. Adicionando-se o componente de hiperfunção laríngea, teríamos na manifestação de grau mínimo, o ajuste *breathy voice* e, no de grau máximo, o ajuste *whisper*. Tais variações de tensão de mecanismos laríngeos encontram correspondência com as propostas de avaliação clínica que também apontam subcategorias de hiperfunção ou hipofunção para *SVEA* e de astenia e tensão para *GRBAS*.

Diante de tais considerações, especialmente da relevância clínica e do respaldo dos procedimentos de avaliação mencionados, no roteiro a ser usado neste trabalho a versão da designação do ajuste original *breathy voice* deixou de existir. Para o termo referente ao ajuste *whisper*, foi adotada a designação escape de ar, quando não haveria combinação com ajuste que levasse à sonoridade, como é o caso do ajuste modal. No caso de existir a combinação, *whispery voice*, adotamos o termo voz soprosa. Tal adaptação foi adotada na presente proposta de avaliação, justamente por se considerar a relevância aos achados clínicos e o embasamento da fisiologia anteriormente delienados.

Nesse tocante, o próprio Laver (1980) esclarece que o ajuste escape de ar (*whisper*) e o murmúrio (*breathiness*) apresentariam em comum a baixa tensão adutora, mas com ações complementares da mesma escala, de forma que sua combinação está excluída por definição. Pode-se supor que altos graus de turbulência ou fricção audíveis seriam equivalentes ao ajuste *whisper*, o qual passaremos a referir como escape de ar. A presença de escape de ar, porém sem ruído audível, resultaria no componente *breathiness*, o qual referimos como murmúrio e que, na realidade, não integrará a bateria utilizada para avaliação perceptiva da qualidade vocal.

Tal consideração torna possível a referência aos termos hiper e hipofunção, os quais revelam dois desmembramentos fisiológicos possíveis para o fechamento glótico incompleto. Para completar a discussão e a justificativa da terminologia adotada, pode-se citar finalmente a referência aos ajustes de tensão na proposta do modelo (Laver, 1980). Foram acrescentados itens de hiper e hipofunção, os quais guardam relação com ajustes de tensão em Laver (1980), com as categorias hiperfuncional (tenso) e hipofuncional (relaxado) para *SVEA* e com as referências de tensão e astenia na *GRBAS*. Além disso, designações como ocorrências de diplofonia, instabilidades e quebras foram consideradas em item distinto à designação de ajustes, enquanto ocorrências de curto termo. Por fim, os aspectos de ocorrência intermitente podem ser referidos com um "i", segundo proposta de Laver (2000).

Dessa forma, apesar da distinção do embasamento de cada uma das propostas, foi possível combinar a descrição fonética, com um forte embasamento fisiológico e, à proposta sueca *SVEA* e à escala *GRBAS*, ambas com o apoio e validação de métodos de percepção auditiva.

Ao esboçar um paralelo entre demais ajustes laríngeos presentes no *VPAS* e descritores de qualidade vocal previstos para a *SVEA*, pudemos constatar que ambas apresentam as categorias registros modal, falsete e *vocal fry*. Outro aspecto interessante refere-se ao fato do *VPAS* incluir o componente *harsh* traduzido como aspereza, o qual combinado com o ajuste de escape de ar resultaria na rouquidão. Na *SVEA*, aparecem as categorias *rough* e *gratings*, referentes respectivamente à aperiodicidade em baixas e altas freqüências, as quais poderiam em nosso meio sinalizar respectivamente as ocorrências de rouquidão e aspereza.

Para complementar o paralelo traçado entre as propostas de avaliação vocal, *SVEA* acrescentou apenas os tópicos de *pitch* e *loudness*, os quais no âmbito da Fonética são considerados elementos do domínio da dinâmica vocal e não propriamente da qualidade vocal (Laver, 1978).

Vale ressaltar que não pretendemos efetuar uma simplificação das propostas *SVEA* e escala *GRBAS*, especialmente pela sustentação teórica e relevância de suas contribuições para a avaliação vocal. A comparação apresentada procurou apenas traçar um paralelo entre aspectos relevantes à percepção das alterações vocais. Guardamos entretanto as devidas proporções de similaridade, bem como as suas particularidades, especialmente marcadas pela característica de enfoque exclusivo de mecanismos laríngeos (Isshiki et al, 1969; Hammarberg et at, 1980; Hirano, 1981; Hammarberg, Gauffin, 1995; Hammarberg, 2000). Tais descritores devem estar presentes nas descrições das vozes disfônicas, independente do referencial adotado.

Nesse sentido, em meio às controvérsias que assolam nosso meio no que tange à terminologia de avaliação das alterações vocais, registra-se proposta de avaliação perceptiva da fonte glótica para a realidade brasileira (RASAT), inspirada na *GRBAS*, constando da descrição de roquidão, aspereza, soprosidade, astenia e tensão para as respectivas iniciais da sigla (Pinho, Pontes, 2002).

O roteiro a ser usado como instrumento de julgamento auditivo da qualidade vocal no presente estudo encontra-se exposto no quadro 1.

Várias propostas de avaliação subseqüentes à descrição do modelo procuraram refinar sua aplicabilidade na situação clínica, bem como a reforçar a necessidade de treinamento para uso do instrumento (Laver et al, 1981; Beck, 1988; Laver, 1994, 2000).

Nolan (1983) realizou uma pesquisa voltada à identificação dos correlatos acústicos de qualidade vocal (espectros de longo termo) das gravações originais de 30 qualidades vocais descritas enquanto ajustes articulatórios por Laver (1980). Nenhuma das medidas diferenciou os ajustes, mas ficou nítida a possibilidade de discriminar os subtipos laríngeos (fonatórios), em detrimento aos supralaríngeos.

Quadro 1. Proposta de avaliação da qualidade vocal com motivação fonética

CATEGORIA	AJUSTES	GRAUS DE ESCALA						
		neutro	1	2	3	4	5	6
SUPRALARÍNGEOS	LARÍNGEOS							
LONGITUDINAIS	Laringe alta							
	Laringe baixa							
	LABIAIS							
	Labiodentalização							
	Protrusão							
SUPRALARÍNGEOS	LABIAIS							
TRANSVERSAIS	Arredondados							
	Estirados							
	MANDIBULARES							
	Fechada							
	Aberta							
	LINGUAIS – PONTA/ LÂMINA							
	Avançada							
	Recuada							
	LINGUAIS - CORPO							
	Avançado							
	Retraído							
	Elevado							
	Abaixado							
	LINGUAIS – BASE							
	Constrição faríngea							
	Expansão faríngea							
VELOFARÍNGEOS	ACOPLAGEM VELAR							
	Nasal							
	Denasal							
FONATÓRIOS	Modal							
	Falsete							
	Escape de ar							
	Voz soprosa							
	Vocal fry/crepitânncia					_		
	Voz crepitante							
	Voz áspera							
	Hiperfunção							
	Hipofunção							

Ocorrências em curto termo: ( )diplofonia ( )quebras ( ) instabilidades

Para ajustes de ocorrência intermitente assinalar (i).

Na sequência de estudos da mesma natureza, Pittam (1987) foi capaz de diferenciar três tipos de vozes correspondentes aos ajustes laríngeos de voz soprosa, voz crepitante e voz cochichada, por meio do espectro de longo termo. Além disso, ressaltou a hipótese da nasalidade ser decorrente de uma série de mobilizações do aparelho fonador, em concordância com Laver (1979, 1980).

Esling (2000), ao comentar de variações da qualidade vocal entre as línguas, retoma a discussão acima reafirmando que os ajustes laríngeos são identificáveis por meio de medidas de longo termo. Salienta que um importante parâmetro do estudo do espectro de longo termo entre diversas línguas é a mudança de qualidade vocal devido à variação dos traços dominantes nos comportamentos articulatórios. Em algumas comunidades, foram encontrados padrões críticos de mudanças associados à atividade de laringe e de faringe.

Há que se registrar também, a investigação da validade do uso do *VPA* realizada por Shewell (1998), comprovando a importância do uso e reforçando a familiaridade dos clínicos da Grã-Bretanha no uso do *VPAS*.

Em nosso meio, poucos trabalhos foram sistematizados a partir do modelo fonético de descrição da qualidade vocal. Cassol et al (1999) efetuaram estudo do *VPAS* (Laver et al, 1981) discutindo sua correspondência com o sistema de avaliação de disfonias proposto por Behlau, Pontes (1995). Encontraram maior facilidade e concordância por parte dos avaliadores na classificação dos ajustes fonatórios. Os julgamentos referentes aos ajustes supralaríngeos foram alvo de discordância e dificuldades entre os juízes, especialmente aqueles relativos à posição do corpo da língua e mandíbula. Destacaram a validade do modelo para análise perceptivo-auditiva da qualidade vocal.

Na concepção adotada neste trabalho, procuramos delinear as bases de grupos com experiência em desenvolvimento de instrumentos de avaliação vocal, inclusive em âmbito clínico. Vale ressaltar que as contribuições de outras propostas de avaliação implementaram de forma significativa a adaptação das descrições de ajustes laríngeos na proposta de modelo fonético de descrição da qualidade vocal.

Vale, finalmente, ressaltar que os ajustes laríngeos são exaustivamente referidos na literatura e enfocados na prática clínica, em detrimento aos supralaríngeos. Dessa forma, esta proposta inicial de implementação do roteiro de avaliação de qualidade vocal esteve restrita ao plano dos ajustes laríngeos, porém, no futuro, poderão envolver o conjunto de ajustes supralaríngeos.

### 4.3. Eletroglotografia (EGG)

Destinamos este tópico à exploração dos fundamentos da técnica de eletroglotografia, partindo de um breve histórico de seu desenvolvimento, incluindo seus princípios básicos, até gradativamente atingir as modalidades de análise que podem colaborar para a abordagem da dimensão fisiológica da qualidade vocal de indivíduos disfônicos.

## 4.3.1. Histórico e princípios básicos

A técnica atualmente referida como eletrogolotografia (EGG) foi concebida a partir da adaptação de método de registro de pressão (arterial) sangüínea: a sigmografia de alta freqüência, possibilitando a detecção de mudanças de impedância ao longo do pescoço, causadas pela vibração de pregas vocais, graças à acoplagem de eletrodos de contato no pescoço (Fabre, 1957). Inicialmente conhecida como glotografia de alta freqüência, seu mecanismo consiste na liberação de uma corrente alternada de alta freqüência, da ordem de 200.000 oscilações por segundo, porém de intensidade muito fraca, para percorrer a laringe, de lado a lado, através da pele do pescoço, graças ao posicionamento de dois eletrodos contra as alas da cartilagem tireóide. Parte das linhas de corrente elétrica usam as vias da glote, de maneira que a passagem é facilitada na medida em que a mesma se fecha.

A amplitude da corrente elétrica passa a ser modulada pelo contato das pregas vocais, apresentando variações proporcionais à amplitude da corrente modulada e ao grau de fechamento glótico. A resistência à corrente diminui em razão

da superfície sobre a qual as bordas das pregas vocais se projetam. A evolução da modulação da corrente detectada traduzirá uma imagem da cinética glótica (Fabre, 1957).

Fabre (1957) também apresenta os primeiros resultados da morfologia do glotograma, suas relações com a onda acústica, então chamada fonograma, e as mudanças bruscas da amplitude do glotograma ao longo de uma emissão em escala. Discutiu as contribuições dos achados glotográficos para uma das teorias fonatórias vigentes na época, a neurocronáxica. Por fim, registrou os traçados das emissões vocálicas de um cantor com f0 por volta de 258 Hz (Do3). Nas emissões em fraca intensidade, o traçado glotográfico registrado foi senoidal. Com o seu aumento, detectou o que chamou de glotograma típico: traçado ascendente da curva (mais lento) sinalizando a abertura e o traçado descendente (mais rápido) demarcando o fechamento de pregas vocais. Na comparação com os chamados fonogramas, revelou que a amplitude máxima da onda acústica foi detectada em tempo próximo (ligeiramente posterior) à abertura glótica máxima.

Segundo o relato anterior, destacava-se inicialmente que a eletroglotografia derivaria, de forma mais refinada, a informação de abertura das pregas vocais. Tal conceito foi reforçado por algumas publicações posteriores até a década de 60 (Fabre, 1959; Vallencien, Faulhaber, 1966, 1967), as quais salientavam que os picos de amplitude do sinal refletiriam a área glótica, ou seja, seriam correspondentes à fase de abertura de pregas vocais. Coube a Fourcin, Abberton (1971) a argumentação de que o sinal de saída da eletroglotografia estaria relacionado ao contato de pregas vocais.

Nesse aspecto, ao longo de décadas de pesquisas com o uso do instrumento, foi reforçado o consenso de que a amplitude do sinal eletroglotográfico representa de forma mais dinâmica a área de contato de pregas vocais e não a área glótica. A fase de fechamento mostrou correspondência ao máximo de excitação do trato vocal, resultando na detecção de picos de amplitude no sinal registrado (Fourcin, Abberton, 1977; Childers, Krishnamurthy, 1985; Baken, 1987, 1992; Childers et al, 1990; Orlikoff, 1991; Bless, Baken, 1992; Vieira, 1997; Baken, Orlikoff, 2000; Fourcin, 2000; Mancini, 2001).

Vale recordar que a acoplagem de dois eletrodos nas alas da cartilagem tireóide gera um elemento de circuito que pode então ser monitorado em termos de sua impedância ou de sua admitância (Fourcin, 2000). A voltagem zero na saída indica que o contato entre pregas vocais não está mudando (Vieira, 1997).

A fim de esclarecer tais características de ordem técnica da eletroglotografia para o âmbito da Fonoaudiologia, Baken (1987, 1992) e Baken, Orlikoff (2000) exploram as suas bases e expõem as características dos tecidos humanos enquanto condutores moderadamente bons de eletricidade. Numa aproximação simples, os autores destacaram que as principais estruturas corporais teriam comportamento semelhante a resistores na proposta da lei de Ohm. Dessa forma, a corrente numa estrutura é proporcional à voltagem aplicada e inversamente proporcional à resistência do circuito, no qual a corrente equivaleria à quantidade de eletricidade (expressa em ampère) e a voltagem à pressão elétrica (expressa em volts). Em outras palavras, a corrente passando por um condutor é proporcional à voltagem que a impulsiona. Se a voltagem (pressão elétrica) permanecer constante, a corrente (fluxo de eletricidade) declina à medida que a resistência aumenta. A resistência à corrente alternada (AC) é chamada de impedância, a qual difere da corrente direta (DC).

Entendendo a impedância como um fenômeno geral de todos os sistemas vibrantes, como uma oposição à vibração ou à oscilação, a impedância acústica passa a ser considerada como a oposição ao movimento vibratório. Os fatores que interferem na impedância do sistema são a massa do aparelho vibrante, a elasticidade ou a rigidez e a fricção. Eletronicamente, a impedância se manifesta na relação entre a voltagem e a corrente ao longo de um circuito. Sua unidade de medida é o ohm. A admitância, por outro lado, é definida como o total de fluxo de energia e, portanto, o contrário da impedância (Kohen, 1985). Diante de tais considerações, Fourcin (2000) descreveu a eletroglotografia como método mais precisamente baseado em admitância, entretanto a maior parte das referências mantém a designação impedância, na qual a baixa impedância representa alta admitância e vice-versa, tratando-se portanto de referências complementares.

Orlikoff (1998) reafirmou que a impedância nesta região é uma função do tamanho, formato, orientação e química dos tecidos e fluídos que repousam

na via de corrente. Epitélio, músculo, mucosa e saliva oferecem pouca impedância à corrente, enquanto tecido adiposo representa um impedimento elétrico importante. Nesse universo, o ar apresenta impedância quase infinita, comportando-se como um isolante. A corrente flutua em ambas direções ao longo do pescoço, sendo a voltagem proporcional à magnitude da impedância ao longo do pescoço. Qualquer contato entre as pregas vocais (incluindo aquele de superfícies mucosas) serve como meio pelo qual a corrente pode tomar uma via direta entre os eletrodos.

Baken (1992) destacou que a estrutura laríngea não é tão proeminente em todos os indivíduos, impondo dificuldades de acesso para acoplagem do instrumental necessário. Outro problema descrito refere-se ao fato da corrente não representar uma linha direta entre os eletrodos. Nesse sentido, o pescoço atua como um condutor de volume, no qual as correntes se espalham em todas as direções, não havendo homogeneidade. Por conseqüência, a corrente passa por estas vias, as quais não são alvo de interesse da eletroglotografia. Qualquer medida de impedância acústica registrada pelos eletrodos posicionados nas alas da cartilagem tireóide inclui todas as estruturas do pescoço situadas nas vizinhanças dos eletrodos (Baken, 1987; Houben et al, 1992), não sendo possível atribuir a modulação do sinal exclusivamente ao movimento de pregas vocais (Rothenberg, Mashie, 1988).

Smith (1981) colaborou com esta temática ao reforçar a hipótese de que as mudanças de impedância observadas não seriam o resultado único do contato de pregas vocais. Nessa visão, o pescoço e o eletroglotógrafo funcionariam como um único microfone complexo.

Baken (1992), retomando algumas observações originais, voltou a destacar que a corrente deve ser reduzida para evitar reações como contração muscular e produção de impulsos nervosos em decorrência da estimulação das fibras nervosas. As correntes de alta freqüência são pouco perceptíveis ao indivíduo testado, desde que aplicadas em amplitudes reduzidas. O uso biológico de corrente elétrica deve ser criterioso. Condutores com freqüências situadas abaixo de 50 Hz poderiam provocar sensações desagradáveis nos indivíduos testados (Vieira, 1997). Por esta razão, todos os cuidados relativos à segurança devem ser mantidos, envolvendo níveis de corrente seguros para uso em seres humanos (Baken, 1987; Vieira, 1997). Segundo

Vieira (1997), transformadores também são usados para propiciar isolamento elétrico aos usuários.

Para melhor compreensão dos aspectos técnicos relacionados ao referido uso de corrente elétrica, vale recordar que o gerador de sinal fornece uma freqüência de sondagem de corrente alternada (AC), a qual depende do modelo comercial, mas facilmente encontra-se na faixa de 0,3 a 5 MHz (Childers, Krishnamurthy, 1985; Baken, 1987). A amplitude é da ordem de 0,5 V (Childers, Krishnamurthy, 1985; Baken, 1992; Orlikoff, 1998).

De forma resumida, pode-se dizer que a corrente de radiofreqüência (RF) detectada é modulada em termos da amplitude, em função da vibração de pregas vocais, mais especificamente pela variação de seu contato. O percentual de modulação de amplitude no sinal reflete a mudança na impedância dos tecidos na via da corrente gerada (Childers, Krishnamurthy, 1985). Um detector de circuito é necessário para registrar a modulação do sinal RF. A forma da onda detectada é apresentada num oscilógrafo, podendo ter deflexão positiva ou negativa para fechamento e abertura.

A linha de base da onda pode variar, especialmente em função de movimentos da laringe, os quais introduzem impedâncias secundárias. Tal instabilidade poderá ser compensada por dispositivos especiais, como filtragem passa-altas. Os aparelhos incluem ganhos automáticos de circuito (AGC), os quais colaboram para manter o sinal EGG num nível constante, compensando as variações no sinal relativas às variações de impedância no pescoço (Colton, Conture, 1990). Levando este fato em conta, torna-se necessário reconhecer as dificuldades ou talvez a impossibilidade de se relacionar o traçado visível diretamente às variações fisiológicas, uma vez que a amplitude original foi corrigida (Childers, Krishnamurthy, 1985; Vieira, 1997). Nesse universo, não é possível a obtenção de medidas absolutas da impedância translaríngea (Vieira, 1997). O sinal detectado, amplificado e sujeito a um controle de ganho automático (AGC), também é limitado a uma certa filtragem banda passa (BP), geralmente situada entre 0 a 10 kHz (Baken, 1987). Vieira et al (1996a) utilizaram filtragem banda-passa de 60 a 5000 Hz para estudo de alterações vocais voltado à detecção de f0 e jitter.

As modalidades de filtragem também variam entre os vários instrumentos disponíveis. Um dispositivo de corte de freqüências passa-altas reduz a influência de pequenos movimentos causados pela mudança da posição vertical da laringe no pescoço e pelos próprios eletrodos (Vieira et al, 1996a; Baken, 1987; Vieira, 1997).

Diante de tais considerações, as mudanças na impedância translaríngea refletem no mínimo dois tipos de evento: rápidos e lentos. Aqueles de natureza lenta aparecem como flutuações na linha de base e são causados por ajustes ventilatórios (incluindo registro de energia subglótica) e articulatórios, além do fluxo sanguíneo nas artérias e veias do pescoço, movimentos de estruturas cervicais (epiglote, base de língua, pregas vestibulares, musculatura de hipofaringe e esôfago) ou do pescoço, da cabeça e dos eletrodos no pescoço (Baken, 1987; Colton; Conture, 1990). Por outro lado, os movimentos rápidos são causados pelo ciclo vibratório de pregas vocais (Fourcin, Abberton, 1971; Smith, 1981; Baken, 1987; Bless, Baken, 1992; Vieira et al 1996a; Vieira, 1997).

Conforme exposto anteriormente, todas as estruturas que estão situadas entre os eletrodos (mesmo as mais distantes) podem interferir no fluxo da corrente elétrica (Baken, 1987; Colton, Conture, 1990). Embora todos os equipamentos contenham recursos para rejeitar sinais não fonatórios e compensar flutuações em linha de base, o instrumento continua sensível a artefatos de movimento (Baken, 1987). O próprio autor do EGG salientou nas particularidades da glotografia elétrica de alta freqüência, os artefatos impostos pela região do pescoço, dos movimentos respiratórios e deslocamentos involuntários (Fabre, 1959). Colton, Conture (1990) também relataram efeitos associados da circunferência do pescoço ao grau de impedância dos tecidos. Em nota técnica recente, Rothenberg (2002) aborda mecanismos de correção das distorções de fase em baixa fregüência.

Colton, Conture (1990) listaram os problemas instrumentais associados à técnica. Em primeiro lugar, abordaram a colocação dos eletrodos, no que se refere a altura, grau de contato com a pele, presença de gordura e movimentos de laringe durante a gravação. Sugeriram a limpeza prévia da pele na região cervical com álcool, ajudando a remover excesso de oleosidade. Além disso, destacaram as razões inerentes ao sujeito avaliado, tais como a conformação física, a presença de gordura

na região cervical e o ângulo da cartilagem, de artefatos induzidos pela fase do sinal. Finalmente, o uso de gel para facilitar a acoplagem dos eletrodos e melhorar as condições gerais de exame. Para Baken (1987) as especificações de aparelhos podem requerer ou limitar o uso de tais substâncias.

Para evitar problemas de acoplagem, Colton, Conture (1990) também propuseram a pressão dos dedos sobre os eletrodos em lugar da fita para manter os eletrodos posicionados durante a avaliação. Anteriormente, Baken (1987) sugeriu que o indivíduo avaliado estivesse posicionado com suporte de cabeça para minimizar movimentos durante a avaliação. Outros cuidados a serem adotados incluem a remoção de correntes e gargantilhas, colares de metal por introduzirem ruído no sinal (Vieira, 1997).

Watson (1995) sintetizou quatro fatores básicos no controle de qualidade dos sinais obtidos: função vocal precária, interferências nas vias de corrente no pescoço em casos de contato precário de pregas vocais, colocação inadequada de eletrodos e técnica deficiente de processamento do sinal (erros de fase, inadequada resposta de freqüência inferior (LF) e de proporção sinal-ruído).

O melhor posicionamento dos eletrodos varia de acordo com aparelhos de diferentes fabricantes (Lecluse et al, 1975; Vieira, 1997). Lecluse et al (1975) revelaram os resultados de estudos com laringes excisadas, variando a posição dos eletrodos e tipo de equipamento. Registraram resultados diferentes tanto para a posição de eletrodos no pescoço, assim como para equipamento, alguns dos quais foram mais susceptíveis à mudança na posição dos eletrodos.

Retornando ao ponto do tratamento da informação obtida, reforçarmos que o sinal analisado é o resultado das mudanças de impedâncias captadas pelo instrumento e previamente filtradas. Dessa forma, o procedimento de filtragem para acentuar os componentes verdadeiramente vocais resultam na onda Lx (Colton, Conture, 1990).

O termo Lx foi introduzido por Fourcin, Abberton (1971, 1972) para denominar a onda de saída, e portanto do sinal glótico resultante, o qual revela a superfície de contato de pregas vocais. Passou a constituir a base do estudo dos

contornos de f0 na fala (Fx), os quais correspondem à entoação e também seriam potencialmente úteis para estudar as condições laríngeas, justamente por avaliarem o contato de pregas vocais. Destacaram que no caso de alterações laríngeas, a onda Lx não necessariamente seria alterada de maneira uniforme ao longo do tempo. Baken (1987) destacou que o estímulo não filtrado foi designado como Gx, o qual seria minimamente utilizado em situações clínicas.

A evolução da terminologia envolveu várias nomenclaturas subsequentes àquela utilizada por Fabre (1957) de glotografia de alta frequência, tais como: eletroglotografia, laringografia e eletrolaringografia, em muitas ocasiões tratadas como sinônimos. Entretanto, alguns autores destacaram que os dois últimos termos seriam mais apropriados por referirem-se à laringe como unidade e não exclusivamente à glote (Fourcin, Abberton, 1977).

Além disso, a menção ao termo laringografia na literatura também envolveu uma adaptação da técnica original de glotografia, com implementação de variações de ordem elétrica, ajustes instrumentais em eletrodos e rede de conexão (Fourcin, Abberton, 1972). Por outro lado, o termo laringografia seria usado preferencialmente pelos radiologistas, há muito tempo, para a visualização laríngea, daí eletroglotografia causar menos confusão, apesar da inadequação da terminologia, no sentido de que o registro da técnica não estaria restrito à atividade glótica (Baken, 1987).

Nesse âmbito, Ondrácková (1972) apontou a necessidade de distinção entre glotografia elétrica e óptica. A primeira seria a técnica descrita neste tópico, enquanto o método óptico daria origem à técnica de filtragem inversa. Por sua vez, Orlikoff (1998) abordou o termo glotografia como uma variedade de técnicas voltadas à descrição do comportamento de pregas vocais durante a fonação, incluindo fotoglotografia (PGG), glotografia de filtragem de fluxo (filtragem inversa), ultrassonografia, fotografia ultrarrápida, além da eletroglotografia.

Na evolução dos estudos da onda Lx, vale destacar que não foi possível estabelecer convenções para o seu traçado. Na descrição de Fabre (1957) o ramo ascendente da onda (formato curvo) marcava abdução e o descendente (formato reto) a adução.

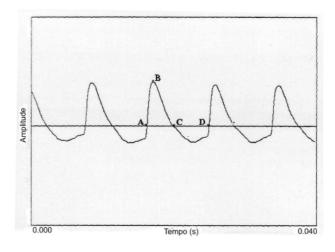
Fourcin; Abberton (1977) utilizaram o EGG e a onda Lx de maneira diversa a de Fabre (1957), ou seja, com contato de polaridade positivo, no qual a adução estaria representada no trajeto ascendente e abdução no descendente. Tal tendência foi preservada em muitos estudos subseqüentes.

Baken (1992) publicou as recomendações da *IALP* (*International Association of Logopedics and Phoniatrics*) para eletroglotografia, com sugestão de padronização da polaridade positiva para o contato de pregas vocais. Entretanto a adoção de curvas com polaridade inversa voltou a ser destacado por usuários que realizam a análise direta sincronizada a outras técnicas como estroboscopia, fotografia ultrarrápida e filtragem inversa anteriormente referidas (Titze, 1984; Karnell, 1989; Holmberg et al, 1995; Fourcin, Abberton, 1999).

Até o momento atual, a grande contribuição da eletroglotografia reside na interpretação das informações contidas no traçado da onda Lx, enquanto análise qualitativa das ações laríngeas, e de algumas tentativas de obtenção de estimativas, as quais passam a ser abordadas no tópico de procedimentos de análise.

#### 4.3.2. Procedimentos de análise

A primeira abordagem refere-se à inspeção da onda Lx, apresentada na figura 1, em que são indicadas três fases do ciclo vibratório: de fechamento, de abertura e aberta.



### Legenda:

A-B – fase de fechamento / B-C – fase de abertura / C-D - fase aberta / A-D - ciclo vibratório

Figura 1. Representação da onda Lx e as informações das fases do ciclo vibratório (fases de fechamento, abertura e aberta).

Fourcin, Abberton (1971, 1972) relataram três partes da onda em indivíduos normais. Na primeira, a subida abrupta representaria o rápido fechamento de pregas vocais, característico da vibração normal e coincidente com a máxima excitação do trato vocal (figura 1 A-B). Em seguida, ocorreria o declínio gradual (figura 1 B-C), associado à separação das pregas vocais, quando a pressão subglótica aumenta. Finalmente, a base alargada da forma da onda corresponderia ao intervalo no qual a glote estaria aberta e as pregas vocais fora de contato (figura 1 C-D). Ao se adotar a polaridade inversa, este traçado se inverte.

Os modelos de funcionamento de pregas vocais foram fundamentais para desvendar a fisiologia das pregas vocais, em parte impressa no traçado da onda Lx. Sua descrição tornou-se presente em vários estudos, especialmente em termos de referências ao modelo de uma massa, fechamento glótico no sentido ântero-posterior, e duas massas, incluindo a movimentação em diferença de fase vertical das bordas superior e inferior de pregas vocais (Childers, Krishnamurthy, 1985; Childers et al, 1986; Baken, 1987, 1992; Colton, Conture, 1990; Orlikoff, 1991; Titze, 1984, 1989; Vieira, 1997).

Segundo Titze (1984, 1989), Childers et al (1990) e Orlikoff (1998) o modelo de duas massas não contemplaria a necessidade de descrição do mecanismo laríngeo, pelo fato do contato glótico também depender, de forma crítica, de fatores anatômicos e biomecânicos de pregas vocais. Childers, Krishnamurthy (1985) apontam que a prega vocal, além do corpo e da cobertura, teria uma camada adicional de muco, secretada pelas glândulas ventriculares, que poderiam causar interferências nos registros.

Tais modelos permitem o refinamento na leitura dos traçados, de forma que há vários pontos a serem demarcados no traçado da onda Lx: o primeiro contato subglótico (figura 1–A), a fase de fechamento (figura 1 A-B, aumentando contato vertical), a fase fechada (figura 1 B-C, separação subglótica das pregas vocais), fase aberta, (figura 1 C-D, separação dos lábios superiores de ambas as

pregas vocais) e aproximação subglótica (figura 1 A ou D) (Baken, 1987; Karnell, 1989).

A abordagem da onda permite também permite identificar aspectos relativos às estimativas das fases de fechamento, fechada, de abertura e aberta:

- fase de fechamento (porção de aumento de contato): definida pelo intervalo de tempo que se estende desde a área glótica máxima até a área zero, ou seja, o instante do fechamento glótico. Segundo Orlikoff (1998), deve preferencialmente ser referida como fase de aumento de contato. Se o indivíduo conseguir estabelecer bom contato de pregas vocais, o sinal será bom, independente das alterações existentes (Colton, Conture, 1990).
- fase fechada: começa no ponto de fechamento glótico completo, que é o instante no qual a área glótica é zero até o momento em que se inicia a separação dos lábios inferiores da prega vocal. Nesse intervalo, o traçado reflete manutenção da área de contato das pregas vocais (Childers, Krishnamurthy, 1985). Para Orlikoff (1998), deve ser chamada de fase de contato, refletindo de maneira mais fidedigna os eventos fisiológicos.
- fase de abertura: caracterizada pelo intervalo demarcado entre o instante de abertura da glote até o máximo valor da área glótica (Childers, Krishnamurthy 1985). A abertura é gradual com grande diferença vertical de fase. Considerada por Colton, Conture (1990) como uma fase que engloba outras: as bordas inferiores das pregas vocais separam-se, a seguir as bordas superiores, tornando-se as pregas vocais completamente abertas. Nos indivíduos normais, podem estar claras, porém nos indivíduos com patologia podem não ser detectadas com nitidez.
- fase aberta: nenhum contato é registrado. Caracterizada como a fase mais sujeita às interferências das adjacências (Vieira, 1997).

Nesse âmbito, uma das maiores dificuldades reside em demarcar o começo e o término do contato de prega vocais na onda Lx, pois a glote fecha um pouco antes do ponto máximo da onda Lx e permanece fechada até um pouco depois do pico, refletindo os pequenos movimentos verticais das pregas vocais que resultam em variações da área de contato (Kelman, 1981; Baken, 1987; Orlikoff, 1995). O critério

adotado para estimativas da duração do contato varia em torno de 25% a 35% do pico de amplitude do sinal (Rothenberg, Mashie, 1988; Orlikoff, 1995).

A interpretação do traçado da onda a partir da comparação com os modelos propostos para atividade de pregas vocais incluem julgamentos de alargamento do pulso de contato, indicativo de aumento de adução de pregas vocais; achatamento de pico de contato para a direita, indicativo de convergência glótica; abaulamento do rodapé, indicativo de protrusão medial das superfícies de pregas vocais e onda de forma mais triangular, indicativo de aumento diferença de fase vertical (Orlikoff, 1998).

Nesse sentido, o traçado da onda seria o dado clínico relevante e as medidas de comportamento glótico dependeriam de conhecimento de se e quando a glote está aberta ou fechada. A técnica de EGG não provê nenhuma informação quando o contato se perde. Isoladamente, não seria possível afirmar se ou quando se perde o contato entre as bordas das pregas vocais, se ou quando a glote fechou, se ou quando o fechamento glótico é completo. As justificativas para seu uso residiriam na necessidade de avaliar o quanto as pregas vocais estão aduzidas firme ou frouxamente, o tipo de contato de pregas vocais usado e o quão regular é o padrão de contato de pregas vocais (Orlikoff, 1998).

Em publicação anterior, Orlikoff (1991) destacou os parâmetros de uniformidade da amplitude dos picos sucessivos da onda Lx, da periodicidade do grau de contato de pregas vocais, da forma de inclinação da onda (do ramo da onda correspondente ao aumento de contato de pregas vocais) e da presença de duração de contato relativamente longa.

Apesar dos conceitos voltados à interpretação dos dados, na última década ocorreu uma proliferação voltada ao desenvolvimento de medidas, as quais, apesar da tentativa de interpretação objetiva, resultaram em complicações e interpretações errôneas. A EGG provê informações específicas sobre área de contato de pregas vocais, mas não permite traçar correlações estreitas com fenômenos de outra ordem, ou seja, acústicos, aerodinâmicos ou glóticos (Orlikoff, 1998). Portanto, são consideradas válidas, apenas as medidas que revelam fatores significativos do comportamento de pregas vocais.

Vieira (1997) destacou o cuidado ao se usar sinais da EGG para análise automática, quanto ao processamento de ondas fracas, ruidosas ou extremamente irregulares, as quais devem ser desqualificadas para medidas objetivas. A determinação precisa e automática dos ciclos é um importante pré-requisito para extração de características relacionadas ao formato da onda, de forma que propôs método para tal finalidade, o qual foi utilizado no processamento das amostras desta esfera de investigação no presente estudo, incluindo etapa prévia de avaliação do traçado da onda Lx, como forma de avaliar a possibilidade de extração de medidas EGG (Vieira et al, 1996a; Vieira, 1997).

A partir do processamento da onda Lx, surgiram as propostas de exploração de várias medidas:

- •Freqüência fundamental: f0 (freqüência fundamental): obtida pelo intervalo dos ciclos.
- •Medidas de perturbação ciclo-a-ciclo: *jitter* (diferença média dos períodos absolutos entre ciclos vocais sucessivos) e *shimmer* (diferença média das amplitudes pico-a-pico dos ciclos sucessivos).
- •Quociente ou coeficiente de contato (razão entre a duração da fase fechada/ contato e a duração total do ciclo): referido, inicialmente, como quociente de adução.
- •Quociente de abertura: Razão entre a duração da fase aberta e o tempo total do ciclo.
- •Quociente de velocidade: análise do ramo da porção de aumento de contato da onda (razão entre velocidade mínima e máxima, revelando o tempo de fechamento)
- •Índice de velocidade: expressa a razão da diferença de duração entre as fases de contato-fechamento e de contato-abertura e a duração total da fase de contato.
- Ciclo ativo: período de atividade registrado no ciclo.
- •Fator forma: diretamente relacionado à soma média das diferenças quadradas em amplitude por ciclos, dependente do nível de ruído no sinal e das perturbações (jitter e shimmer)
- •Fator crista: razão entre valor (amplitude) do pico e valor do quadrado médio do período

Considerações técnicas envolvem as extrações de medidas, especialmente no que se refere a eliminação das instabilidades de linha de base e a apreciação do nível de ruído do sinal.

As primeiras descrições do uso da técnica para avaliação de alterações vocais revelaram a complexidade do sinal e o fato de não poder ser incorporada naquele momento ao ambiente clínico (Fourcin, Abberton, 1971). Em

Childers, Krishnamurthy (1985) são apontadas interpretações clínicas, incluindo correspondências com traçados acústicos, de forma que a uniformidade de picos da onda Lx corresponderiam a um sinal acústico uniforme. Destacaram que a alteração laríngea só seria detectada pela técnica quando ocasionasse alterações no ciclo vibratório. Orlikoff (1998) combateu decisivamente essa hipótese salientando que os sinais são de natureza diversa.

O fato de que é necessário cuidado na comparação de dados eletroglotográficos com acústicos é salientado nos trabalhos de um grupo de autores que estudaram a correspondência entre medidas de perturbação acústicas e eletroglotográficas na situação de alterações vocais (Vieira et al, 1996a, 1996b; Vieira, 1997; Lindsey, Vieira, 1997), os quais passam a ser comentados, justamente por representarem parte do respaldo de método relativo à eletroglotografia no presente estudo.

Lindsey, Vieira (1997) descreveram método de análise das modulações em freqüência (FM) e em amplitude (AM) como uma nova forma de exploração da perturbação do ciclo vibratório. Destacaram a aplicabilidade do método para revelar a relação de *jitter* com a modulação de freqüência e *shimmer* com a de amplitude e a dificuldade de representar *jitter* e *shimmer* em um modelo de produção e de percepção.

Vieira et al (1996a) compararam medidas de f0 e *jitter* acústico e eletroglotográfico (EGG) em vozes disfônicas durante as emissões sustentadas e encadeadas de [a] [i] [u] e relataram dificuldades quanto ao uso das duas últimas vogais por ocorrerem imprecisões na determinação dos ciclos e, conseqüentemente, falhas na extração de medidas. Procedimentos de síntese podem ser úteis, porém as simulações não necessariamente refletem as anormalidades presentes nas disfonias. A vogal [a] revelou-se potencialmente promissora para estudo do comportamento laríngeo na presença de desordens.

Vieira et al (1996b), utilizando as referidas medidas no estudo anterior, detectaram pequenas anormalidades nos movimentos da cartilagem aritenóide em 45 indivíduos disfônicos. Tal alteração laríngea foi relacionada às diferenças na tensão entre prega vocal direita e esquerda, levando à irregularidade vibratória.

Vieira et al (1997) realizaram estudo comparativo das médias de *jitter* extraídas por meios acústicos e eletroglotográficos, destacando a escassez de estudos desta natureza. Relataram a aplicabilidade de algoritmo para estimativas de f0 baseado em cruzamento de zero para implementação das medidas obtidas para indivíduos disfônicos.

Vieira (1997) desenvolveu uma série de algoritmos para extração de medidas de f0 e de *jitter*, conjugando as análises acústica e eletroglotográfica, com o objetivo de desenvolver um método de suporte ao diagnóstico de patologias laríngeas. Em termos da aplicabilidade, ressaltou a importância das medidas de *jitter* EGG na detecção de casos de assimetria de comissura posterior de pregas vocais, decorrentes de paralisia seletiva de músculos intrínsecos da laringe.

Na continuidade deste nível de investigação, Vieira et al (2002) discutiram a influência das patologias laríngeas em medidas de *jitter* acústico e *jitter* EGG. Os maiores desvios de valores de *jitter* acústico ocorreram para pequenas lesões nodulares sem assimetrias de comissura posterior, enquanto o *jitter* EGG esteve relacionado a situações de hiperfunção vocal, pequenas assimetrias de comissura posterior e fendas glóticas, denotando sua aplicabilidade clínica.

O mesmo grupo destacou que a eletroglotografia representa basicamente o comportamento dinâmico da área de contato de pregas vocais, de forma que o sinal não depende somente da integridade física das pregas vocais, mas dos aspectos aerodinâmicos da fonação. Além da importância de algumas medidas relatadas, a anormalidade do sinal, ou seja, do traçado da onda Lx chama a atenção para a necessidade de investigação mais detalhada da atividade laríngea.

# 5. MÉTODOS DE COLETA DE DADOS

Este capítulo foi estruturado em tópicos referentes ao delineamento de método para a coleta de amostras voltadas à abordagem da qualidade vocal na situação de disfonia. No primeiro item são descritos os critérios de composição do grupo estudado e seleção do indivíduo referência. Na seqüência, segue exposta a concepção geral das etapas de coleta das amostras de fala por meios acústicos e eletroglotográficos, incluindo o preparo do instrumental e do indivíduo avaliado. Os procedimentos complementares de análise fisiológica por coleta de dados de exame otorrinolaringológico representam o item final deste capítulo.

Conforme exposto anteriormente, a avaliação dos parâmetros vocais acústicos, perceptivo-auditivos e fisiológicos envolve respectivamente os procedimentos de gravação e processamento digital do sinal de fala, a análise das amostras por grupos de ouvintes e, finalmente, o registro de atividade de alguns segmentos do aparelho fonador durante a produção vocal.

A opção pela associação de dados de ordem acústica e eletroglotográfica revelou a necessidade de cuidadoso planejamento de coleta de amostras de fala, de forma a eliminar artefatos que poderiam sinalizar possíveis irregularidades de sinal, impostas por variáveis não propriamente relativas ao quadro de alteração vocal. Dessa forma, a concepção dos procedimentos envolveu vários testes preliminares, os quais colaboraram para o aprimoramento do método e para a garantia da qualidade das amostras coletadas, seguindo tendências apontadas na literatura, ao incorporarmos procedimentos de monitoramento dos registros, com garantia do nível de gravação, da manutenção de ganho e da exclusão da possibilidade de saturação do sinal para as várias etapas de registro, além do procedimento de edição das amostras coletadas para a etapa posterior de análise (Baken, 1987; Colton, Conture, 1990; Orlikoff, 1991; Titze, 1994, 1995; Orlikoff, 1995; Vieira, 1997; Orlikoff, 1998; Baken, Orlikoff, 2000; Fourcin, 2000; Mancini, 2001; Vieira et al, 2002).

## 5.1. Composição do grupo estudado e seleção de indivíduo referência

Os procedimentos apresentados para avaliação conjugada das amostras de fala de indivíduos portadores de disfonia foram aprovados pela Comissão de Ética Médica do Hospital do Servidor Público Municipal da cidade de São Paulo (HSPM), segundo memorando n.º 11/00 e respeita os preceitos de pesquisas envolvendo seres humanos segundo resolução 196, de outubro de 1996, promulgada pelo Conselho Nacional de Saúde.

Cinco indivíduos portadores de disfonia subseqüente à instalação de paralisia unilateral de prega vocal por lesão do nervo laríngeo recorrente (ramo do X par craniano) ou quadros similares, como fixação unilateral de prega vocal, foram inicialmente selecionados. Todos encontravam-se em acompanhamento nas clínicas de Otorrinolaringologia e Cirurgia de Cabeça e Pescoço da referida instituição hospitalar e eram atendidos no setor de Fonoaudiologia.

Foram selecionados indivíduos duração variada da com sintomatologia, como forma de contemplar as fases de evolução de alteração vocal, enquanto a variabilidade dos efeitos sonoros inerente a tal situação. Tal escolha incluiu a possibilidade de retomada de mobilidade de prega vocal, uma vez que as causas de quadros de referida natureza podem levar a manifestações transitórias, de forma que indivíduos com histórico recente possam apresentar recuperação da função glótica, principalmente no período inicial de seis meses, como ocorreu para dois indivíduos previamente selecionados. Um deles foi eliminado por não mais apresentar as condições anteriormente descritas. O segundo caso, relativo à retomada de mobilidade por fixação de prega vocal após intubação prolongada, apresentava importante fibrose em região glótica, tanto em porção membranosa, quanto cartilaginosa.

Dessa forma, aspectos cicatriciais e atrofia de tecidos compunham um quadro igualmente relevante para a caracterização de alteração vocal do ponto de vista fisiopatológico delineado, em que se destacam as alterações de controle de fluxo aéreo por falhas de fechamento glótico, de rigidez e de mobilidade das prega vocais, bem como de desequilíbrio de tensão e massa entre ambas, em acordo com Dejonckere, Wienecke (1994) e Isshiki (1999b).

A casuística estudada constou de quatro indivíduos adultos do sexo feminino, com idades variando entre 51 e 72 anos, portadores de disfonia subseqüente à instalação de quadro de paralisia ou fixação unilateral de prega vocal, com evoluções variadas, referidos como i1, i2, i3 e i4 e um indivíduo referência da mesma faixa etária (52 anos) e sexo, sem antecedentes de alterações vocais, consumo de fumo e patologia de base respiratória, referido como iR. Todos os integrantes do grupo estudado (i1 a i4) e indivíduo referência (iR) eram falantes nativos do português, sem histórico de bilingüismo.

Apenas um dos indivíduos avaliados era representativo de faixa etária mais avançada (72 anos) em relação aos demais (entre 51 e 54 anos) e foi mantido no estudo por considerarmos a sintomatologia apresentada (disfonia), mesmo diante das possibilidades de associação de componentes relativos à manifestação de presbifonia, uma vez que o enfoque do estudo é justamente a investigação da esfera de alteração do sinal vocal. Além disso, os efeitos da presbifonia apenas acentuam as condições fisiopatológicas das disfonias anteriormente delineadas e podem inclusive enriquecer a discussão dos achados.

No quadro 2 seguem descritas as características referentes à idade, naturalidade, grau de escolaridade e profissão e do grupo estudado (i1 a i4) e do indivíduo referência (iR).

Quadro 2. Descrição da idade, naturalidade, grau de escolaridade e profissão dos indivíduos portadores de alterações vocais (i1 a i4) e do indivíduo referência (iR)

Indivíduo	Idade (anos)	Naturalidade (cidade e estado)	Grau de Escolaridade	Profissão
i1	54	Tatuí- SP	Universitário	Assistente social aposentada
i2	51	Princesa Isabel -PB	Segundo grau incompleto	ADI (auxiliar de desenvolvimento infantil)
i3	53	Palmital - SP	Primeiro grau completo	Inspetora de alunos
i4	72	Carlópolis - PR	Primeiro grau incompleto	Auxiliar de enfermagem aposentada
iR	52	Santos- SP	Universitário	Professora

No quadro 3 são expostas a causa e a duração do quadro clínico apresentado pelos indivíduos do grupo estudado (i1 a i4). Nenhum deles recebeu tratamento cirúrgico voltado à minimização das seqüelas impostas pelas alterações de mobilidade de pregas vocais. Tais informações foram obtidas em prontuário médico-hospitalar seguindo roteiro de consulta apresentado no anexo 2.

Quadro 3. Descrição de causa e duração (em meses) do quadro clínico apresentado pelos indivíduos portadores de alterações vocais (i1 a i4)

Indivíduo	Quadro clínico	Causa	Duração (meses)
i1	Paralisia de prega vocal direita	tireoidectomia total e paratiroidectomia	19
i2	Paralisia de prega vocal esquerda	tireoidectomia total	120
i3	Fibrose e atrofia de pregas vocais	Intubação prolongada	13
i4	Paralisia de prega vocal esquerda	tireoidectomia total	13

A composição de casuística concentrada em indivíduos do sexo feminino é compatível com referências da literatura, no tocante à maior demanda de atendimento para indivíduos do sexo feminino em serviços hospitalares (Askenfalt, Hammarberg, 1986), inclusive na instituição hospitalar em questão (Camargo et al, 1996, 2001). Além disso, vale reforçar a menor quantidade de estudos sobre a casuística feminina, em comparação aos estudos na população do sexo masculino (Titze, 1989).

Dessa maneira, procurou-se compor um grupo o mais homogêneo possível em termos sexo e causa de alteração vocal, guardando as características institucionais de prevalência de demanda concentrada no sexo feminino, bem como da necessidade assinalada de desenvolvimento de maiores estudos na voz da população feminina e, conseqüentemente de suas alterações.

Os motivos que permeiam a composição do grupo estudado referem-se justamente àqueles apontados por Dejonckere, Villarosa (1986) para exclusão da manifestação de paralisia de prega vocal no estudo das disfonias, em que

se destaca a variabilidade de achados vocais. Tal fato, na presente pesquisa, representou o potencial para estudo das possibilidades de adaptações do aparelho fonador ao mesmo impedimento de base: as limitações de mecanismos de ação glótica (laríngea) e suas interferências nos demais níveis de produção vocal.

Nesse sentido, a descrição que se segue está centrada na alteração do sinal vocal e não propriamente no estudo das conseqüências da paralisia de prega vocal ou de fibrose de região glótica na produção vocal. O grupo eleito representou apenas uma forma de incursão ao campo das disfonias. Tal consideração é fundamental para a devida apreciação dos dados a serem discutidos, divergindo de outros estudos que utilizaram a análise acústica como meio de caracterizar os mecanismos compensatórios de populações portadoras de alterações de funcionamento do trato vocal no momento da fonação, como é o caso de trabalho anteriormente realizado pela autora (Camargo, 1996), o qual suscitou uma série de indagações que culminaram na elaboração da presente pesquisa.

# 5.2. Concepção geral das etapas de coleta de amostras

Todos os indivíduos do grupo estudado (i1 a i4) foram submetidos a duas sessões de coleta de dados realizadas no mesmo dia, no período da tarde, a fim de minimizar o registro de variações de voz em função de ajustes compensatórios instalados em diversos momentos do dia (Löfqvist, 1986). As etapas de coleta foram realizadas em ambientes distintos à instituição hospitalar, referindo-se ao Laboratório de Rádio da Faculdade de Comunicação e Filosofia COMFIL (PUC/SP) e a clínica médica privada especializada em otorrinolaringologia.

Nessa concepção, os indivíduos denominados i1, i2, i3 e i4 passaram por coleta de amostras de fala e exame otorrinolaringológico, enquanto iR apenas pela primeira etapa. Como justificativa para tal procedimento destacamos que a adoção de indivíduo referência para estudos de base fonético-acústica envolveu o levantamento de particularidades da emissão oral de falantes de determinada comunidade lingüística e, nesta situação delineada, de ausência de condições características de quadro de disfonia.

Nos anexos 3 e 4 encontram-se respectivamente cópias de termo de solicitação e esclarecimento para participação em pesquisa e de termo de consentimento pós informado do sujeito do estudo ou responsável legal, segundo normas da Comissão de Ética do Hospital do Servidor Público Municipal, nos quais os indivíduos participantes inicialmente selecionados foram informados sobre as etapas de coleta de dados, sua finalidade e ausência de risco para os procedimentos em questão. Na situação de coleta das amostras, todos os indivíduos avaliados revelaram autodeterminação no consentimento ao uso da informações clínicas de histórico e exames físicos, laboratoriais, gravações em áudio e vídeo para fins de pesquisa e publicação exclusivamente de cunho científico, sendo garantidos seus direitos no que se refere à intimidade, anonimato, confidencialidade e possibilidade de se retirar do estudo sem que seu acompanhamento institucional seja prejudicado.

Na primeira etapa de coleta de amostras de fala, os indivíduos do grupo estudado (i1 a i4) e o indivíduo referência (iR) passaram por sessão de gravação, a qual constou do registro simultâneo de gravação digital e eletroglotográfica das amostras de fala. Num momento posterior, com intervalo mínimo de uma hora e meia e máximo de três horas, foi realizada a coleta de dados laringológicos provenientes do procedimento de exame otorinolaringológico com apoio de avaliação instrumental de técnica de telescopia laríngea, associada aos recursos de videoestroboscopia e de videoquimografia para os indivíduos do grupo estudado (i1 a i4). As justificativas para a gravação da amostra de fala anteriormente ao exame otorrinolaringológico devem-se às possíveis interferências dos efeitos da anestesia tópica nasal e oral em padrões vibratórios e acústicos das emissões (Sorensen et al, 1980; Södersten, Lindestad, 1992).

# 5.2.1. Coleta das amostras de fala por meios acústico e eletroglotográfico (EGG)

A primeira sessão de coleta de dados envolveu o registro das amostras de fala no Laboratório de Rádio da COMFIL (PUC/SP). Os procedimentos descritos foram previamente estruturados e revistos por equipe composta por três professores do departamento de Lingüística da COMFIL (PUC/SP), atuantes no LIAAC, resultando no planejamento das etapas de preparo do instrumental e do ambiente de

gravação, do indivíduo avaliado e do registro dos dados de ordem acústica e eletroglotográfica, os quais passam a representar os próximos itens deste tópico.

## 5.2.1.1. Preparo do instrumental e do ambiente de gravação

O eletroglotógrafo *EG2* da *Glottal Enterprises* (cedido pelo Laboratório de Voz e Fala do Centro de Especialização em Fonoaudiologia Clínica - CEFAC) foi conectado à entrada de linha da mesa de som *Mackie Microseries 1202 VLZ* (12 canais). Na entrada de microfone foi conectado o modelo de acoplagem em cabeça *audiotechnica ATM 75*, unidirecional, diagrama polar cardióide, elemento condenser, resposta de freqüência de 60 a 15000 Hz, curva de ganho linear. Nas saídas da mesa de som descrita, o PAN no lado esquerdo carregou a informação do microfone e o PAN à direita conduziu o sinal elevado à entrada de linha proveniente do eletroglotógrafo (*EG2*), permitindo o registro simultâneo de ambos sinais, acústico e eletroglotográfico.

O preparo da sala de gravação iniciou-se com a checagem da melhor posição a ser ocupada pelo falante e pelos instrumentos necessários ao teste. Foram necessários ajustes de regulagem de altura do suporte de cabeça da cadeira e previsão da sua distância em relação à mesa na qual se apoiavam o eletroglotógrafo, considerando-se a extensão do cabo com os eletrodos de contato. Tais cuidados visaram evitar sensações desagradáveis ao indivíduo testado.

A cadeira para acomodação do indivíduo avaliado foi adaptada para conter apoio de braços e de cabeça, a fim de minimizar movimentos de cintura escapular que pudessem interferir nos registros. No local do solo previsto para o apoio de seus pés foram posicionados tapetes de borracha para promover melhor isolamento e proteção em função da corrente elétrica provida pelo EGG, lembrando que a mesma é de baixa amplitude e alta freqüência, sendo praticamente imperceptível ao indivíduo avaliado.

Outro cuidado constou da fixação do texto estímulo em suporte sobre mesa situada em frente ao local a ser ocupado pelo falante, de forma a minimizar

a necessidade de movimentação corporal, especialmente de cintura escapular e membros superiores para leitura do texto. O texto apresentado foi impresso em duas folhas A4 com fonte Comic Sans (tamanho 18), com o cuidado em efetuar a divisão de página em ponto que não comprometesse a leitura ou gerasse pausa não relacionada ao padrão habitual do falante.

Vale ressaltar que, no momento da coleta, permaneceram no estúdio de gravação com tratamento acústico apenas o indivíduo testado e um examinador. Neste ambiente, os equipamentos restringiram-se ao eletroglotógrafo e ao microfone de cabeça, sendo que o restante do instrumental descrito encontrava-se em outro ambiente anexo também tratado acusticamente, bem como profissional técnico responsável pelo monitoramento do sinal registrado e manipulação do instrumental. Tal modo de disposição primou pela qualidade das amostras registradas, sem interferências nos traçados eletroglotográficos, especialmente pela manutenção de distância maior do indivíduo avaliado e da unidade de eletroglotografia em relação aos demais equipamentos, situados em sala anexa, sendo recomendável aquela mínima de um metro. Dessa forma, foram adotados vários procedimentos recomendados em nota técnica recente voltada ao procedimento de eletroglotografia (Rothenberg, 2002). No momento da coleta dos dados, o indivíduo avaliado esteve posicionado a uma distância mínima de 1,5 m dos equipamentos, numa sala com área de 9 m<sup>2</sup> e isolamento acústico, diminuindo de forma considerável as possibilidades de interferência de ruído externo e de demais equipamentos nas amostras de fala registradas.

Na sala anexa ao estúdio de gravação foram ligados e testados os equipamentos a serem utilizados, checados níveis de interferências e ruído. Um tripé com câmera de vídeo *RCA DSP3 Hi8 24x Zoom Plus* e fita *Sony* 8mm vídeo MP 120 foi posicionado nesta mesma de sala, de forma a manter o foco no indivíduo testado (cabeça e tronco), bem como no painel frontal do eletroglotógrafo. A imagem registrada incluiu a movimentação do falante, a evolução da posição dos eletrodos reproduzida no painel do eletroglotógrafo, referindo-se à oscilação de luzes que reflete as variações da distância relativa entre o nível da glote e os pontos médios superior e inferior do par de eletrodos, apontando conseqüentemente ajustes dos eletrodos e/ou da posição de laringe no pescoço, graças a indicadores de laringe alta (ou eletrodo baixo) e viceversa. O registro de som foi efetuado graças aos equipamentos de controle de

gravação, com reprodução do sinal registrado no estúdio de gravação por meio de altofalantes *JBL* modelo *control S*. Dessa forma, apesar de permanecer fora da sala de teste, a câmera de vídeo registrou o estímulo de fala sincrônico aos movimentos do falante e sinais do painel do eletroglotógrafo, possibilitando a detecção de movimentos corporais, especialmente na região cervical, que pudessem ter interferências no traçado eletroglotográfico registrado.

As fitas digitais (DAT) Sony PDP-35C 18m/59ft Pro Dat Plus para uso profissional utilizadas na gravação foram previamente instaladas no equipamento de gravação, sendo acionada a opção rápido avanço, de forma a permitir que rodassem uma vez em toda sua extensão anteriormente à gravação. Além disso, não foi utilizado o trecho inicial da fita, de forma que as gravações tiveram início no tempo correspondente a dois minutos. Tais cuidados foram adotados na tentativa de eliminar interferências de pequenas ondulações na superfície da fita que pudessem comprometer os registros sonoros. Quanto ao microfone, foi efetuada a reposição de bateria (maxell alkaline ace AA 1,5 V) mantendo a opção de resposta plana de freqüência.

No anexo 5 encontra-se a síntese dos procedimentos relatados no roteiro de preparo do instrumental (etapas 1 e 2), referida no presente item, além de preparo do indivíduo avaliado (etapa 3), a ser descrita no próximo tópico.

### 5.2.1.2. Preparo do indivíduo avaliado

O preparo de cada indivíduo avaliado na etapa de coleta de amostras de fala incluiu sua acomodação no ambiente de teste descrito, em posição sentada, quando foi solicitada a remoção de colares ou demais objetos metálicos que pudessem interferir no circuito de corrente elétrica estabelecido pelo equipamento de eletroglotografia. A pele do pescoço foi previamente limpa com lenços de papel para remover o excesso de oleosidade, e, em seguida, com algodão embebido em álcool etílico hidratado anti-séptico tópico. Os eletrodos de superfície (34 mm de diâmetro; voltagem máxima entre eletrodos de 1,5 V; corrente entre eletrodos prevista para aproximadamente 10 mA na dependência da impedância do pescoço e condutor de fregüência de 2 MHz) receberam uma camada de gel condutivo hipoalergênico para

eletrodos *Spectra 360* (*Parker Laboratories*) em suas porções metálicas, sendo acoplados às alas da cartilagem tireóide manualmente e presos com uma fita de velcro em torno da circunferência do pescoço, garantindo firme contato com a pele da região cervical. O microfone de cabeça, cujas especificações foram anteriormente descritas, foi posicionado do lado direito, com distância mantida a 3 cm da comissura labial direita.

Anteriormente ao início das gravações foi observada a posição dos eletrodos, monitorada pela opção trajetória da laringe no painel frontal *EG2* anteriormente descrito. Enquanto os eletrodos eram ajustados ao pescoço, foi realizado um treinamento prévio da gravação, o qual permitiu familiaridade com os estímulos de fala, microfone e eletrodos posicionados, bem como com a situação de avaliação por parte do indivíduo avaliado. Durante a coleta de dados procurou-se registrar as emissões em padrão habitual e mais confortável possível. Também houve registro de emissões com elevação ou diminuição dos níveis de freqüência e intensidade habituais.

# 5.2.1.3. Registro dos dados de natureza acústica e eletroglotográfica

Os estímulos captados pelo microfone e eletrodos de contato foram digitalizados com o auxílio da placa de som Delta 44 e processados pelo *software Sound Forge 4.0*, monitorados pela reprodução simultânea do traçado acústico (saída do microfone) na porção superior e eletroglotográfico (saída do EG2) na inferior do monitor do PC com processador Pentium III, HD 40 *gigabytes*, memória 256 RAM *megabytes*. Tais amostras também foram registradas em gravador digital Sony DTC A6 em modalidade de gravação estéreo com freqüência de amostragem de 44100 Hz, 16 bits de quantização, para os dois canais, com taxa sinal/ruído situada entre 30 e 60 dB.

Os sinais registrados em freqüências de amostragem de 44100 Hz e 22050 Hz puderam ser editados para finalidade de análise e processamento por outros softwares. Os dados processados e digitalizados pelo programa Sound Forge 4.0 foram gravados em CD-R Nashua 650 megabytes com o uso de gravador de CD Sony CDR – 2x. Cada indivíduo do grupo estudado e indivíduo referência teve suas amostras registradas separadamente em fitas de áudio DAT e CDs, constando de duas faixas

com registro em modalidade estereofônica referente aos sinais acústico e eletroglotográfico, respectivamente digitalizadas em 44100 e 22050 Hz e uma em modalidade monofônica com amostra acústica digitalizada na freqüência de amostragem 22050 Hz.

A opção pela digitalização em duas freqüências de amostragem, 22050 Hz e 44100 Hz foi adotada a fim contemplar diversas possibilidades de análise. Para o caso da primeira, existe a possibilidade de análise até a freqüência de 11025 Hz, sendo possível reduzi-la para valores inferiores na tentativa de obter estimativas mais confiáveis, quando da utilização do algoritmo *LPC*, levando-se em conta sua relação com o ajuste de número de ordem do algoritmo. No caso de 44100 Hz, a indicação reside na possibilidade de investigação acústica de faixas de freqüência mais altas do sinal (até 22050 Hz), as quais podem conter ruído, especialmente em casos de disfonias.

As amostras de fala registradas incluíram repetições de seqüência de três emissões breves, de cerca de um segundo de duração, da vogal [a] (denominada emissão entrecortada de vogais [aaa]), emissões sustentadas da vogal [a] ([a:]) com duração mínima de dois segundos em padrões de freqüência e intensidade habituais e confortáveis, em freqüência e intensidade mínima e máxima e finalmente três leituras de um mesmo texto.

Conforme abordado no tópico referente à avaliação das disfonias, as emissões em padrão encadeado passaram a ser indicadas para avaliação fidedigna da qualidade vocal habitual. A inexistência de um texto padronizado para tal finalidade em nossa língua revela as dificuldades de composição dos procedimentos de pesquisa, bem como limita a comparação de dados obtidos por diferentes estudos. Nesse âmbito, optamos por compor texto que resume várias das necessidades de estudos no LIAAC (PUC/SP), como uma primeira etapa de concepção de instrumentos a serem utilizados em pesquisas subseqüentes no campo da Fonética Acústica.

A partir destas necessidades, foi elaborado um texto inspirado em Attar (1993), prevendo-se duração mínima de 90 segundos de emissão, com os cuidados de composição, no sentido de evitar uso de vocábulos ou expressões pouco conhecidos do público em geral. A etapa de adaptação do referido texto constou da

gravação de emissões de seis falantes do sexo feminino em variadas faixas etárias (21 a 57 anos) e níveis de instrução (primeiro grau incompleto a universitário completo) sem queixas vocais em três tarefas distintas: gravação de leitura no primeiro contato do falante com o texto, ou seja, sem leitura prévia, gravação após leitura silenciosa prévia do texto e, finalmente, gravação após leitura não silenciosa prévia do texto. Após as etapas descritas, o estímulo para registro de emissão encadeada utilizado refere-se ao seguinte texto, composto de 209 palavras, não balanceado foneticamente:

Há um tempo atrás, li uma lenda sobre a jornada de um grupo de pássaros à procura do rei ideal. Para líder do grupo, os pássaros escolheram a águia que era admirada pelas aves por ter vencido o medo de voar a lugares desconhecidos. No dia marcado para o início da viagem, ela reuniu o grupo e procurou motivá-lo a percorrer o caminho. O papagaio, adornado com seu colar de fogo, foi o primeiro a declarar que estava pronto para partir. A arara parecia estar animada e convocou a todos para iniciar a marcha. Sua alegria contagiou o tímido pato, a bela patativa, o valente falcão, a educada codorna, o sabiá branco, o querido uirapuru, a aplicada coruja, o delicado canário, o orgulhoso pavão, a elegante garça, o esperto bicudo, a meiga rolinha, o delicado pardal e o animado pombo. A águia sabia que o trajeto era difícil, e que o rei só seria encontrado por aquele que tomasse o rumo correto. A sábia águia já havia percorrido o caminho e descobrira que só aquele que segue em direção aos vales do amor e da humildade encontra a realeza dentro de si. Um pássaro em viagem representa o homem com suas fraquezas, ideais e qualidades em busca do criador.

O roteiro das gravações digital e eletroglotográfica das amostras de fala dos indivíduos do grupo estudado e indivíduo referência é apresentado no anexo 6, o qual sinaliza os passos seguidos de forma sistemática para todos os falantes avaliados nesta etapa. A partir da acoplagem dos eletrodos e adaptação de microfone, teve o início dos registros dos estímulos acústico e eletroglotográfico, encerrados apenas com a remoção dos eletrodos e microfone.

Na primeira fase de gravação, o eletroglotógrafo permaneceu na opção de análise em VFCA (área de contato de pregas vocais), de forma que a polaridade da onda representada é positiva para o contato. O limite inferior de freqüência (LF) esteve ajustado para 40 Hz, o mais indicado para vozes nas faixas feminina e infantil. O ganho foi ajustado em baixo, pois havia controle técnico no caso de necessidade de aumento de ganho do sinal.

A primeira solicitação de emissão para gravação constou de

seqüências de emissões entrecortadas de vogais [aaa], variando de duas a cinco tentativas, de acordo com a facilidade de execução e padrão vocal. Em seguida, solicitou-se a emissão de duas vogais [a] em padrão sustentado com duração mínima de dois segundos em nível habitual e confortável de freqüência e intensidade.

Na seqüêcia, cada indivíduo avaliado foi solicitado a manter pressão digital sobre a superfície dos eletrodos. A examinadora ofereceu o padrão, segundo a preferência manual do examinando, estimulando o contato firme do eletrodo com a pele, porém sem realizar pressão sobre a ala da cartilagem tireóide, a qual caracterizaria manobra de coaptação glótica a favorecer a adução e modificar o ajuste habitual laríngeo. A partir deste ponto, todo o teste incluiu o apoio manual nos eletrodos. Tal manobra encontrou respaldo em diversas referências de estudos na área (Colton, Conture, 1990; Vieira, 1997), bem como em teste prévio realizado para a condução do presente trabalho.

Nesta situação, foram mantidos os ajustes do eletroglotógrafo em VFCA (padrão de contato de pregas vocais), LF 40 Hz e ganho baixo. As falantes foram estimuladas a emitirem a vogal [a] em padrão habitual e confortável por uma vez, em seguida emitir a vogal [a] em variadas situações na seqüência que se segue: grave, o mais grave possível, aguda, o mais aguda possível, padrão habitual e confortável, fraca, o mais fraca possível, forte, o mais forte possível, padrão habitual e confortável, cinco tentativas em padrão entrecortado e finalmente nova leitura do texto.

Na última etapa da gravação, ainda com pressão digital sobre os eletrodos, o eletroglotógrafo teve seu ajuste modificado para DEGG (eletroglotograma diferenciado). Neste momento, o sinal apareceu registrado com polaridade VFCA (área de contato de pregas vocais) e refletiu a taxa de mudanças do sinal até a freqüência de 3 kHz, sendo indicado para localizar os momentos de abertura e fechamento glóticos. Por enfatizar os componentes de ruído de alta freqüência, o sinal registrado pode conter ruído. Nesta opção, foram solicitadas a emitir vogal em padrão entrecortado [aaa] por cinco tentativas e sustentada [a:] por uma ocasião, finalizando com uma nova leitura do texto.

As etapas de gravação descritas foram repetidas sistematicamente

por um mesmo examinador que permaneceu no estúdio de gravação. Na sala anexa estiveram presentes um técnico de gravação, um examinador que auxiliava o monitoramento da qualidade de gravação e registros e um técnico de vídeo responsável pelo controle e registro das imagens.

A coleta de dados foi realizada individualmente, de forma que as demais falantes avaliadas permaneceram fora das dependências do estúdio de áudio, enquanto aguardavam o momento da teste. Além disso, cada falante recebeu o texto para leitura previamente, juntamente com treinamento e esclarecimento de dúvidas referentes ao texto por parte de um dos examinadores.

# 5.2.2. Exame otorrinolaringológico

No mesmo dia, logo em seguida à coleta das amostras de fala, os integrantes do grupo estudado (i1 a i4) passaram por exame otorrinolaringológico em clínica particular, conduzido por profissional médico especialista em otorrinolaringologia, com experiência clínica na interpretação de achados laringológicos e seguindo roteiro apresentado no anexo 7.

Os exames foram realizados com o equipamento de telescopia de laringe (endoscópio rígido 70° Modelo LY-C530 Machida), associado a videoestroboscopia (estroboscópio *Brüel & Kjaer type 4914 -Rhino-Larynx Strobocope*) e videoquimografia (videoquimógrafo modelo 8900 da *Kay Elemetrics Corp.*, manufaturado pela *Lambert Instruments Company-* CDD Câmera) para avaliação de atividade laríngea.

Os examinadores da etapa anterior permaneceram no ambiente de exame e auxiliaram a solicitação de emissões específicas para a situação de coleta de dados. Quando o procedimento proposto não foi possível devido ao excessivo reflexo nauseoso apresentado pelo indivíduo avaliado, optou-se pela realização do exame por fibroscopia laríngea (endoscópio flexível) associando-se à técnica de estroboscopia.

Tal situação ocorreu apenas com i2, a qual mesmo após anestesia tópica apresentou acentuado reflexo de náusea, o que impedia a observação da laringe em atividade por meio do endoscópio rígido. Pelo mesmo motivo, não houve possibilidades de realização da videoquimografia. Nessa situação, foi utilizado o endoscópio flexível *ENT* 30 P3 *Machida* juntamente com o estroboscópio referido. Os exames foram realizados seguindo a mesma rotina, com variações impostas pelas próprias condições do exame ou tolerância do indivíduo testado ao procedimento em questão.

O preparo para esta etapa de coleta de dados incluiu a acomodação do indivíduo avaliado em cadeira, a remoção de próteses dentárias e a aplicação de anestesia tópica por xylocaina 10% em spray direcionada à orofaringe. Durante o período em que se aguardava o efeito, foram colhidos dados relevantes de histórico e de saúde pelo profissional médico.

Em seguida, foi introduzido na cavidade oral o telescópio de laringe. A fonte de luz utilizada foi modelo *RG* 2500A *Machida*. Num primeiro momento foram observados o aspecto geral e a integridade das estruturas laríngeas. Na seqüência, registrou-se o comportamento laríngeo à fonação, inicialmente durante a emissão sustentada da vogal [e], a qual favorece a elevação da laringe, facilitando sua exposição e iluminação do complexo laríngeo no momento do exame.

A avaliação estroboscópica permitiu a observação de detalhes da periodicidade do sinal, do fechamento glótico, da onda mucosa, da simetria de fase, de amplitude de movimentos vibratórios e da presença de atividade em região supraglótica (Hirano, Bless, 1993; Tsuji et al, 1999a, 1999b), colaborando para avaliação funcional da fonação dos indivíduos do grupo estudado. Quando possível, foram registradas emissões próximas ao som da vogal [a], por ser a mais representativa em termos de avaliação de medidas e representações acústicas, de acordo com referências apresentadas no capítulo específico de avaliação vocal acústica, reconhecendo-se a dificuldade imposta às mobilizações de língua pelo posicionamento do telescópio.

Na etapa final da referida avaliação foi observada a emissão de vogal sustentada, preferencialmente, em padrão de emissão habitual, sendo alternadamente efetuada pressão leve dos dedos de um dos examinadores nas alas da cartilagem tireóide, simulando a ação exercida sobre os eletrodos, de forma a avaliar a

possível interferência da manobra de apoio digital durante a etapa anterior de coleta de dados. Houve certa dificuldade inicial em se executar a manobra dada a necessidade de acoplagem do microfone do estroboscópio ao pescoço do indivíduo avaliado, justamente na ala da cartilagem tireóide. Vencida tal dificuldade, foi possível observar que não houve interferências da referida manobra no padrão de fechamento glótico, bem como de ondulação de mucosa. Vale ressaltar que o examinador que efetuou a manobra foi a mesma presente no estúdio de gravação na primeira sessão de coleta de dados, tendo orientado a pressão digital sobre os eletrodos de superfície ao pescoço.

Como última etapa da realização do exame, foi acoplado o equipamento de videoquimografia ao telescópio rígido anteriormente descrito e os falantes foram solicitados a manter emissões em padrão sustentado que permitissem a sincronização do videoquimógrafo à emissão. A partir do julgamento das imagens em padrão quadro a quadro do vídeo, foram detalhados os aspectos de periodicidade, fechamento em fase, nível vertical de aproximação das pregas vocais, simetria de fase e de amplitude de movimentos vibratórios e onda mucosa (Švec, Schutte, 1996; Tsuji, Sennes, 1998; Larsson et al, 1999, 2000; Tsuji et al, 2001).

Todas as imagens dos exames foram registradas em fitas de vídeo VHS T-30, novamente destinando-se uma fita para cada indivíduo avaliado. Vale ressaltar que o equipamento utilizado não possuía sistema de conexão para eletroglotografia, de forma que não foi possível o registro eletroglotográfico simultâneo das informações para esta etapa de teste.

Ao término do exame, os indivíduos avaliados foram orientados quanto aos cuidados relativos à anestesia tópica e liberados. A partir deste momento as gravações referentes ao exame foram reproduzidas e cuidadosamente analisadas pelo examinador médico, compondo os dados de avaliação laringológica segundo anexo 7, com referência às condições gerais de exame, integridade e mobilidade das estruturas laríngeas, bem como de parâmetros relacionados à dinâmica vibratória, anteriormente enfocados com relação à videoestroboscopia e videoquimografia.

Dessa maneira, a interpretação dos achados de exame foi realizada imediatamente após o exame, permitindo maior detalhamento do quadro clínico e a impossibilidade de troca das imagens ou de laudos. Ao fim dos exames,

todos os registros de imagem e laudos estavam completos, encontrando-se disponíveis, para cada indivíduo testado, uma fita de vídeo e um laudo de exame, cujas cópias foram direcionadas ao SAME (serviço de arquivo médico) do HSPM para arquivo em prontuário médico-hospitalar.

# 6. MÉTODOS DE EDIÇÃO, TRATAMENTO E ANÁLISE DOS DADOS DE ORDEM ACÚSTICA, PERCEPTIVA E FISIOLÓGICA

A apresentação deste capítulo contempla as etapas de edição e tratamento das amostras de fala e de dados laringológicos coletados, das quais resultaram as possibilidades de análise acústica, perceptivo-auditiva e fisiológica da qualidade vocal, representando respectivamente as seções do presente tópico de exposição de métodos de análise. O encerramento do capítulo foi destinado à descrição de experimento voltado à validação do estímulo utilizado na análise acústica de longo termo, o qual é enfocado enquanto extensão de método da presente pesquisa.

# 6.1. Edição e tratamento das amostras de fala

As amostras de fala inicialmente registradas segundo procedimentos descritos no capítulo anterior foram submetidas a uma sessão de inspeção geral, a fim de verificarmos a qualidade do sinal, o nível de ganho e a ausência de saturação, a garantirem as possibilidades de análise. Diante desta comprovação inicial, os estímulos foram editados de acordo com os critérios de inspeção adotados.

Do total das amostras, os estímulos selecionados para análise acústica e eletroglotográfica referiram-se às emissões de vogal [a] nas modalidades sustentada e entrecortada e trechos de emissão encadeada. Daquelas vogais emitidas em padrão sustentado [a:], foram selecionadas as cinco primeiras emissões. Para o padrão entrecortado de emissões vocálicas [aaa], optou-se por selecionar duas seqüências da primeira etapa de gravação (sem apoio digital nos eletrodos e avaliação do padrão de contato de pregas vocais), duas da segunda (com apoio digital nos eletrodos e avaliação do padrão de contato de pregas vocais) e uma da última etapa (com apoio digital nos eletrodos e avaliação pelo DEGG). Algumas mudanças ocorreram em função da necessidade de eliminação de gravações pouco representativas da emissão habitual do indivíduo em questão.

Além das vogais sustentadas e entrecortadas, a edição das amostras incluíram trechos de sentença contendo a vogal [a] em diferentes contextos fonéticos: "a arara", na qual, além do artigo inicial, encontrava-se nas sílabas prétônica, tônica e pós-tônica, além de vocábulos contendo sons consonantais plosivos surdos e sonoros. Tais estímulos também foram destacados das leituras do texto e submetidos a várias modalidades de análise acústica na dimensão de curto termo, bem como na eletroglotografia. Os últimos estímulos editados referiram-se às gravações correspondentes às leituras do texto na íntegra, apenas com exclusão de pausas silenciosas para processamento por análise de longo termo. O estudo apresentado ao final deste capítulo justifica tal procedimento.

Quanto à composição do *corpus* de análise, especialmente com relação à seleção de vogal, vale ressaltar que apesar de controvérsias a respeito da seleção de vogal para a rotina de avaliação vocal e, mais especificamente para análise acústica, vários estudos recentes conduzidos por um mesmo grupo de pesquisadores, utilizando metodologia similar, reforçaram a importância da vogal [a] no contexto de avaliação vocal, especialmente no caso de disfonias (Vieira et al, 1996a, 1996b; Vieira, 1997; Vieira et al, 1997). Dessa maneira, a edição das amostras possibilitou condições de análise da referida vogal em diferentes contextos de emissão, colaborando para detectar o impacto da disfonia nas condições de comunicação oral do falante.

Os estímulos silábicos também foram submetidos a procedimentos de análise acústica de curto termo, mais especificamente à extração do tempo de início de sonorização (*VOT*), de sílabas inseridas nos vocábulos **pa**ra, es**per**to e **be**la para representarem o par de consoantes plosivas bilabias [p] e [b], es**tar** e humil**da**de para o par de consoantes plosivas dentais/alveolares [t] e [d] e apli**ca**da e lu**ga**res para o par de consoantes plosivas velares [k] e [g].

A seleção de consoantes plosivas surdas ([p] [t] [k]) e correspondentes sonoras ([b] [d] [g]) situadas em posição tônica guardam similaridade de contexto fonético, na medida em que o som subseqüente à vogal da sílaba investigada apresentava ponto de articulação alveolar ([r], [l] e [d]). Dessa forma, o contexto de análise foi semelhante, minimizando interferências nas análises, de acordo com estudo de Shimizu (1996), o qual aponta variações de medidas de *VOT* em função

do sexo e do contexto em falantes de uma mesma língua. Levando-se em conta que não foi utilizada sentença-veículo, devemos considerar a possibilidade de variantes estilísticas, traduzidas em variação de componentes de ordem supra-segmental no momento das leituras, das quais foram extraídas tais sílabas.

Retornando ao tratamento das demais amostras coletadas, aquelas registradas em modalidades estereofônica e, portanto contendo informações acústicas e eletroglotográficas (Lx), digitalizadas em freqüência de amostragem de 22050 Hz (16 bits) foram submetidas a uma série de programas desenvolvidos para eliminar os efeitos de flutuação de linha de base nas estimativas de f0 (Vieira et al, 1996a; Vieira, 1997)) e em medidas de perturbação para amostras registradas por meio da eletroglotografia (Vieira, 1997).

Inicialmente, o programa *S2mono.exe* converteu cada estímulo estereofônico em dois arquivos monofônicos, separando o sinal acústico (esquerdo) e o eletroglotográfico (direito). Em seguida o programa *delay0bat.exe* processou as amostras EGG (sinal direito), efetuando sua filtragem a fim de atenuar as flutuações de linha de base. Tal sistema opera conjugando as ações dos programas *lxfilter.exe* e *revfile.exe*. Graças a esse método, o arquivo original passou por um processo de filtragem passa-altas com corte de 60 Hz e atenuação de 45 dB na faixa de rejeição. A filtragem foi automaticamente efetuada duas vezes: uma no sentido da onda e, em seguida, no sentido inverso, para reverter deformações no sinal referentes à alteração de fase. A atenuação do sinal abaixo de 60 Hz é da ordem de 90 dB para a operação em questão.

Conforme exposto previamente na revisão da avaliação fisiológica por eletroglotografia, a modalidade de filtragem passa-altas é efetuada com o objetivo de eliminar os componentes de baixa freqüência, os quais não guardam relação com movimentos de pregas vocais, mas com outras estruturas cervicais e demais artefatos resultando em traçado irregular no que se refere à estabilidade da linha de base, conforme exposto na figura 2a.

-

<sup>&</sup>lt;sup>(1)</sup> Os programas foram disponibilizados pelo autores, autorizando e orientando o seu uso para o projeto em questão. Tais recursos foram desenvolvidos como parte de tese de doutorado (Vieira, 1997) – The University of Edinburgh e são atualmente utilizados em pesquisas no Departamento de Física da Universidade Federal de Minas Gerais (UFMG).

Do referido processo de filtragem resultaram as ondas EGG filtradas, conhecidas como Lx, as quais representam amostra passível de análise e inclusive de extração de algumas medidas, podendo-se detectar a estabilidade da linha de base na figura 2b.

Para finalidade de análise sincrônica dos estímulos EGG e acústico, respectivamente representadas pelas ondas Lx e acústica, foi necessária nova edição a fim de compor o estímulo estéreo com o sinal acústico (canal esquerdo) e o estímulo EGG submetido à filtragem pelo programa delay0.bat (canal direito). Tal edição foi realizada com o auxílio do programa Sonic Foundy Vegas Áudio 2.0 instalado no mesmo computador do Laboratório de Rádio da COMFIL utilizado para as edições anteriormente referidas, resultando em parte dos estímulos que compõem o banco de dados do presente estudo.

Em etapa subsequente, as amostras foram editadas em estímulos menores, seguindo a mesma divisão realizada para as amostras utilizadas em modalidades de análise acústica.

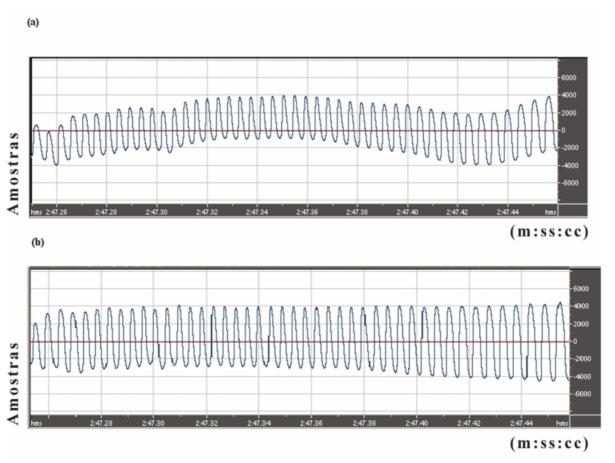


Figura 2. Ilustração dos traçados de eletroglotografia (a) com flutuação de linha de base (b) filtrada para a mesma emissão de vogal [a] destacada de contexto de fala encadeada ("a arara").

#### 6. 2. Procedimentos de análise dos dados

As amostras coletadas foram analisadas em suas vertentes acústica, perceptivo-auditiva e fisiológica (eletroglotográfica), complementadas pelos dados laringológicos. Os critérios de inspeção neste domínio foram delineados em função de descrições e propostas de padronização de análise neste âmbito das alterações vocais (Yanagihara, 1967; Rontal et al 1975; Titze, 1994, 1995; Vieira, 1997; Vieira et al, 2002).

#### 6.2.1. Análise acústica

Os procedimentos de análise acústica aplicados ao banco de dados incluíram a observação de padrões espectrais, principalmente a natureza e a distribuição dos componentes de energia em diversas faixas de freqüência do sinal e a extração de medidas acústicas para as dimensões de longo e de curto termo.

Os equipamentos utilizados nesta etapa referiram-se ao CSL (Computer Speech Laboratory) e software Multispeech, ambos da Kay Elemetrics Corporation, respectivamente instalados no Laboratório de Voz e Fala do CEFAC e no LIAAC da PUC/SP.

Para análises do domínio de longo termo, foram utilizados os estímulos de leitura, para gerar o espectro de longo termo. Na dimensão de análise de curto termo, os estímulos de vogais em padrão sustentado foram utilizados para análise de espectrografia de banda estreita e espectros *FFT*, bem como espectrogramas banda larga e espectros *LPC* (*Linear Predictive Coding*). Das emissões de vogais em padrão entrecortado foram eleitas para análise a vogal central da seqüência de estímulos, nas modalidades de espectrografia de banda estreita e espectros *FFT*, banda larga e espectros *LPC*. Nos trechos "a arara", os quais privilegiaram a inspeção da vogal [a] em diversos contextos fonéticos, foram aplicadas modalidades de tratamento referentes à espectrografia de banda larga e espectros *LPC*, além daqueles espectros *FFT* e análise cepstral (por *cepstrum*).

# 6. 2.1.1. Dimensão acústica de longo termo

Nesta fase, foram registrados os traçados dos espectros de longo termo (ELT), com parâmetros de análise em 256 pontos, pré-ênfase de 0,98, sem suavização, janelamento *Hanning*, eixo de intensidade representando mínimo em 0 (zero) e máximo em 80 dB e limite superior de análise de freqüência em 11 kHz. Foram eliminados trechos iniciais e finais da leitura, nas quais não havia identificação de produção sonora, bem como pausas demarcadas por período de silêncio. As medidas de longo termo correspondentes à análise de 256 pontos permitiram traçar estimativas das tendências de declínio da energia espectral em intervalos de freqüência correspondentes a 86,13 Hz entre 0 e 11 kHz para cada uma das três leituras efetuadas pelos indivíduos do grupo estudado (i1 a i4) e indivíduo referência (iR).

Os trechos de leitura processados registraram variações de 77,47s a 304,97s de duração, considerando-se cada uma das três repetições dos falantes do grupo estudado e indivíduo referência (iR). Nessa situação, foram utilizados os estímulos digitalizados na freqüência de amostragem de 22050 Hz. Os procedimentos foram seguidos de maneira consistente em todas as etapas do processamento das informações.

Os valores referentes à variação de amplitude ao longo de 128 pontos de freqüência compreendidos na faixa de 0 a 11 kHz foram submetidos a tratamento estatístico por meio de análise fatorial de componentes principais e análise de construção de *clusters* com auxílio dos programas SPSS 10.0 *for windows* e *SPAD* 3.5, instalados no computador com processador Pentium II, 300 MHz, 32 *megabytes*, memória RAM, HD 41.1 do Laboratório de Informática da Faculdade de Psicologia da PUC/SP.

Numa nova etapa, a análise fatorial de componentes principais e análise de construção de *clusters* foi aplicada às mesmas amostras, com exclusão de iR, seguindo os mesmos passos anteriores. O tratamento de tais medidas também possibilitou a abordagem de médias de distribuição de energia por faixas de freqüência diferenciadas da emissão (0-1 kHz, 1-3 kHz, 3-5 kHz e 5-11 kHz), indicativas do declínio espectral, representadas em modalidades de gráficos e também submetidas

ao procedimento estatístico referido.

#### 6. 2.1.2. Dimensão acústica de curto termo

Para inspeção dos espectrogramas de banda estreita, espectros *FFT* (*power spectrum*) e análise cepstral, foram eleitos os critérios de caracterização da estrutura harmônica, com identificação de indícios de componentes harmônicos e não harmônicos e sua distribuição ao longo das faixas de freqüência, com limites superiores estabelecidos em 2,5 kHz e 5 kHz (amostras digitalizadas em 11025 Hz), 11 kHz (amostras digitalizadas em 22050 Hz) e 22 kHz (amostras digitalizadas em 44100 Hz). A inspeção incluiu a observação e registro da presença de bifurcações (subharmônicos ou séries de f0) e sua respectiva distribuição.

Os espectrogramas de banda estreita foram gerados a partir do processamento de cinco amostras de [a:] e [aaa] com seleção de 1024 pontos, préênfase de 0,98, janelamento *Hanning*, variando-se a extensão de limite superior de freqüência gradativamente entre 5kHz e 10 kHz (quando as amostras estavam digitalizadas em 22050 Hz) a 22 kHz (para o caso de digitalização em freqüência de amostragem em 44100 Hz).

Os espectros *FFT* foram efetuados a partir da escolha de trecho estável da emissão vocálica, na qual os efeitos de início e declínio da emissão são menores. Optamos pela demarcação do referido ciclo na onda acústica, correspondente ao seu ponto de máxima de amplitude. Os parâmetros de análise referiram-se a 1024 pontos, pré-ênfase de 0,98, janelamento *Hanning*, sem nível de suavização, variando-se a extensão de limite superior de freqüência gradativamente entre 5kHz, 11 kHz (quando as amostras estavam digitalizadas em 22050 Hz) e 22 kHz (para o caso daquelas com freqüência de amostragem de 44100 Hz), de forma a corresponder a cada um dos limites de freqüência estabelecidos.

Para análise cepstral, o intervalo mínimo permaneceu em 0,10 ms e máximo em 12 ms, dependente de parâmetros de *FFT* 11 kHz, janelamento *Hanning*. Em tal etapa de análise, procedemos a extração de medidas de freqüência e amplitude do primeiro harmônico (H1) e segundo harmônicos (H2) e finalmente os valores de amplitude e qüefrência na representação cepstral, os quais propiciaram várias

abordagens de extração de f0.

Os espectrogramas de banda larga e envelopes espectrais (espectros *LPC*) foram eleitos para analisar respectivamente os padrões de representação da energia espectral e a do padrão de formantes, graças à possibilidade de representação dos picos de amplitude ao longo das faixas de freqüência de representação espectral, ampliada para freqüências agudas, gradativamente com o estabelecimento de limites superiores em 5 kHz, 10 kHz e 22 kHz, novamente respeitando os critérios de análise em função da freqüência de amostragem.

Além da caracterização das medidas que refletem o padrão de formantes (freqüência e amplitude dos quatro primeiro formantes), também utilizamos o apoio do algoritmo *FFT*, especialmente para determinar os componentes de freqüência e amplitude próximos ao pico espectral. Os envelopes espectrais não foram traçados para as vogais sustentadas, nem valores de formantes foram registrados em função de nossa opção por caracterizar os eventos representativos da emissão em padrão encadeado, como habitual ao falante em questão e portanto mais próximos da correspondência à dimensão de qualidade vocal.

Os espectrogramas de banda larga foram gerados a partir do processamento das emissões de [a:], [aaa] e trechos de leitura "a arara" com seleção de 75 pontos para análise, pré-ênfase de 0,98, janelamento *Hanning*, variando-se a extensão de limite superior de freqüência gradativamente de 5 kHz, 11 kHz e 22 kHz, novamente respeitando os critérios de seleção de amostras em função da freqüência de amostragem.

No mesmo ponto eleito para análise dos harmônicos na modalidade de análise *FFT*, foi traçado o envelope espectral com ordem do filtro *LPC* 24, novamente para freqüência de amostragem de 22050 Hz, duração da janela 10ms, autocorrelação, pré-ênfase 0,98, amplitude entre –20 a 80 dB, ordem do filtro *LPC* em 24, duração da janela em 10ms, avanço de janela igualmente em 10 ms e janelamento *Hanning*.

No âmbito da espectrografia de banda larga, também foram processados os estímulos eleitos para extração das medidas de tempo de início de

sonorização (*VOT*), cujo procedimento inicial envolveu a seleção de trecho de amostra digitalizada em freqüência de amostragem de 22050 Hz. A extração de medidas envolveu várias etapas, em que constam inicialmente sincronização de janelas, contendo as representações de traçado de forma de onda e espectrograma de banda larga. Os parâmetros de análise nesta etapa seguiram aqueles expostos anteriormente, acrescentando-se o histórico dos formantes, o qual foi gerado a partir de ordem do filtro *LPC* em 24 para amostras digitalizadas em 22050 Hz, duração da janela e avanço em 10 ms, autocorrelação, pré-ênfase 0,98 e janelamento *Hanning*.

Na janela superior, contendo o traçado de forma de onda também foi solicitada a marcação dos períodos nos quais houve sonorização. Em seguida, foi demarcada a plosão (ruído transiente), apontada com seta de traçado contínua na figura 3. Em seguida, comparou-se o intervalo de tempo entre a plosão e o início do vozeamento (seta pontilhada na figura 3). Dessa forma, delimitamos as ocorrências de vozeamento previamente, sincronicamente ou posteriormente à plosão, correspondendo às respectivas referências de *VOT* negativo, zero ou positivo.

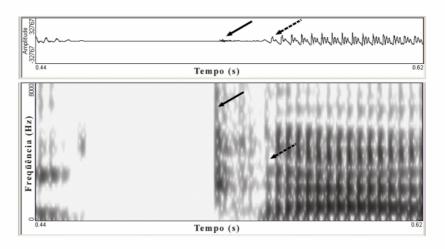


Figura 3. Ilustração de procedimento voltado à extração dos valores de *VOT*, com plosão indicada por seta de traçado contínuo e início da sonorização da vogal com seta de traçado pontilhado nas representações de forma da onda acústica (tela superior) e espectrograma de banda larga (tela inferior) para a emissão da sílaba [ka].

Alguns critérios foram adotados para análise, seguindo a tendência de estudos anteriores (Shimizu, 1996; Madureira et al, 2002). A demarcação do *VOT* positivo foi feita na forma da onda em ponto onde esta realiza cruzamento de zero para a seguir iniciar inclinação ascendente, tendo-se como correspondência, no espectrograma de banda larga, o início da estrutura formântica de F1 e F2.

Quando da presença de pré-vozeamento e da existência de sílaba pré-tônica, casos dos vocábulos "lu**ga**res" e "humil**da**de", o início do vozeamento da consoante plosiva sonora correspondente foi demarcado a partir do ponto em que o som prévio à consoante plosiva deixava de apresentar estrutura formântica (com características de F1 e F2). Os valores obtidos para as medidas acústicas de *VOT* foram submetidos a tratamento estatístico por meio de teste de comparação de médias, com margem de erro estimada em 5%.

No âmbito do tempo de início de sonorização, os eventos representativos na espectrografia foram comparados àqueles dos traçados de forma de onda Lx. Dessa maneira, a onda Lx foi adotada como complemento à inspeção da onda acústica, no sentido de averiguar a provável existência e duração de atividade de pregas vocais para os contextos investigados. Nessa situação, a onda acústica forneceu a informação referente à ocorrência da liberação da oclusão dos articuladores (seta com traçado contínuo na figura 4), enquanto a onda Lx possibilitou a delimitação da informação do provável início de atividade glótica (seta pontilhada na figura 4).

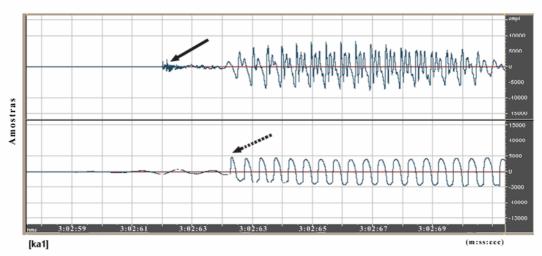


Figura 4. Ilustração de procedimento voltado à extração dos valores de *VOT*, com presença de ruído transiente indicada por seta com traçado contínuo na representação da onda acústica (tela superior) e início da sonorização da vogal por seta com traçado pontilhado na representação da onda Lx (tela inferior) para a emissão da sílaba [ka].

Como consideração final a respeito da extração de medidas de *VOT*, vale ressaltar que além da possibilidade de demarcação do intervalo entre estouro da oclusão e início da atividade glótica, a inspeção da onda Lx criou condições para obtenção de informações adicionais em relação às estimativas de vozeamento anterior, sincrônico ou posterior à liberação da oclusão, especialmente quando levamos em conta as possibilidades de irregularidades do sinal para a população em questão. Tal

procedimento levou em conta as limitações da técnica para a situação, em termos da possibilidade de informação de atividade glótica não ser registrada e de algum efeito que interfira na aparente sincronicidade das amostras no tempo. Conseqüentemente, o registro de ausência ou diminuição de contato foi considerado nesta etapa de avaliação.

Para as informações acústicas relativas à estrutura harmônica, padrão de formantes, bem como aquelas do âmbito da eletroglotografia, a assessoria estatística adotou abordagem qualitativa, identificando suas distribuições e padrões de variabilidade inter e intra-sujeitos, muitos dos quais expressos em gráficos. Foram utilizados os programas SPSS 10.0 for Windows e Microsoft Excel instalados no computador microtec, Windows 95, 64 megabytes de memória RAM do LIAAC (PUC/SP).

Quanto ao tratamento estatístico, os valores de variação de amplitude ao longo de 200 pontos de freqüência compreendidos na faixa de 0 a 11 kHz, obtidos por meio de análise *LPC*, foram submetidos a análise fatorial de componentes principais e análise de construção de *clusters* com auxílio dos programas SPSS 10.0 *for windows* e *SPAD* 3.5 do Laboratório de Informática da Faculdade de Psicologia da PUC/SP, sendo os mesmos procedimentos adotados para análises de longo termo anteriormente descritas.

#### 6.2.2. Avaliação perceptivo-auditiva

As amostras de fala (versão áudio) foram editadas de forma a compor duas gravações para julgamento auditivo, referentes respectivamente à avaliação perceptiva da qualidade vocal com motivação fonética e ao julgamento de inteligibilidade de sílabas. Ambas edições foram realizadas no Laboratório de Rádio da COMFIL (PUC/SP) e gravadas em CD Rom, utilizando-se dos recursos técnicos e profissionais anteriormente descritos quando da coleta dos dados. Cada gravação foi apresentada a grupos distintos de examinadores, selecionados em função da formação e experiência para os julgamentos em questão e passam a ser descritos separadamente nos dois itens subseqüentes.

## 6.2.2.1. Proposta de avaliação da qualidade vocal com motivação fonética

O material elaborado para a avaliação da qualidade vocal foi apresentado em forma de um CD e de um roteiro impresso. Para maior compreensão do teor do material, serão descritos os detalhes da edição dos estímulos de áudio, que segue o roteiro apresentado no anexo 8.

O material foi estruturado para a pesquisa em questão, utilizando-se como apoio a gravação que acompanha a publicação original de Laver (1980), de forma que todo o cuidado foi tomado para se preservar o teor da obra, inclusive em face das adaptações expostas no capítulo 4 (item 2) e que se fizeram necessárias em função da evolução de conceitos na área em duas décadas posteriores de estudos, bem como de propostas de avaliação vocal com validação científica e clínica do *VPAS* (Beck, 1988; Laver et al, 1981; Laver, 1994; Laver, 2000).

O referido CD totaliza nove faixas distribuídas em 65 minutos de gravação, contendo os seguintes estímulos e solicitações de tarefas:

#### Etapa de apresentação:

- faixa 1: indicações gerais de identificação do projeto, vinculação institucional dos autores e solicitação para preenchimento de dados referentes à formação e experiência por parte dos juízes (anexo 8a)

#### Etapa 1:

- faixa 2 (etapa 1A)
  - a) solicitação para avaliação do nível geral de alteração vocal e grau de agradabilidade em escala de seis pontos de igual intervalo, na qual 0 (zero) representa ausência de manifestação do componente avaliado e 5 (cinco) a ocorrência em grau máximo
  - b) descrição de qualidade vocal sem determinação de critérios específicos (atribuição livre)

Estímulos: emissões de vogal [a] em padrões sustentado e entrecortado dos indivíduos do grupo estudado (i1 a i4) e do indivíduo referência (iR) apresentadas em ordem aleatória e

com uma repetição para dois dos indivíduos do grupo estudado (anexo 8b).

## - faixa 3 (etapa 1B)

- a) solicitação para avaliação do nível geral de alteração vocal e grau de agradabilidade em escala de seis pontos de igual intervalo, na qual 0 (zero) representa ausência de manifestação do componente avaliado e 5 (cinco) a ocorrência em grau máximo
- b) descrição de qualidade vocal sem determinação de critérios específicos (atribuição livre)

Estímulos: três emissões da sentença "A arara parecia estar animada" dos indivíduos do grupo estudado (i1 a i4) e do indivíduo referência (iR) apresentadas em ordem aleatória e com uma repetição para dois indivíduos do grupo estudado, excluindo-se aqueles repetidos na etapa anterior (anexo 8c).

A estruturação das amostras apresentadas na etapa 1 teve como objetivo verificar a confiabilidade dos julgamentos. A edição envolveu ordem aleatória de apresentação para as duas etapas (1A e 1B), de maneira que não havia correspondência entre a seqüência de apresentações dos falantes nessas duas situações.

Dado o número de falantes e de tarefas, foi possível estruturar duas repetições para 1A, bem como para 1B, de maneira que a repetição envolveu todos os indivíduos do grupo portador de disfonia. Tal tarefa teve a finalidade de investigar o impacto da avaliação vocal de um mesmo falante, por meio de estímulos de natureza diversa, como é o caso das amostras em padrão sustentado e encadeado.

Na fase (1), os juízes tiveram condições de atribuir seu livre julgamento de qualidade vocal, permitindo a investigação do ponto de referência do qual partem para o julgamento de qualidade vocal, anteriormente à solicitação de avaliação baseada em modelo teórico, como será exposto a seguir para a etapa 2.

- Etapa 2 (análise perceptivo-auditiva da qualidade vocal com motivação fonética):
  - Faixas 4 e 5: exposição da fundamentação teórica do procedimento a ser realizado na etapa 2A, constando de considerações sobre o modelo teórico, especificamente aos ajustes supralaríngeos e laríngeos, com referência à demonstração do ajuste de referência: o ajuste neutro. Ilustração de exemplos dos ajustes descritos no roteiro de avaliação, a partir da gravação original da publicação (anexo 8d).

Em seguida, os juízes analisaram a qualidade vocal de acordo com o roteiro oferecido em duas etapas, denominadas 2A e 2B, nas quais deveriam identificar respectivamente os ajustes presentes e atribuir-lhes graus de influência na qualidade vocal.

- Faixa 6 (etapa 2A): identificação dos ajustes presentes segundo roteiro disponível.

Estímulo: mesmas amostras apresentadas na etapa 1B, referentes a estímulos de três repetições da mesma sentença, mantendo nova repetição de emissões de dois dos indivíduos do grupo estudado em seqüência aleatória (anexo 8e).

- Faixa 7 (etapa 2B): Atribuição de graus de influência (1 a 6) dos ajustes presentes na qualidade vocal.

Estímulo: trechos de emissão com duração mínima de 45 segundos, referentes à segunda leitura do texto, desde o início até o enunciado "A arara parecia estar animada". A opção pela segunda leitura refere-se à possibilidade de familiaridade com o texto e com a situação de gravação (anexo 8f).

#### Etapa Final:

- Faixa 8: agradecimentos e encerramento, sugerindo que examinador destaque do roteiro as duas últimas páginas, referentes ao anexo abaixo indicado, como forma de obter a síntese dos parâmetros utilizados para compor o roteiro de avaliação da qualidade vocal com motivação fonética apresentado (anexo 8g).

Vale ressaltar que a avaliação da qualidade vocal em duas passadas (etapas 2A e B) foi referida na literatura pertinente às primeiras propostas de avaliação por meio do roteiro *VPAS* (Laver et al, 1981), porém adaptações mais recentes não mais utilizam-na (Laver, 1994; Laver, 2000). Optamos pela manutenção da avaliação em duas passadas para propiciar a familiaridade dos juízes, especialmente por não ser um instrumento de domínio em nosso meio.

Algumas medidas que foram tomadas para garantir a confiabilidade dos julgamentos e evitar tendências à contaminação nos julgamentos encontram respaldo na literatura pertinente à avaliação perceptiva da voz (Hammarberg, 1980). Seguindo tais tendências, entre as gravações apresentadas aleatoriamente nas etapas 1A-1B e 2A-2B foram inseridos estímulos musicais com cerca de dois minutos de duração, com o objetivo de desviar a atenção do ouvinte dos estímulos anteriormente avaliados, permitindo que seus julgamentos não fossem influenciados por aqueles da etapa anterior.

O grupo de ouvintes experientes em julgamento de qualidade vocal participante dessa etapa foi composto por três fonoaudiólogos e um lingüista, todos familiarizados com os conceitos do modelo fonético para descrição da qualidade vocal proposto por Laver (1980) e procedimentos clínicos e didáticos envolvendo avaliação vocal perceptivo-auditiva. Eram atuantes em avaliação perceptiva da qualidade vocal há no mínimo doze anos, com título de doutor e pesquisas desenvolvidas no campo de estudos da voz.

Cada examinador recebeu um CD e um roteiro de avaliação devidamente encadernado, nos quais os itens apresentados em anexo 8b, 8c, 8e e 8f aparecem em número de sete, um para cada indivíduo do grupo estudado (4) e indivíduo referência (1) e duas repetições (2) variáveis a cada etapa. Foram solicitados a devolver o material no prazo de um mês e a esclarecer as dúvidas que surgissem na execução da tarefa. De um total de cinco roteiros distribuídos, quatro retornaram com respostas dentro do prazo previsto e, portanto, foram utilizados para compor o perfil de qualidade vocal dos falantes em questão (i1 a i4 e iR).

As referências que representaram consenso entre os quatro

examinadores e/ou aquelas referidas em graus a partir de 4 foram consideradas na composição do perfil de qualidade vocal de cada falante.

## 6.2.2.2. Avaliação da inteligibilidade de sílabas

A edição para a presente etapa reuniu os estímulos e instruções para avaliação da inteligibilidade das sílabas plosivas surdas e sonoras utilizadas nas extrações dos valores de VOT na etapa de análise acústica, com roteiro para preenchimento apresentado no anexo 9. As emissões das sílabas [pa], [pɛ], [bɛ], [tar], [da], [ka] e [ga] de três leituras foram destacadas dos respectivos vocábulos.

Procedimentos de edição com aplicação de filtros para atenuar os extremos iniciais e finais dos sinais acústicos dos estímulos propiciaram o devido tratamento das amostras, no sentido de evitar cortes abruptos que interferissem negativamente na percepção dos sons. Tal etapa de seleção e edição das sílabas foi realizada no LIAAC (PUC/SP) com o programa *Cool Edit* 2000 *Syntrillium Software Corporation* instalado no mesmo computador anteriormente descrito e complementada com o programa *Sound Forge* do Laboratório de Rádio da COMFIL (PUC/SP), também descrito anteriormente, para revisão das edições e mixagem. De um conjunto de sete sílabas para cada leitura efetuada por cinco falantes, atingimos o total de 105 sílabas na edição em questão.

Para a avaliação de inteligibilidade de sílabas foram selecionados três grupos de ouvintes, compostos por alunos de cursos de pós-graduação *lato sensu* nas especialidades Voz, Motricidade Oral e Audiologia Clínica do CEFAC nas cidades de Goiânia, São José dos Campos e Campinas.

Foi realizada uma sessão de coleta de dados com cada grupo eleito, totalizando quatro salas de aula, nas quais cada juiz recebeu o roteiro apresentado no anexo 9, sendo solicitado a preenchê-lo de acordo com as instruções contidas na gravação e restritas a uma etapa de análise. Foi utilizado o ambiente de sala de aula, em momento de silêncio em seu interior e nas redondezas para reprodução da gravação em CD com duração total de 12 minutos.

Os juízes foram solicitados a identificar as sílabas apresentadas em 15 seqüências de sete estímulos, três para cada falante. O formulário para preenchimento apresentava espaços que sinalizavam e facilitavam o controle e a identificação das amostras apresentadas. Ao final da apresentação de cada seqüência ocorria a repetição da mesma. Os examinadores foram solicitados, quando necessário, a identificar a vogal com marcação de acento agudo (´) ou circunflexo (^) para sinalizarem as modalidades de vogais abertas, referentes a [ɛ] [ɔ], e fechadas, referentes a [e] [o].

Dessa forma, foi possível garantir a adequada identificação de estímulo, com resposta mais rápida em relação à possível adoção de transcrição fonética, especialmente quando levamos em conta que não há padronização quanto ao uso da simbologia de transcrição fonética na Fonoaudiologia em nosso meio. Novamente, as emissões dos falantes avaliados foram apresentadas de forma aleatória ao longo das quinze seqüências, bem como as sílabas em cada uma delas.

Os juízes solicitados a participar da tarefa puderam optar pela entrega do instrumento de avaliação devidamente preenchido, de maneira que de um total de 80 roteiros distribuídos, 78 foram devolvidos ao longo de quatro sessões de avaliação. Os resultados referentes ao número de acertos para cada sílaba e cada falante foram computados e submetidos a tratamento estatístico por meio de teste de comparação de médias com margem de erro estimada em 5%.

## 6.3. Avaliação fisiológica

Nesta etapa de apresentação de métodos são expostos os procedimentos relativos à análise dos dados de origem fisiológica, representados neste trabalho pelos procedimentos de eletroglotografia e avaliação laringológica.

## 6.3.1. Eletroglotografia (EGG)

As ondas Lx foram analisadas de forma simultânea àquelas de natureza acústica, por meio do processamento pelo programa *Cool Edit 2000 Syntrillium Software Corporation* instalado no computador microtec, *Windows* 95, 64 *megabytes* de memória do LIAAC (PUC/SP). Os procedimentos de análise seguiram a mesma rotina de inspeção de toda a seqüência de estímulos registrada, investigandose a variação dos traçados em função do estímulo de fala.

A primeira avaliação da onda Lx envolveu proposta de classificação semi-automática (Vieira, 1997), como etapa prévia à tentativa de extração de medidas de f0, perturbação, coeficiente de fechamento e índice de velocidade. A avaliação constou de atribuição de uma nota à onda, variando em escala de 1 a 4, na qual 1 representa onda impossível de análise numérica e altamente sugestiva de inexistência de contato de pregas vocais; 2 indica sinal contendo excessivo ruído e fraca amplitude, portanto pouco recomendadas para fins de extração de valores automáticos. Considerou-se sinal fraco aquele abaixo de 14000 amostras e forte o que continha acima de 2800, de acordo com referência usada para medida de amplitude no software Cool Edit. Na modalidade de avaliação adotada, 3 representa onda de amplitude aumentada, porém com ruído visível e 4 para aquelas que não apresentam problemas à inspeção. Apenas as duas últimas seriam recomendadas para o processamento e extração de medidas com extração automática.

Após análise geral das emissões, optamos pela análise detalhada das amostras da vogal em padrão entrecortado [aaa] e na emissão "a arara", uma vez que seus registros revelaram qualidade superior, em termos de traçados para todos os indivíduos estudados, quando comparadas às emissões sustentadas.

Para tais estímulos foram adotados procedimentos para identificar, caracterizar e especificar a forma e a geometria da onda e irregularidades, tais como: alongamento, deslocamento e achatamento do pico, onda triangular, abaulamento do rodapé, joelho, além de indícios de escape de ar prévio à sonorização, aperiodicidade e séries de f0. As medidas extraídas referiram-se a:

f0= computo da duração de um ciclo vibratório (T0), em amostras na figura 5
 (tela inferior) de pregas vocais e aplicação na fórmula f0=freqüência de

- amostragem/duração do ciclo (expressa em Hz);
- CC coeficiente de contato: calculado a partir da razão entre a duração (em amostras) da fase fechada (Tf) na figura 5 (tela superior) e a total do ciclo (T0) na figura 5 (tela inferior), sendo equivalente a Tf/T0 x 100 e expressa em %.

A extração das medidas de *VOT* por meio dos procedimentos eletroglotográficos centrou-se na verificação do início de registro de atividade glótica em relação à soltura de oclusão, conforme ilustrado na figura 4, e serviu de contraponto às medidas de *VOT* extraídas por meio de análise acústica, visto que as mesmas sílabas serviram de amostras para ambos os procedimentos.

As emissões em padrões grave/agudo e forte/fraco serviram de apoio para a identificação de possíveis mudanças de registro das ondas em função da mudança de estímulo.

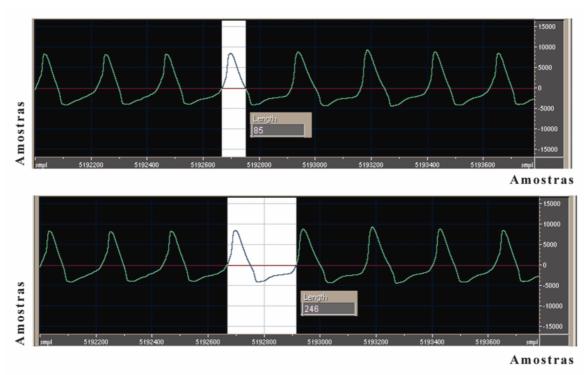


Figura 5. Procedimento de extração de medidas f0 e coeficiente de contato (CC) da onda Lx a partir das estimativas de duração da fase fechada (tela superior) e da duração total do ciclo vibratório (tela inferior).

Para extração das medidas automáticas, as amostras, previamente submetidas ao processo de tratamento anteriormente descrito, foram processadas pelo

programa egg.exe<sup>2</sup>, o qual gera arquivo de f0 em função do tempo, baseado em marcador de períodos glóticos, favorecendo o cálculo automático de valores relativos a:

-f0

-perturbação em freqüência- jitter (%)

-perturbação em amplitude – shimmer RMS (%)

-coeficiente de contato

-índice de velocidade

Vale recordar que apenas as ondas Lx que obtiveram classificação de avaliação semiautomática maior ou igual a 3 foram submetidas ao procedimento descrito, de acordo com proposta original de Vieira (1997). Foram utilizadas apenas as ondas registradas em padrão de contato de pregas vocais e não aquelas em DEGG.

## 6.3.2. Exame otorrinolaringológico

Os laudos de exame otorrinolaringológico, mais especificamente de avaliação laringológica, foram fornecidos pelo profissional médico especialista em otorrinolaringologia, seguindo parâmetros do roteiro apresentado no anexo 7.

As imagens foram revistas após o procedimento e os trechos representativos dos padrões habituais de emissão durante os exames foram selecionados e impressos, constituindo um banco de imagens digitalizadas. O processo de impressão das imagens foi realizado no serviço de Otorrinolaringologia do Hospital Edmundo Vasconcelos, com uso dos equipamentos de vídeo cassete Sony SVO 1450, placa de vídeo *Diamond* de 32 *megabytes* e software de captura de imagens Laudo & Imagem 3.0 instalados no computador com processador Pentium III, 800 MHz.

\_

<sup>&</sup>lt;sup>2</sup> Programa integrante de série de instrumentos desenvolvidos por Vieira et al (1996a) e Vieira (1997) para estimativas relativas a f0, taxas de perturbação e demais índices de eletroglotografia, disponibilizados para uso no LIAAC- PUC/SP, mediante autorização dos autores.

# 6.4. Experimento voltado à investigação dos efeitos da composição do estímulo de fala na análise acústica de longo termo

A atenção a este aspecto emergiu de referências conflitantes na literatura quanto à consideração de interferência negativa de períodos não vozeados, fricativos e pausas nas emissões processadas por análises de longo termo (Frøkjær-Jensen, Prytz, 1976; Hammarberg et al, 1980; Dejonckere, Cuvelier, 1983; Dejonckere, Villarosa, 1986; Löfqvist, Mandersson, 1987; Hammarberg et al, 1986; Hammarberg, Gauffin, 1995; Mendonza et al 1996; Linville, Rens, 2001), em detrimento de outros, que utilizaram estímulos (textos) de composição mista em relação aos sons surdos e sonoros e não adotaram a exclusão dos trechos não vozeados (Dejonckere, 1983; Dejonckere, 1984; Wendler et al, 1986; Pittam, 1987; Figueiredo, 1993; Cleveland, Sundberg, 2001). Pittam (1987) não excluiu sons surdos, porém compôs amostra sem sons nasais. Wendler et al (1986) demonstraram que não houve diferença nos registros de longo termo em função de composição mista e exclusão dos trechos não vozeados num banco de dados de 473 vozes, incluindo indivíduos com e sem disfonia. O último estudo salienta a importância de tratamento dos dados do ELT, cuja garantia propiciará plano perceptivo, inclusive em associações ao casos de disfonias. suas Ananthapadmanabha (1995) também não encontrou diferenças em termos de composição de amostra mista ou exclusiva por sons sonoros.

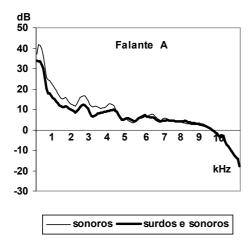
Como forma de investigar o possível impacto da amostra coletada nos registros de longo termo no presente estudo, de composição mista em termos de sons surdos e sonoros, foram elaborados dois textos distintos, de forma que um apresentava composição exclusiva de sons vozeados e outro mista com sons vozeados e não vozeados, com conteúdo e número de palavras similares, inspirados no original utilizado na coleta de dados, os quais encontram-se no anexo 10.

Foram efetuadas gravações de leitura dos dois textos acima referidos por dois falantes sem alterações vocais ou de aparelho fonador, ambos do sexo feminino, pesquisadores do LIAAC, na faixa etária de 35 a 52 anos de idade, sem antecedentes de alterações vocais, denominados falantes A e B. As gravações seguiram o mesmo padrão exposto para a coleta de dados anteriormente descrita, com

a ressalva de que foi realizada apenas gravação do sinal acústico, no mesmo ambiente do laboratório de áudio da COMFIL (PUC/SP), com o mesmo equipamento e concepção de coleta das amostras de fala anteriormente descritos.

Tais estímulos foram digitalizadas em freqüência de amostragem de 22050 Hz, quantizadas em 16 bits, no programa CSRE 45 (*Computer Speech Reserch Environment* da *Avaaz Inovations*) em formato .adf e, finalmente, convertidas em amostras com extensão .wav, as quais foram analisadas no programa *Multispeech* (*Kay Elemetrics Corp.*). Ambos programas integram os recursos do LIAAC (PUC-SP). Os parâmetros de análise sofreram ajustes relativos à aplicação e não aplicação de pré-ênfase (0,98). O número de pontos utilizado nas várias etapas permaneceu em 256, com janelamento *Hanning* e limite superior de freqüência em 11000 Hz. Diante da mesma abordagem exposta anteriormente para análise acústica de longo termo, procuramos comparar as alterações de traçado e medidas correspondentes, enquanto variação de amplitude ao longo de 128 pontos de freqüência, com relação às amostras compostas exclusivamente por sons sonoros e por sons surdos e sonoros. Os resultados também foram apreciados em função de variações em termos da aplicação de pré-ênfase e da remoção de trechos de pausas silenciosas para cada texto.

Na figura 6 são expostos os gráficos comparativos dos espectros de longo termo para texto composto exclusivamente por sons sonoros (linha pontilhada), em relação àquele de composição mista (linha contínua) para os falantes A e B. A observação dos traçados revelou basicamente o mesmo contorno, com diferença apenas em termos da amplitude registrada, a qual apresenta-se mais intensa para aquele com predomínio de sons vozeados, especialmente nas faixas de freqüência abaixo de 4300 Hz, incluindo o primeiro pico do espectro. Deve-se recordar que para ambos houve o uso de pré-ênfase (0,98), portanto os dados comparativos para análise de dois textos distintos envolvem os mesmos falantes e os mesmos parâmetros de análise. Poucas mudanças foram evidenciadas no traçado acima de 4 kHz.



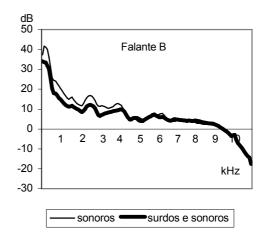


Figura 6. Espectros de longo termo para duas leituras com estímulos compostos somente com sons sonoros (linha pontilhada) e com composição mista de sons surdos e sonoros (linha contínua) para os falantes A e B, participantes da etapa de teste dos efeitos do contexto nos aspectos de longo termo da emissão.

Os testes também foram realizados com observação dos traçados e cômputo de variação por faixas de freqüência, em que constaram as amplitudes dos primeiros picos e o declínio entre ambos, não sendo detectadas diferenças. Dessa forma, a inclinação espectral não se alterou, sugerindo-nos que, diante da manutenção do mesmo estímulo e das mesmas condições de análise, a comparação entre os indivíduos do grupo estudado é válida.

Outro achado que merece destaque é fato da remoção de pausas ter levado ao aumento do máximo de amplitude detectado, para ambos textos, levando a certo aumento de amplitude na região do pico espectral, de forma que a exclusão de pausas poderia ser adotada como procedimento de edição das amostras. Nesse caso, novamente não foram registradas diferenças acima da faixa de 4000 Hz. Tal referência reveste-se de importância ao retomarmos as constantes referências a interferências de sons fricativos promovendo aumento do nível de energia espectral entre 3 e 4 kHz (Frøkjær-Jensen, Prytz, 1976; Hammarberg et al, 1980, 1984; Löfqvist, 1986; Hammarberg, Gauffin, 1995; Mendonza et al, 1996).

Os valores referentes à variação de amplitude ao longo de 128 pontos de freqüência, decorrente da análise numérica dos pontos do traçado de longo termo compreendidos na faixa de 0 a 11 kHz também foram submetidos a tratamento estatístico por meio de análise fatorial de componentes principais e análise de construção de clusters com auxílio dos programas SPSS 10.0 for windows e SPAD 3.5

do Laboratório de Informática da Faculdade de Psicologia da PUC/SP, sendo o mesmo procedimento aplicado nas amostras do grupo estudado e indivíduo referência.

A comparação entre as gravações por meio da análise fatorial e de formação de *clusters* demonstrou não haver diferença significativa por influência da modalidade do texto ou por introdução de pré-ênfase.

Em suma, a composição exclusiva de texto por sons sonoros não impôs alterações significativas nos traçados de ELT quando comparadas àqueles obtidos por amostras com composição mista, especialmente no que se refere à inclinação espectral. Apesar de certo rebaixamento do máximo de amplitude registrada, a principal informação espectral impressa no declínio manteve-se inalterada, guardando proporção entre maior e menor registro de energia, sem diferenças acima de 4300 Hz. Tais achados foram compatíveis com aqueles apresentados por Wendler et al (1986) e Ananthapadmanabha (1995).

Dessa maneira, optamos por manter o texto apresentado com o parâmetro de análise de longo termo para a população em questão, com o cuidado de remoção de pausas marcadas por silêncio na emissão.

Vale finalmente ressaltar que as emissões com textos de composição mista são inclusive representativos da fala espontânea, na qual há alternância entre sons surdos e sonoros. Além disso, mesmo o texto composto exclusivamente pelos sons sonoros, gera situações de final de enunciado, vogais em posição pós-tônica e pausas que também incluem períodos de desvozeamento na emissão. Portanto, o uso de texto com predomínio de sons vozeados (sonoros) não elimina a possibilidade de trechos de desvozeamento, para os quais alguns autores sugerem procedimentos automáticos de eliminação (Hammarberg et al, 1980; Dejonckere, Cuvelier, 1983; Löfqvist, Mandersson, 1987; Hammarberg, Gauffin, 1995; Linville, Rens, 2001).

No enfoque do presente estudo, adotamos o procedimento de comparação entre os integrantes do grupo estudado, de forma que o cuidado fundamental consistiu em manter os procedimentos de edição e tratamento das amostras nas várias etapas, como forma de imprimir confiabilidade ao banco de dados.

## 7. ANÁLISE DOS DADOS DAS ESFERAS ACÚSTICA, PERCEPTIVA E FISIOLÓGICA EM UM GRUPO DE INDIVÍDUOS DISFÔNICOS

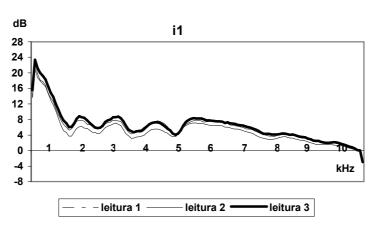
A exposição dos resultados está estruturada em dados referentes à análise das amostras de fala e do exame laringológico. As amostras de fala foram avaliadas do ponto de vista acústico, perceptivo-auditivo e fisiológico (eletroglotográfico) e complementadas pelos dados do exame otorrinolaringológico, com ênfase nos achados telelaringoestroboscópicos e videoquimográficos. Tais esferas de descrição representam os itens da seqüência do texto, enquanto apresentação dos resultados.

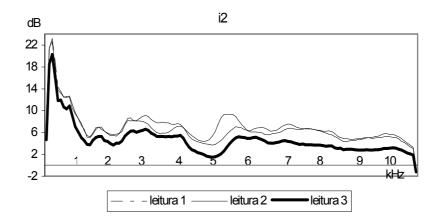
#### 7.1. Análise acústica

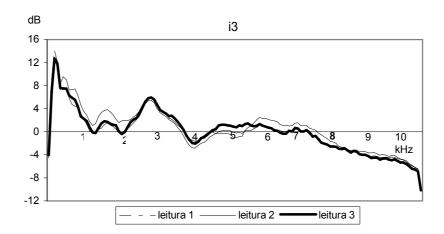
Os resultados da inspeção acústica da amostras de fala são expostos enquanto abordagens de longo e de curto termo, respectivamente representadas pelos itens 7.1.1 e 7.1.2 desta seção.

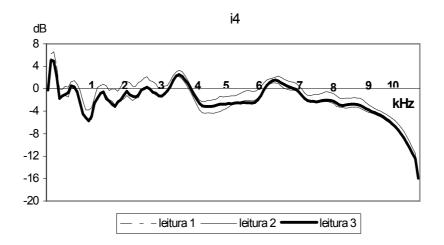
## 7.1.1. Dimensão acústica de longo termo

Os traçados referentes aos espectros de longo termo representam a distribuição de 128 pontos de freqüência analisados e são apresentados em três traçados referentes às leituras do mesmo texto pelos indivíduos do grupo estudado (i1 a i4) e referência (iR) na figura 7.









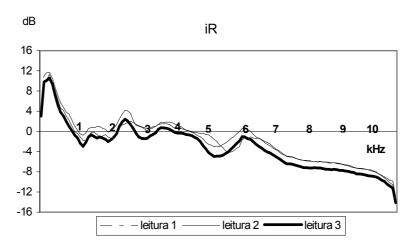


Figura 7. Espectros de longo termo (ELT) de três leituras de texto realizadas por i1, i2, i3, i4 e iR.

Os dados referentes à análise fatorial de tais informações resultou em seis classes, nas quais os falantes i1, i3, i4 e iR representaram classes distintas, no sentido de que agruparam as suas três emissões. Para i2, foram definidas duas classes. O fator subjacente a dividir as classes com 88,9 % de influência referiu-se à amplitude registrada em faixas de freqüências diferenciadas para emissões de i1, i2, i3, i4 e IR. Dessa forma, i1 e i2 apresentaram amplitude acima da média do grupo, com concentração entre 4300 a 6700 Hz para i1 e entre os intervalos de 3500 a 3900 Hz e 9700 a 10400 Hz para i2. Em contrapartida, i3, i4 e iR apresentaram-nas abaixo da média do grupo, sendo que a faixa de 4100 a 4200 Hz destacou-se para i3 e os intervalos entre 260 a 1400 Hz, 2600 a 2850 Hz e 3450 a 5000 Hz para i4. No caso de iR predominaram os pontos situados entre 6600 a 8500 e 8700 a 9900 Hz.

A análise fatorial também realizada com a exclusão das emissões de iR resultou em três classes com 91,03 % de influência, em que i1 e i2 tiveram suas emissões agrupadas, enquanto i3 e i4 permaneceram como classes distintas.

As médias de concentração de energia em faixas de freqüências calculadas para os intervalos de freqüência 0 a 1 kHz, 1 a 3, 3 a 5 kHz e 5 a 11 kHz são apresentadas na figura 8.

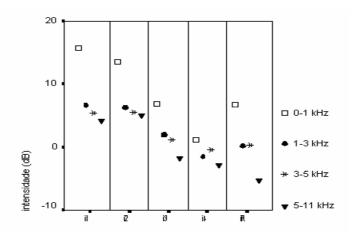


Figura 8. Distribuição das médias de concentração de energia (em dB) por faixas de freqüência (0-1 kHz, 1-3, 3-5 kHz e 5-11 kHz) dos espectros de longo termo para três leituras efetuadas pelos indivíduos do grupo estudado (i1 a i4) e referência (iR).

A análise fatorial aplicada ao cômputo das médias anteriormente representadas revelou o agrupamento das três emissões distintas de i3, i4 e iR, enquanto i1 e i2 mesclaram suas emissões por dois grupos. O fator subjacente a dividir as classes com 95,10 % de influência referiu-se à maior amplitude das faixas para i1 e i2 em relação ao restante dos falantes. A diminuição de amplitude revelada por i4 estendeu-se de 3-11 kHz. Apenas alguns pontos em torno de 1-3 kHz representaram respectivamente a diminuição e o aumento de amplitude para as emissões de iR, enquanto i3 não definiu faixas distintas perante o restante do grupo.

#### 7.1.2. Dimensão acústica de curto termo

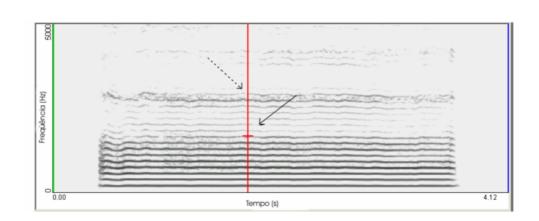
Os resultados apresentados nesta etapa referem-se àqueles das análises de espectrografia de banda estreita e de banda larga, medidas de tempo de início de sonorização (VOT), análises da estrutura harmônica e formântica por meio de análises cepstral e espectral por procedimentos FFT e LPC.

## 7.1.2.1. Espectrografia de banda estreita

Os espectrogramas de banda estreita revelaram particularidades das emissões dos falantes avaliados quanto à presença e distribuição de harmônicos e não harmônicos. Os achados para cada indivíduo do grupo estudado (i1 a i4) e para o indivíduo referência (iR) são apresentados separadamente com relação às emissões de [a:] e [aaa].

Registro de componentes harmônicos restrito a nove componentes (H9) no limite máximo de 1600 Hz, sinalizado na figura 9a por seta com linha contínua. Abaixo deste limite de freqüência os componentes apresentam nível de energia praticamente constante, traduzido em escassas variações no grau de escurecimento das barras horizontais. Acima deste limite, registra-se esboço de harmônicos até 3000 Hz, limite apontado Figura 3a com uma seta traçada em linha pontilhada. Tal característica volta a ser detectada em maior concentração entre 5000 e 8800 Hz, não havendo registro de energia acima de tais limites, como apontado por setas na 9b.





Etequencia (Hz)

Figura 9. Espectrogramas de banda estreita com limites superiores de freqüência estabelecidos em 5000 Hz (a) e 11000 Hz (b) da mesma emissão da vogal [a:] por i1, digitalizada na freqüência de amostragem de 22050 Hz.

## [aaa]

O registro de componentes harmônicos amplia-se para a quantidade máxima de 17 componentes (H17), com perda gradativa de definição até a freqüência de 3900 Hz (figura 10). O ruído de fundo de baixa intensidade pode ser observado entre H3 e H7 e ruído difuso concentra-se em bandas em regiões mais altas até 8500 Hz (Figura 10). Registros eventuais de ataque vocal brusco foram revelados por irregularidade no trecho inicial de algumas vogais, com presença de pequenas bifurcações, as quais não são registradas no restante da sustentação das emissões.

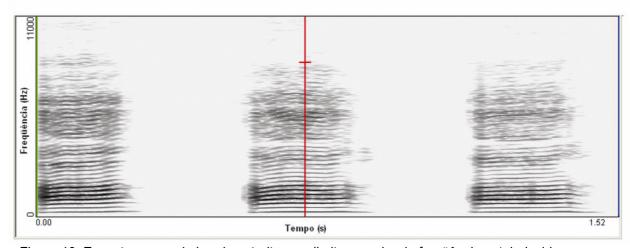


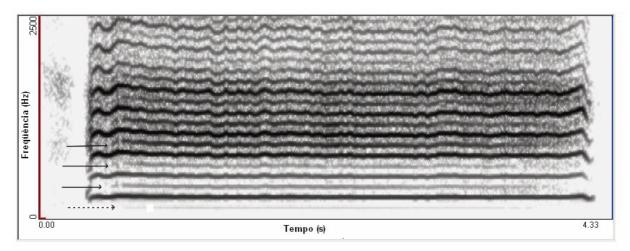
Figura 10. Espectrograma de banda estreita com limite superior de freqüência estabelecido em 11000 Hz da emissão [aaa] por i1, digitalizada na freqüência de amostragem de 22050 Hz.

i2

• [a:]

O registro de componentes harmônicos concentrados em faixas de freqüências mais graves revela bifurcações (subharmônicos) de forma constante em todos os estímulos avaliados, sendo alguns deles sinalizados por setas com linha contínua na figura 11a. Outro achado que merece destaque é o fato de apresentar bifurcação anteriormente ao registro de f0, de amplitude diminuída, porém constante ao longo da sustentação, indicada por seta com linha pontilhada na figura 11a. A adição de ruído pode ser registrada a partir do terceiro harmônico (H3) e bifurcações. Por volta da freqüência de 4500 Hz, os componentes não harmônicos passam a predominar, concentrados em bandas até 8800 Hz e de forma difusa até 11000 Hz (figura 11b).

(a)



(b)

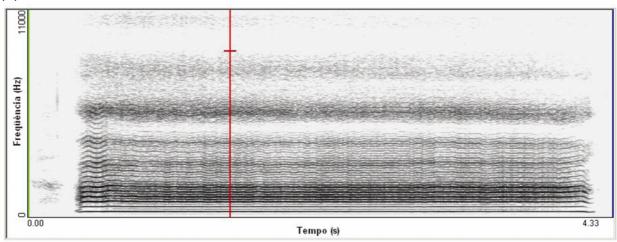
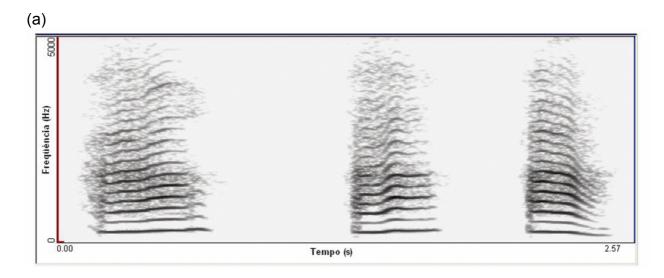


Figura 11. Espectrogramas de banda estreita com limites superiores de freqüência estabelecidos em 2500 Hz (a) e 11000 Hz (b) da mesma emissão da vogal [a:] por i2, digitalizada na freqüência de amostragem de 22050 Hz.

[aaa]

Os componentes harmônicos, registrados até a freqüência de 3300 Hz (H13), aparecem mesclados ao ruído, especialmente a partir do terceiro harmônico (H3), concentrado em bandas até 6200 Hz (figura 12a e b). Duas das emissões registraram ruído adicional difuso até o máximo de 11000 Hz, apontado por seta na figura 12b. Algumas bifurcações foram detectadas apenas em trechos iniciais de emissão, caracterizando ataque vocal brusco.



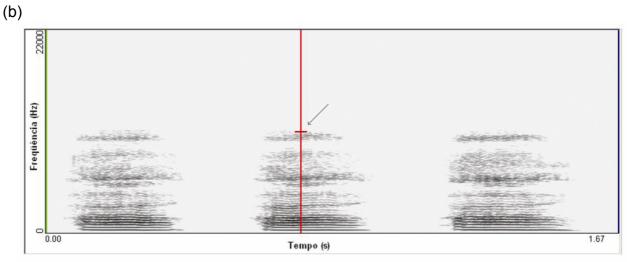


Figura 12. Espectrogramas de banda estreita com limites superiores de freqüência estabelecidos em 5000 Hz (a) e 22000 Hz (b) da emissão [aaa] por i2, digitalizada respectivamente nas freqüências de amostragem de 22050 Hz e de 44100 Hz.

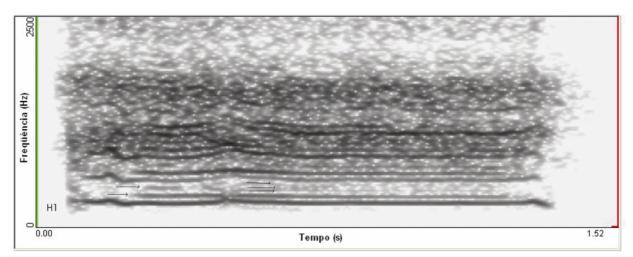
- I3 ■ [a:]

Predomínio de características não harmônicas, com traçado de harmônicos ao fundo de registro de ruído dominante (figura 13a e 13b), concentrado em bandas até o limite superior de 9000 Hz, indicado por seta na figura 13b. De forma

difusa e em menor intensidade, os componentes não harmônicos foram detectados até a freqüência de 10700 Hz.

As bifurcações aparecem de forma marcante, sendo possível registrar séries de duas a três bifurcações entre dois harmônicos esboçados em faixas de freqüência graves (figura 13a), indicadas por setas. Abaixo do traçado referente a f0, indicado na figura 13a pelo símbolo (H1), a bifurcação pode ser detectada, com registro inconstante ao longo da sustentação da emissão.

(a)



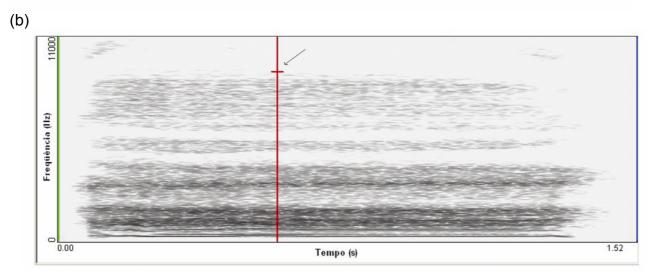
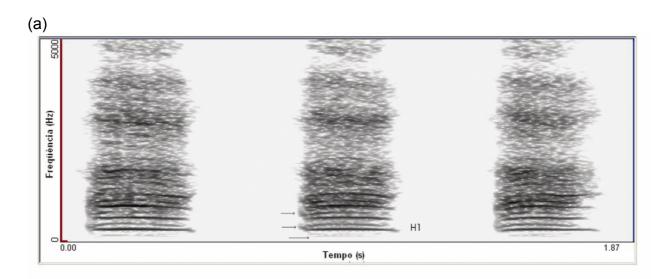


Figura 13. Espectrogramas de banda estreita com limites superiores de freqüência estabelecidos em 2500 Hz (a) e 11000 Hz (b) da mesma emissão da vogal [a:] por i3, digitalizadas respectivamente nas freqüências de amostragem de 11025 e 22050 Hz.

## [aaa]

Predominância de componentes não harmônicos, identificados a partir do traçado de H1 (figura 14a). As bifurcações aparecem de forma constante ao

longo das emissões, sendo possível registrá-las inclusive em meio ao ruído, como parece indicado por setas na figura 14a, incluindo aquela abaixo de H1. Os componentes não harmônicos concentram-se em bandas até a freqüência de 11000 Hz (figuras 14b).



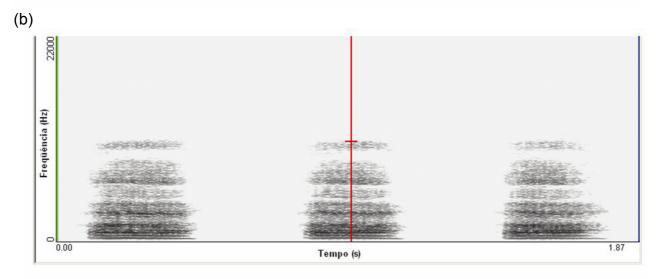


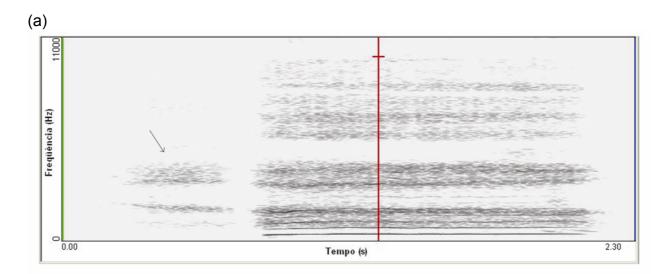
Figura 14. Espectrogramas de banda estreita com limites superiores de freqüência estabelecidos em 5000 Hz (a) e 22000 Hz (b) da emissão [aaa] por i3, digitalizada respectivamente nas freqüências de amostragem de 22050 Hz e de 44100 Hz.

- i4 ■ [a:]

Predominância de componentes não harmônicos, com concentração em bandas até a faixa de freqüência de 10000 Hz, quando aparece de forma difusa e em pouca concentração até no máximo de 11200 Hz. O registro de

harmônicos permanece restrito à quantidade máxima de seis entre as várias emissões avaliadas.

Apresentou ruído inspiratório prévio a todas as emissões vocálicas, o qual tendeu a apresentar bandas de concentração de energia semelhantes às da vogal, no que se refere às faixas de concentração correspondentes aos formantes das vogais, numa caracterização possível de fonação inspiratória, sinalizado por setas nas figuras 15a e 15b.



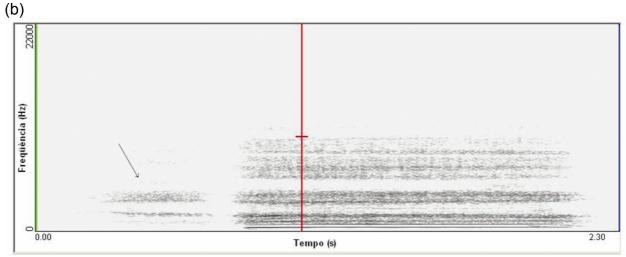
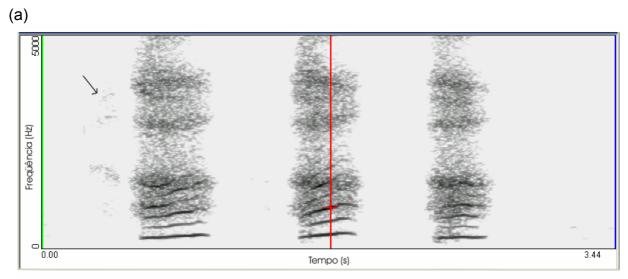


Figura 15. Espectrogramas de banda estreita com limites superiores de freqüência estabelecidos em 11000 Hz (a) e 22000 Hz (b) da mesma emissão da vogal [a:] por i4, digitalizada nas freqüências de amostragem de 22050 Hz e 44100 Hz.

## [aaa]

Os componentes não harmônicos predominam nas representações espectrográficas (figuras 16a e 16b), e podem ser detectados a partir

do traçado do terceiro harmônico (H3) na figura 16a. Sua concentração foi registrada até o limite de freqüência de 11500 Hz. Apresentou ruído inspiratório prévio a algumas emissões vocálicas, sinalizado nas figuras 16a e 16b, o qual acompanha a divisão por faixas de concentração de energia da vogal subseqüente.



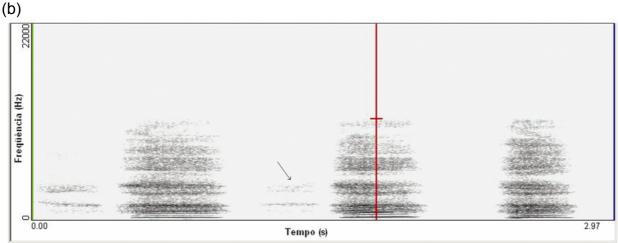
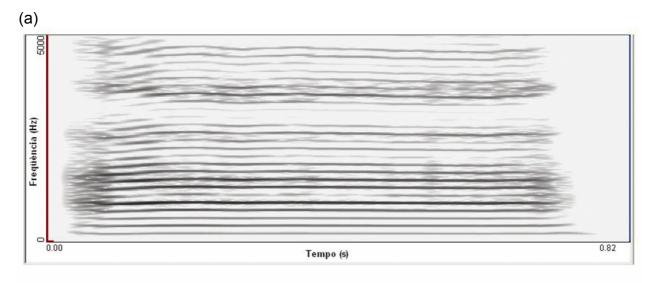


Figura 16. Espectrogramas de banda estreita com limites superiores de freqüência em 5000 Hz (a) e 22000 Hz (b) da emissão [aaa] por i4, digitalizada respectivamente nas freqüências de amostragem de 22050 Hz e de 44100 Hz.

- iR ■ [a:]

Predomínio de componentes harmônicos (figura 17a), registrados até o limite de freqüência máximo de 7500 Hz, o qual aparece indicado por seta na figura 17b. Em algumas das sustentações são observadas flutuações em freqüência dos componentes harmônicos. Há registro de ataque vocal brusco em várias vogais analisadas, caracterizado pela perturbação inicial do traçado dos harmônicos nas várias figuras apresentadas.



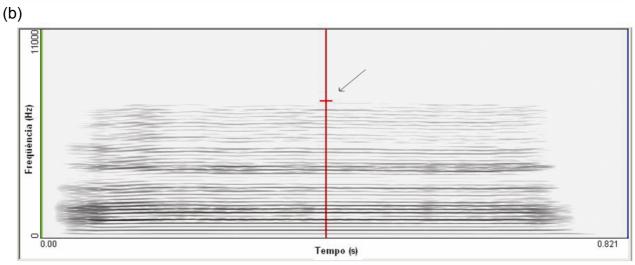


Figura 17. Espectrogramas de banda estreita com limites superiores de freqüência estabelecidos em 5000 Hz (a) e 11000 Hz (b) da mesma emissão da vogal [a:] por iR, digitalizada na freqüência de amostragem de 22050 Hz.

#### [aaa]

As representações destacam a predominância de componentes harmônicos (figura 18), detectados até o limite máximo de freqüência de 7800 Hz, indicado por seta. Observam-se alguns componentes de distribuição difusa e em reduzida concentração até, no máximo, 11000 Hz. Detectam-se algumas perturbações nos traçados dos harmônicos, especialmente nos trechos iniciais, compatíveis com ataque vocal brusco, quando aparecem bifurcações de curta duração. Tal achado pode ser observado na figura apresentada, porém não foi constante para todas as seqüências analisadas.

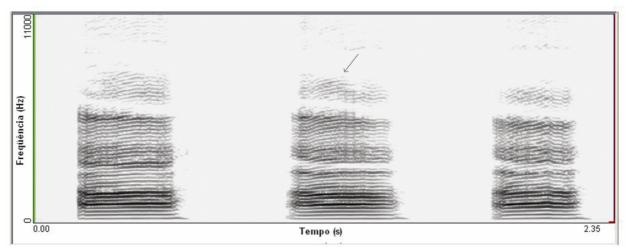


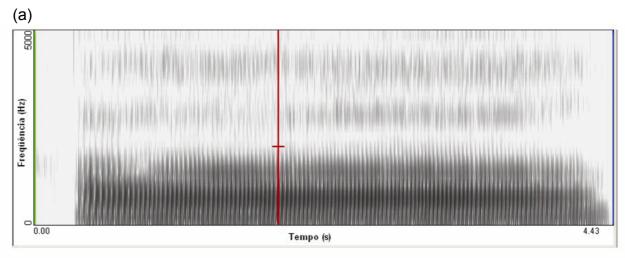
Figura 18. Espectrograma de banda estreita com limite superior de freqüência em 11000 Hz da emissão [aaa] por iR, digitalizada na freqüência de amostragem de 22050 Hz.

## 7.1.2.2. Espectrografia de banda larga

Os espectrogramas de banda larga revelaram as características das emissões avaliadas pela distribuição da energia acústica por faixas de freqüência ao longo da duração do sinal, revelando aspectos de sua periodicidade. Os resultados são apresentados separadamente para cada indivíduo do grupo estudado (i1 a i4) e indivíduo referência (iR) em função da emissões de [a:], [aaa] e "a arara".

-i1 ■ [a:]

Predomínio de registro de periodicidade em porção inferior do gráfico, até o intervalo de freqüências compreendido entre 2000 e 3150 Hz, (figuras 19a e 19b), quando progressivamente inicia a adição de ruído, concentrado em bandas até no máximo 10800 Hz (Figura 19b).



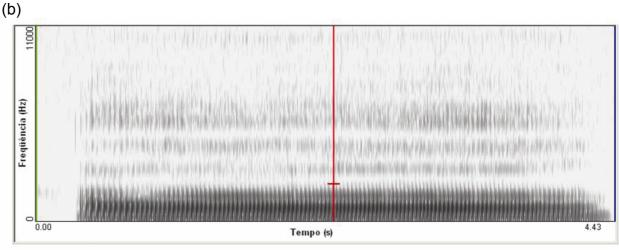


Figura 19. Espectrogramas de banda larga com limites superiores de freqüência estabelecidos em 5000 Hz (a) e 11000 Hz (b) da mesma emissão da vogal [a:] por i1, digitalizada na freqüência de amostragem de 22050 Hz.

## [aaa]

Periodicidade registrada em maior concentração até 4000 Hz, quando o sinal apresenta maiores irregularidades, estendendo-se até 9000 Hz (figura 20). Presença de ataques vocais bruscos em todas as seqüências analisadas, sinalizado por seta nas figura indicada. Em duas emissões aparecem áreas sugestivas de anti-ressonâncias (não constantes ao longo da sustentação).

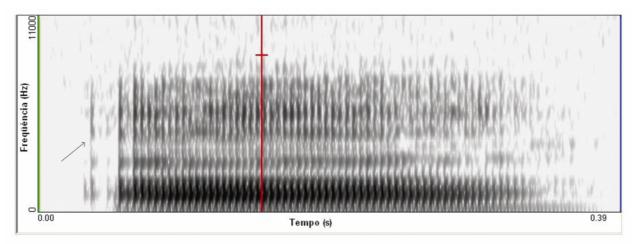
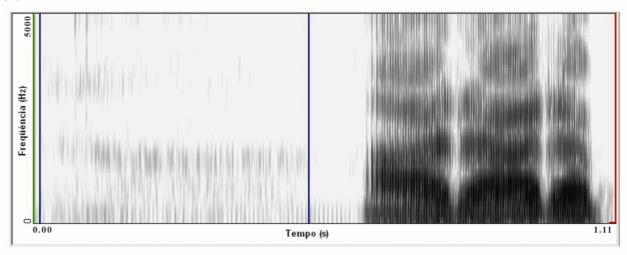


Figura 20. Espectrograma de banda larga com limite superior de freqüência estabelecido em 11000 Hz da mesma emissão da vogal central de [aaa] por i1, digitalizada na freqüência de amostragem de 22050 Hz.

#### Trechos de emissão encadeada "a arara"

Característica harmônica chega a ser registrada até 9000 Hz de forma concentrada, atingindo em seu padrão difuso o limite de 11100 Hz (figura 21b). Foram registrados indícios de fonação inspiratória de longa duração, chegando a superar aquela do trecho de fala analisado, delimitados por cursores no intervalo de 0,51 segundos, numa emissão com duração de total de 1,11 segundos (figura 21b).

(a)



(b)

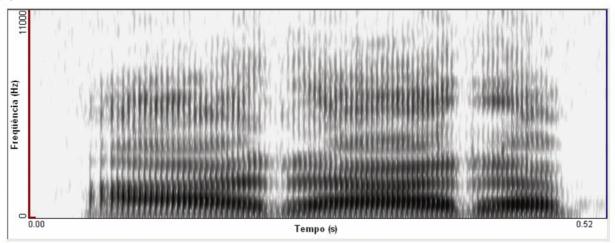


Figura 21. Espectrogramas de banda larga com limites superiores de freqüência estabelecidos em 5000 (a) Hz e 11000 Hz (b) da mesma emissão de "a arara" por i1, digitalizada na freqüência de amostragem de 22050 Hz.

-i2

\_

• [a:]

Ruído inspiratório prévio a todas as emissões analisadas, o qual pode ser observado no início da representação gráfica na figura 22. Além disso, foram registrados indícios de escape de ar não sonorizado ao início e final de algumas das emissões analisadas, com concentração especialmente nas freqüências de 1200 a 1500 Hz e eventualmente em 5500 Hz. Os traçados relativos à dimensão de periodicidade encontram-se restritos à faixa de freqüências inferiores a 2000 Hz, conforme indica a seta na figura 22.

O trecho inicial de emissão revela periodicidade para faixas mais altas de freqüência, além daquelas expostas acima. A energia aperiódica segue concentrada em bandas até 11000 Hz e ruído difuso pode ser detectado em freqüências superiores a esta faixa (figura 22).

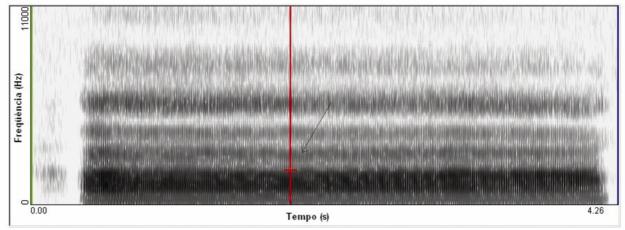


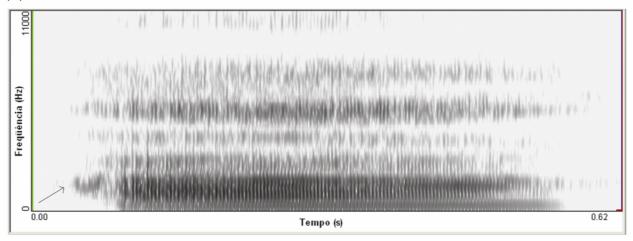
Figura 22. Espectrograma de banda larga com limite superior de freqüência estabelecido em 11000 Hz (b) da emissão da vogal [a:] por i2 digitalizada na freqüência de amostragem de 22050 Hz.

#### [aaa]

Registro de aperiodicidade predominante, com concentração em bandas de freqüência. Alguns indícios de periodicidade encontram-se situados na faixa inferior a 2100 Hz (figura 23a). A partir de tal ponto definem-se duas bandas de ruído, a primeira até 7000 Hz (figura 23a) e, a segunda até a freqüência máxima de 11600 Hz.

Registro de ar não sonorizado ao início e/ou final das emissões também foi constante e aparece indicado por setas com traçado pontilhado em ambas as figuras referentes aos espectrogramas, denotando concentração de energia em 1500 Hz, como revela envelope espectral do referido trecho, apresentado na figura 23b. Ataques vocais bruscos também foram registrados em algumas das vogais analisadas.

(a)



(b)

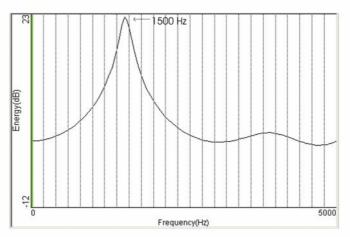


Figura 23. Espectrograma de banda larga (a) com limite superior de freqüência estabelecido em 11000 Hz e envelope espectral (b) da mesma emissão da vogal central de [aaa] por i2, digitalizada na freqüência de amostragem de 22050 Hz.

#### Trechos de emissão encadeada "a arara"

As representações revelam componentes de aperiodicidade (figuras 24a e 24b) concentrados até a freqüência de 11800 Hz, conforme indicação de seta em figura 24b. Tendência geral em demarcar uma faixa de freqüências inferior a 4500 Hz com menos ruído, sinalizada na figura 24a, e uma superior no qual predomina energia aperiódica. Ruído inspiratório prévio de longa duração à emissão também foi detectado, com concentração de energia semelhante à vogal, novamente numa caracterização possível de fonação inspiratória. As faixas referentes a F1 e F2 revelam algumas descontinuidades em seus traçados, como revelado na figura 24a. Os trechos que apresentam ar não sonorizado revelam concentrações em torno de 1500 e 5500 Hz.

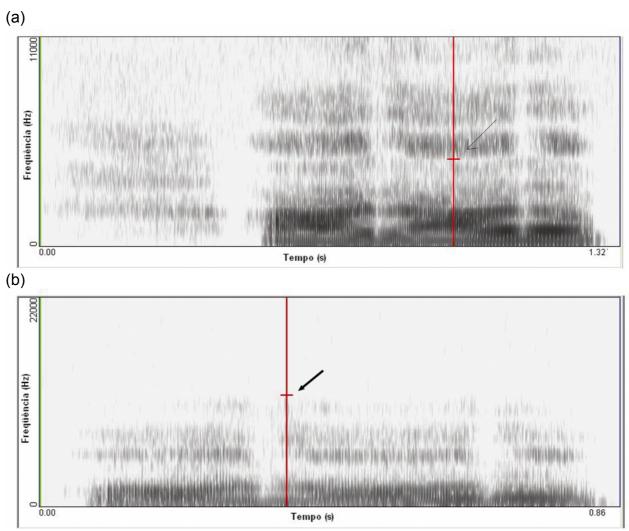
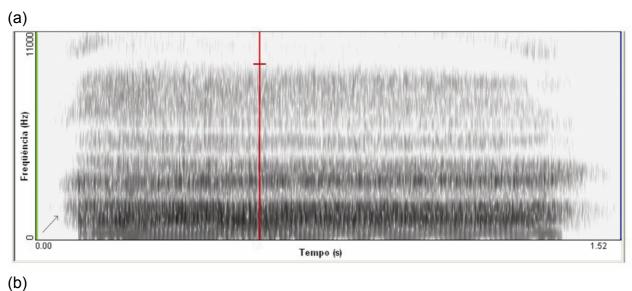


Figura 24. Espectrogramas de banda larga com limites superiores de freqüência estabelecidos em 11000 (a) Hz e 22000 Hz (b) da mesma emissão de "a arara" por i2, digitalizada nas freqüências de amostragem de 22050 Hz e 44100 Hz.

- i3 ■ [a:]

Registro predominante de aperiodicidade (figura 25a e 25b), concentrada em bandas até 9400 Hz, e de forma difusa até 11800 Hz (figura 25b). Certo registro de regularidade do traçado restringe-se à faixa inferior a 600 Hz. Indícios de escape de ar não sonorizado em início e/ou final de emissão, apontada por setas no início das figuras 25a e 25b. Há rupturas no traçado de F1 e F2.



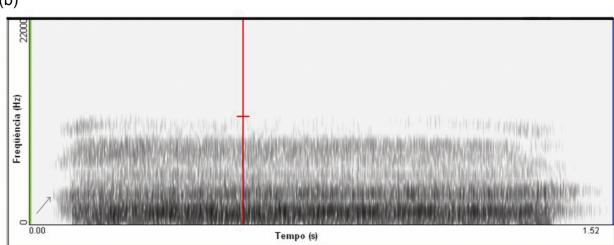


Figura 25. Espectrogramas de banda larga com limites superiores de freqüência estabelecidos em 11000 Hz (a) e 22000 Hz (b) da mesma emissão da vogal [a:] por i3, digitalizada nas freqüências de amostragem de 22050 Hz e 44100 Hz.

#### [aaa]

Registro predominante de aperiodicidade nos espectrogramas, com concentração em bandas até 12100 Hz (figuras 26a e 26b). Maior regularidade do traçado restringe-se à faixa inicial de 1000 Hz da representação, de forma inconstante ao longo das emissões. Ataques vocais bruscos foram demarcados em todas as emissões. Indícios de escape de ar não sonorizado, especialmente ao final das emissões, conforme assinalado por seta na figura 26a, com concentração em freqüências de 1600 e 3000 Hz.

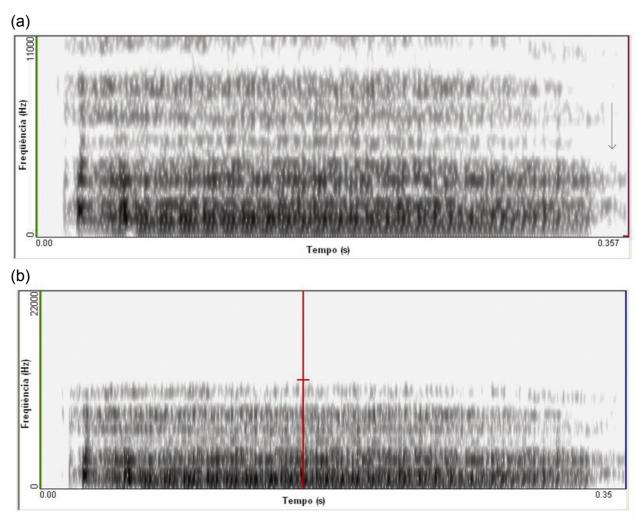


Figura 26. Espectrogramas de banda larga com limites superiores de freqüência estabelecidos em 11000 (a) Hz e 22000 Hz (b) da mesma emissão da vogal central de [aaa] por i3, digitalizada nas freqüências de amostragem de 22050 Hz e 44100 Hz.

#### Trechos de emissão encadeada "a arara"

Os espectrogramas analisados revelam predomínio de ruído (figuras 27a e 27b), atingindo os registros superiores máximos de 11500 Hz, sinalizado por seta com traçado contínuo na figura 27b. Ataques vocais bruscos consistentes ao início das seqüências. Maior irregularidade do traçado do sinal da vogal em posição pós-tônica foi detectado (figura 27a). Indícios de ar não sonorizado registrado nos trechos iniciais da emissão aparecem indicados por setas com traçado pontilhado em ambas as figuras, novamente com concentrações de amplitude em 1600 e 3000 Hz.

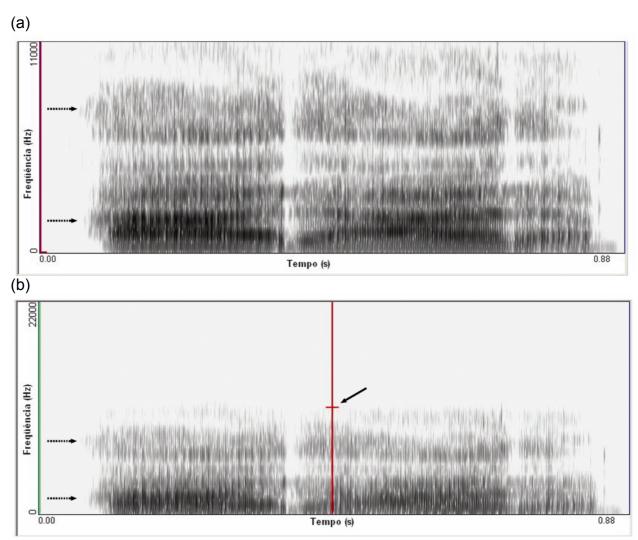


Figura 27. Espectrogramas de banda larga com limites superiores de freqüência estabelecidos em 11000 (a) Hz e 22000 Hz (b) da mesma emissão de "a arara" por i3, digitalizada nas freqüências de amostragem de 22050 Hz e 44100 Hz.

-i4 ■ [a:]

Predomínio de aperiodicidade (figura 28a e 28b), com concentração em bandas até a freqüência de 11600 Hz (figura 28b). O registro de certa regularidade do traçado restringe-se à faixa inferior a 700 Hz. Achado de escape de ar não sonorizado em início e/ou final de emissão aparece indicado por setas na figura 28a. A representação de fonação inspiratória revela ruído prévio à emissão, cujas concentrações assemelham-se dos formantes da vogal subseqüente, assinalado na mesma figura.

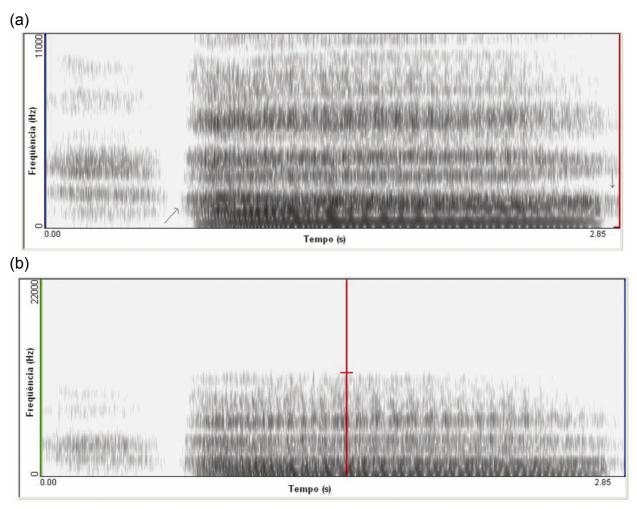
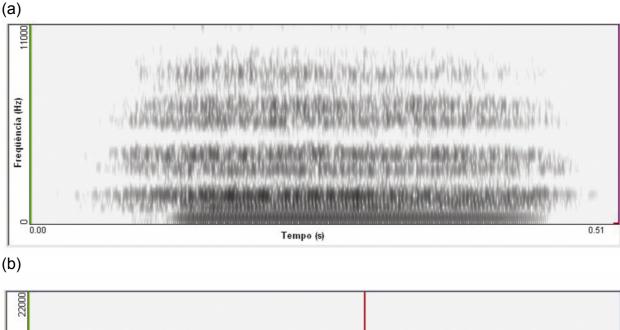


Figura 28. Espectrogramas de banda larga com limites superiores de freqüência estabelecidos em 11000 Hz (a) e 22000 Hz (b) da mesma emissão da vogal [a:] por i4, digitalizada nas freqüências de amostragem de 22050 Hz e 44100 Hz.

## [aaa]

A característica dominante de aperiodicidade do sinal revela concentrações em bandas na extensão máxima de freqüência de 11500 Hz (Figura 29a e 29b). Certa regularidade do traçado restringe-se à faixa inferior a 1300 Hz dentre as várias emissões estudadas. Há indícios de fonação inspiratória ao início registro da emissão entrecortada, guardando a particularidade de apresentar a distribuição por faixa de freqüências similar ao trecho sonorizado da emissão da vogal seguinte. Nesses trechos da emissão, destacam-se especialmente as concentrações em torno das freqüências de 1500 e 3000 Hz.



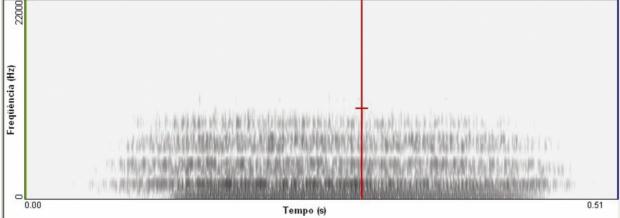


Figura 29. Espectrogramas de banda larga com limites superiores de freqüência estabelecidos em 11000 (a) Hz e 22000 Hz (b) da mesma emissão da vogal central de [aaa] por i4, digitalizada nas freqüências de amostragem de 22050 Hz e 44100 Hz.

Trechos de emissão encadeada "a arara"

Predomínio de ruído em toda a extensão da representação espectrográfica (figura 30), o qual atinge limites máximos de 11200 Hz.

Apresentou de forma constante ruído inspiratório prévio a todas as emissões, o qual foi registrado anteriormente ao artigo, quando de maior duração, e antes da emissão "arara", quando de menor duração. Em ambas as situações, apresenta bandas de concentração de energia, que guardam semelhança com o traçado dos formantes das vogais subseqüentes. O primeiro intervalo, estimado em 0,68 segundos, e o segundo, em 0,20 segundos de emissão são representados na figura 30.

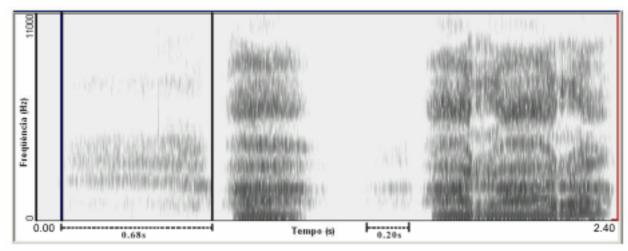


Figura 30. Espectrograma de banda larga com limite superior de freqüência estabelecido em 11000 (a) Hz da mesma emissão de "a arara" por i4, digitalizada na freqüência de amostragem de 22050 Hz.

- iR

[a:]

Os espectrogramas revelam predomínio de periodicidade em toda extensão das representações, concentrada até 8000 Hz, conforme indicado por seta na figura 31. Registra-se energia difusa com características de periodicidade até cerca de 11500 Hz.

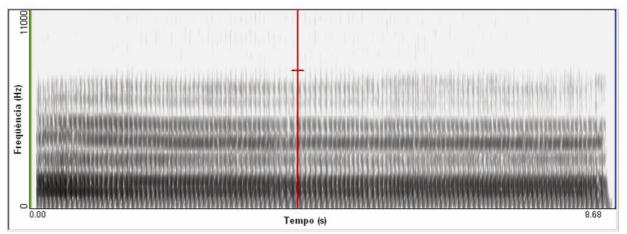
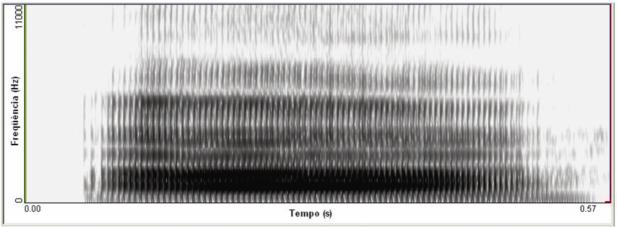


Figura 31. Espectrograma de banda larga com limite superior de freqüência estabelecido em 11000 Hz da emissão da vogal [a:] por iR, digitalizada na freqüência de amostragem de 22050 Hz.

[aaa]

Registro dominante de periodicidade em toda extensão das representações, concentrada até 12000 Hz (figuras 32a e 32b). Registro de ataques vocais bruscos.





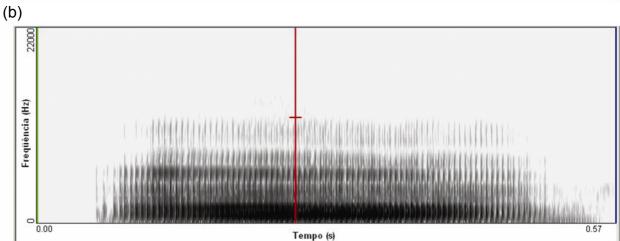
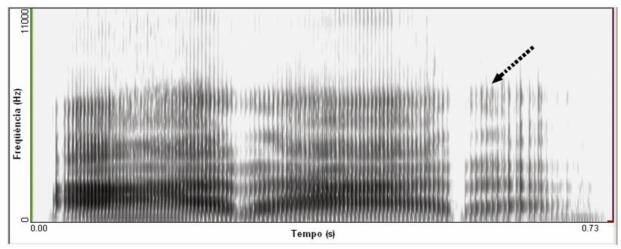


Figura 32. Espectrogramas de banda larga com limites superiores de freqüência estabelecidos em 11000 (a) Hz e 22000 Hz (b) da mesma emissão da vogal central de [aaa] por iR, digitalizada nas freqüências de amostragem de 22050 Hz e 44100 Hz.

#### Trechos de emissão encadeada "a arara"

Energia periódica registrada até o limite máximo de 12100 Hz, conforme indicado por seta na figura 33b. Até o limite de freqüência de 8150 Hz, há maior concentração de energia em relação às demais faixas referidas. Certas irregularidades no traçado podem ser evidenciadas nas vogais em posição pós-tônica, numa representação compatível com ajuste de crepitância ao final da emissão, sinalizada por seta com traçado pontilhado (Figura 33a).

(a)



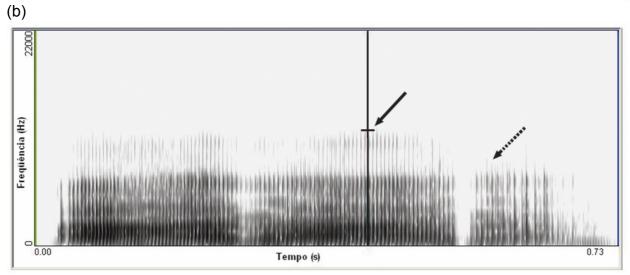


Figura 33. Espectrogramas de banda larga com limites superiores de freqüência estabelecidos em 11000 (a) e 22000 Hz da mesma emissão de "a arara" por iR, digitalizada nas freqüências de amostragem de 22050 Hz e 44100 Hz.

## 7.1.2.3. Tempo de início de sonorização - VOT

As medidas de *VOT* expressas em segundos(s) para três emissões de sons consonantais plosivos surdos ([p] [t] [k]) e sonoros ([b] [d] [g]) efetuadas pelos indivíduos do grupo estudado (i1 a i4) e do indivíduo referência (iR) são expostas na tabela 1.

TABELA 1. VALORES DE *VOT* EXPRESSOS EM SEGUNDOS (S) PARA EMISSÕES DE CONSOANTES PLOSIVAS SURDAS E SONORAS DOS INDIVÍDUOS DO GRUPO ESTUDADO (i1 A i4) E DO INDIVÍDUO REFERÊNCIA (iR)

Indivíduos	VOT (s)								
	[pa]	[per]	[bε]	[tar]	[da]	[ka]	[ga]		
i1a	0,013	0,010	-0,074	0,006	-0,096	0,035	-0,041		
i1b	0,015	0,013	-0,069	0,003	-0,086	0,034	-0,038		
i1c	0,010	0,015	-0,074	0,007	-0,085	0,044	-0,045		
i2a	0,010	0,010	-0,157	0,031	-0,105	0,033	-0,074		
i2b	0,013	0,014	-0,143	0,033	-0,098	0,049	-0,087		
i2c	0,020	0,012	-0,160	0,028	-0,099	0,030	-0,090		
i3a	0,006	0,012	-0,110	0,021	-0,086	0,023	0,024		
i3b	0,009	0,009	-0,109	0,014	-0,078	0,021	0,025		
i3c	0,008	0,016	-0,106	0,014	-0,084	0,020	0,023		
i4a	0,010	0,004	0,000	0,030	0,000	0,040	0,028		
i4b	0,016	0,004	-0,081	0,026	0,005	0,050	0,031		
i4c	0,014	0,025	-0,072	0,024	0,003	0,042	0,029		
iRa	0,012	0,007	-0,073	0,015	-0,080	0,024	-0,089		
iRb	0,012	0,008	-0,099	0,014	-0,115	0,029	-0,076		
iRc	0,013	0,012	-0,085	0,021	-0,087	0,040	-0,081		
jenda:						(p	<0,05)		

Legenda:

a= leitura 1

b= leitura 2

c= leitura 3

A distribuição das medidas de *VOT* é apresentada na figura 34 de forma diferenciada para o grupo de consoantes plosivas surdas e sonoras do português, de acordo com as médias calculadas para a emissão de cada uma das sílabas analisadas nas três leituras dos indivíduos do grupo estudado (i1 a i4) e indivíduo referência (iR).

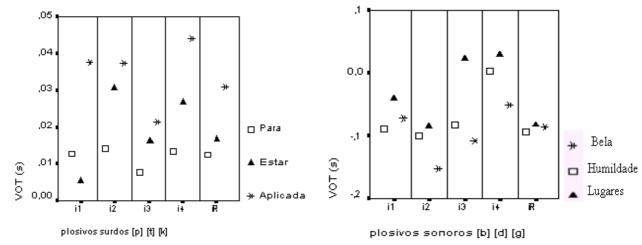
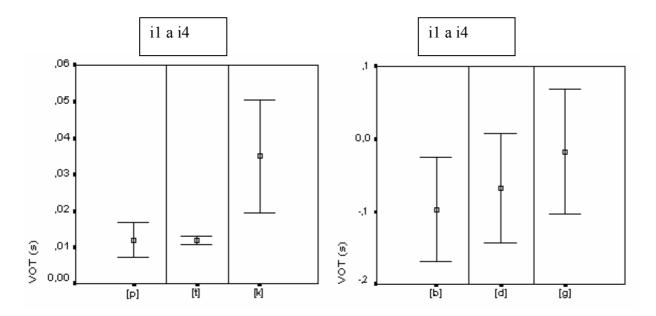


Figura 34. Distribuição das médias de medidas de *VOT* em segundos de três emissões para as consoantes plosivas surdas nas sílabas [p], [t] e [k] e sonoras nas sílabas [b] [d] e [g] dos indivíduos do grupo estudado (i1 a i4) e do indivíduo referência (iR).

Os dados referentes à análise das sílabas com sons consonantais plosivos surdos e sonoros também são apresentados separadamente com relação às médias para o grupo grupo estudado (i1 a i4) e referência (iR) na figura 35.



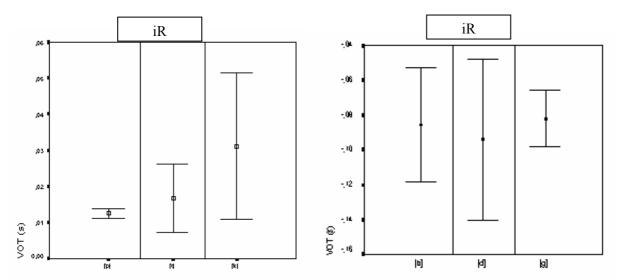


Figura 35. Distribuição das médias dos valores de VOT em segundos de consoantes plosivas surdas [p], [t] e [k] nos respectivos vocábulos **pa**ra, es**tar** e apli**ca**da e sonoras [b], [d] e [g] nos respectivos vocábulos humil**da**de e lu**ga**res emitidos pelo grupo estudado (i1 a i4) e indivíduo referência (iR).

A análise estatística revelou diferenças significativas (p<0,05) para as sílabas [ta] e [ga] de i1, [bε] e [ta] de i2, [pa] [bε] [ga] de i3 e [ta] [da] [ka] e [ga] de i4, todos comparativamente aos valores revelados por iR. No anexo 11 seguem dispostos os gráficos referentes à distribuição de medidas para cada uma das sílabas.

Os valores de *VOT* foram extraídos por complementação de inspeção de onda acústica e Lx, cujos valores foram dispostos em gráficos comparativos no anexo 11.

## 7.1.2.4. Estrutura harmônica e padrão de formantes

Os espectros *FFT* da vogal [a] em posição tônica das emissões "a arara" revelaram os limites de máximo registro de harmônicos, mesmo que mesclados a ruído, em freqüências máximas de 4700Hz para i1, 4688 Hz para i2, 3942 Hz para i3, 1147 Hz para i4 e, finalmente, 6521 Hz para iR. As medidas correspondentes à freqüência (em Hz) e amplitude (em dB) dos dois primeiros harmônicos (H1 e H2) dos indivíduos do grupo estudado (i1 a i4) e do indivíduo referência (iR) são expostas na tabela 2, cujas médias de três emissões são representadas na figura 36.

TABELA 2. VALORES DE FREQÜÊNCIA (EM Hz) E AMPLITUDE (EM dB) DOS DOIS PRIMEIROS HARMÔNICOS (H1 E H2) DA EMISSÃO DA VOGAL [a] EM POSIÇÃO TÔNICA EM "A ARARA" EM TRÊS LEITURAS PELOS INDIVÍDUOS DO GRUPO ESTUDADO (i1 A i4) E DO INDIVÍDUO REFERÊNCIA (iR)

	Estrutura Harmônica								
Indivíduo	Η´	1	H2						
maividuo	Freqüência	Amplitude	Freqüência	Amplitude					
	(Hz)	(dB)	(Hz)	(dB)					
i1a	211	37	412	38					
i1b	200	34	400	37					
i1c	200	39	412	41					
i2a	200	38	434	31					
i2b	223	42	456	35					
i2c	290	38	579	37					
i3a	301	37	579	37					
i3b	278	35	556	34					
i3c	267	34	523	32					
i4a	189	38	400	23					
i4b	178	36	367	19					
i4c	212	34	390	28					
iRa	189	31	367	34					
iRb	200	29	390	34					
iRc	267	36	557	31					

Legenda:

a= leitura 1

b= leitura 2

c= leitura 3

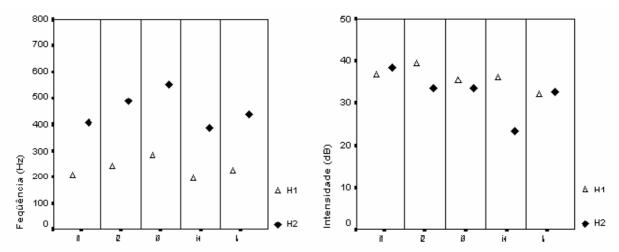
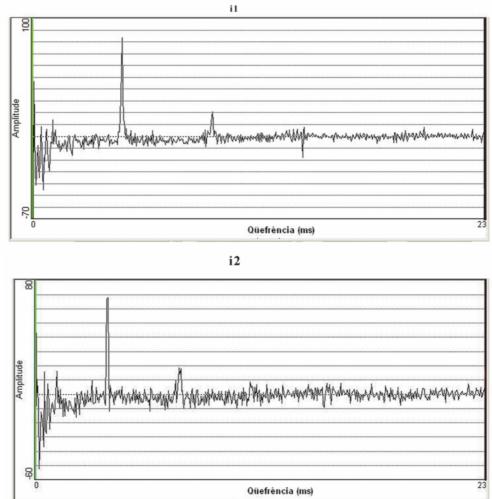


Figura 36. Médias de freqüência (Hz) e amplitude (dB) do primeiro (H1) e segundo (H2) harmônicos de três emissões da vogal [a] em posição tônica da emissão de "a arara" em três leituras dos indivíduos do grupo estudado (i1 a i4) e referência (iR)

A figura 37 apresenta os gráficos correspondentes à análise cepstral da emissão que revelou o pico de maior amplitude dentre as três leituras realizadas pelos indivíduos do grupo estudado (i1 a i4) e indivíduo referência (iR).



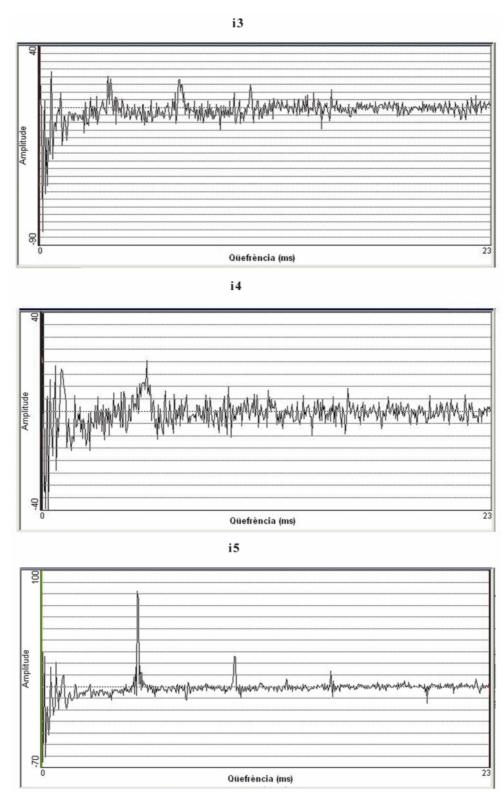


Figura 37. Traçados cepstrais com informações de qüefrência (em ms) e nível de amplitude para a emissão de "a arara" por i1, i2, i3, i4 e iR.

Os valores da análise cepstral são expressos em função do tempo (qüefrência, em ms) e do nível de amplitude dos picos detectados nos correspondentes gráficos e

registrados na tabela 3, incluindo estimativa de f0, para as emissões efetuadas pelos indivíduos do grupo estudado (i1 a i4) e indivíduo referência (iR).

TABELA 3. VALORES DE QÜEFRÊNCIA (EM ms) E DO NÍVEL DE AMPLITUDE DO PRINCIPAL PICO E DO VALOR DE f0 (EM Hz) CORRESPONDENTE À ANÁLISE CEPSTRUM DA EMISSÃO DA VOGAL [a] EM POSIÇÃO TÔNICA EM "A ARARA" EM TRÊS LEITURAS PELOS INDIVÍDUOS DO GRUPO ESTUDADO (i1 A i4) E INDIVÍDUO REFERÊNCIA (iR)

L. P. Zil	Parâmetro	cepstral	f0
Indivíduo	Qüefrência	Nível de	_
	(ms)	Amplitude	(Hz)
i1a	4,7487	71	210
i1b	5,022	90	199
i1c	4,6462	101	215
i2a	4,8854	41	204
i2b	4,5437	36	220
i2c	3,7238	68	268
i3a	3,553	20	281
i3b	4,6462	14	215
i3c	3,3822	14	295
i4a	4,9537	19	201
i4b	5,3295	20	187
i4c	5,4662	21	182
iRa	4,9537	93	201
iRb	4,9879	83	200
iRc	3,6213	53	276

Legenda:

a= leitura 1

b= leitura 2

c= leitura 3

As medidas correspondentes à freqüência (em Hz) e à amplitude (em dB) dos quatro primeiro formantes (F1, F2, F3 e F4) das emissões da vogal tônica de "a arara" pelos indivíduos do grupo estudado (i1 a i4) e do indivíduo referência (iR), resultantes da aplicação dos algoritmos *LPC* e *FFT*, são apresentadas na tabela 4, cujas médias de três emissões são representadas na figura 38.

TABELA 4. VALORES DE FREQÜÊNCIA (Hz) E AMPLITUDE (dB) DOS QUATRO PRIMEIROS FORMANTES (F1, F2, F3 E F4) DA EMISSÃO DA VOGAL [a] EM POSIÇÃO TÔNICA EM "A ARARA" EM TRÊS LEITURAS PELOS INDIVÍDUOS DO GRUPO ESTUDADO (i1 A i4) E DO INDIVÍDUO REFERÊNCIA (iR)

	Padrão de formantes								
Indivíduo	F	1	F2		F:	F3		4	
	F <sub>1</sub>	A1	F <sub>2</sub>	A2	F <sub>3</sub>	A3	F <sub>4</sub>	A4	
	(Hz)	(dB)	(Hz)	(dB)	(Hz)	(dB)	(Hz)	(dB)	
i1a	868	48	1681	44	2639	36	3830	26	
i1b	824	42	1726	39	2639	37	3786	9	
i1c	824	54	1692	46	2628	42	3953	8	
i2a	946	40	1670	39	2734	19	3819	12	
i2b	968	45	1680	41	2425	28	3393	18	
i2c	935	47	1536	47	2383	34	4031	26	
i3a	1057	44	1703	43	2862	34	3452	20	
i3b	981	41	1793	41	2915	36	3812	24	
i3c	1072	40	1825	34	2834	31	3612	22	
i4a	991	39	1748	38	2695	26	4020	25	
i4b	935	37	1637	32	2683	31	3953	23	
i4c	801	32	1637	34	2695	27	3997	19	
iRa	980	46	1648	44	2472	31	3775	28	
iRb	989	43	1808	39	2404	30	3914	33	
iRc	1127	41	1680	30	2531	18	3414	15	

Legenda:

a= leitura 1

b= leitura 2

c= leitura 3

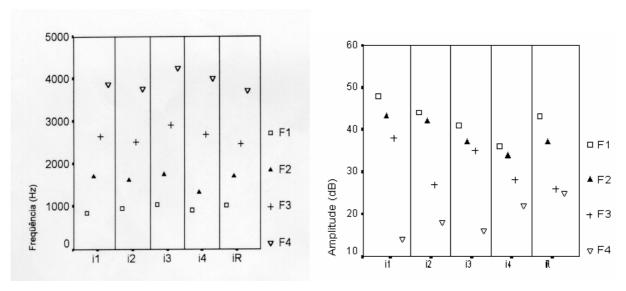


Figura 38. Médias de freqüência (Hz) e amplitude (dB) dos quatro primeiros formantes (F1, F2, F3 e F4) da emissão da vogal [a] em posição tônica em "a arara" para três leituras pelos indivíduos do grupo estudado (i1 a i4) e do indivíduo referência (iR).

A análise fatorial revelou o agrupamento em quatro classes correspondentes às três leituras de cada um dos quatro falantes do grupo estudado (i1, i2, i3 e i4), em que o fator subjacente com 45,41% de influência denotou diminuição de energia espectral em faixas acima de 9000 Hz e aumento por volta de 4000 Hz para i1, diminuição nas faixa de 6600 Hz a 7100 Hz na classe referente a i2, aumento em intervalo de 2860 a 3100 Hz seguida de declínio até 3400 Hz na classe referente a i3 e, finalmente, diminuição de energia espectral por muitos pontos, concentrados entre as proximidades das faixas compreendidas desde 1900 a 2200 Hz e de 10600 a 11025 Hz para os agrupamentos referentes a i4.

#### 7.2. Análise perceptivo-auditiva das amostras de fala

A apresentação dos resultados está dividida de acordo com as duas modalidades de avaliação nesta esfera, qualidade vocal e inteligibilidade, representadas nos próximos itens desta seção.

#### 7.2.1. Avaliação da qualidade vocal com motivação fonética

Os resultados dos julgamentos de grau geral de alteração vocal e de agradabilidade atribuídos às amostras de fala de i1 a i4 e iR nas etapas 1 A e 1 B do roteiro de avaliação perceptiva da qualidade vocal com motivação fonética são apresentadas na figura 39.

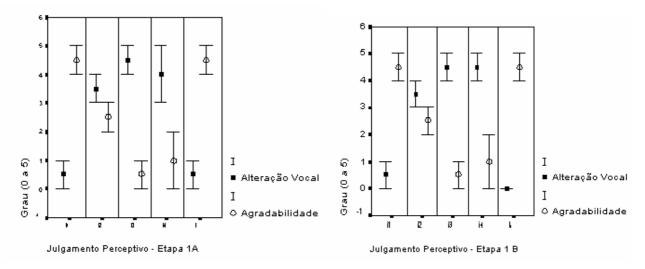


Figura 39. Julgamentos de grau geral de alteração vocal e de agradabilidade atribuídos às amostras de fala dos indivíduos do grupo estudado (i1 a i4) e indivíduo referência (iR) nas etapas 1A e 1B de julgamento da qualidade vocal.

Os resultados referentes à livre atribuição de qualidade vocal realizada pelos juízes anteriormente à proposta de roteiro com motivação fonética são apresentados no quadro 4 para cada um dos indivíduos do grupo estudado (i1 a i4) e indivíduo referência (iR) na etapas 1A e 1B.

QUADRO 4. JULGAMENTOS DE QUALIDADE VOCAL ATRIBUÍDOS ÀS AMOSTRAS DE FALA DOS INDIVÍDUOS DO GRUPO ESTUDADO (i1 A i4) E INDIVÍDUO REFERÊNCIA (IR) NAS ETAPAS 1A E 1B.

Qualidade Vocal	i1	i2	i3	i4	iR
Normal/adaptada					• 🗆
Modal/neutra					
Rouca		• 🗆		• 🗆	
Áspera		•	• 🗆	• 🗆	
Soprosa	• (-) 🗆 (-)	•	• 🗆	• 🗆	
Diplofônica			• 🗆	• 🗆	
Hiponasal					
Hiperfunção	•	• □	•		•
Ressonância	□ (-)				• (-)
Laringo-faríngea					
Instabilidades	•	• 🗆			

Legenda:

- Etapa 1A
- □ Etapa 1B
- (-) referida em menor grau ou por metade dos juízes

Os resultados referentes aos julgamentos de combinações de ajustes laríngeos e supralaríngeos para caracterização do perfil de qualidade vocal segundo roteiro de avaliação foneticamente motivado são apresentados no quadro 5.

QUADRO 5. PERFIL DA QUALIDADE VOCAL SEGUNDO JULGAMENTOS ATRIBUÍDOS ÀS AMOSTRAS DE FALA DOS INDIVÍDUOS DO GRUPO ESTUDADO (i1 A i4) E INDIVÍDUO REFERÊNCIA (iR) NAS ETAPAS 2A E 2B.

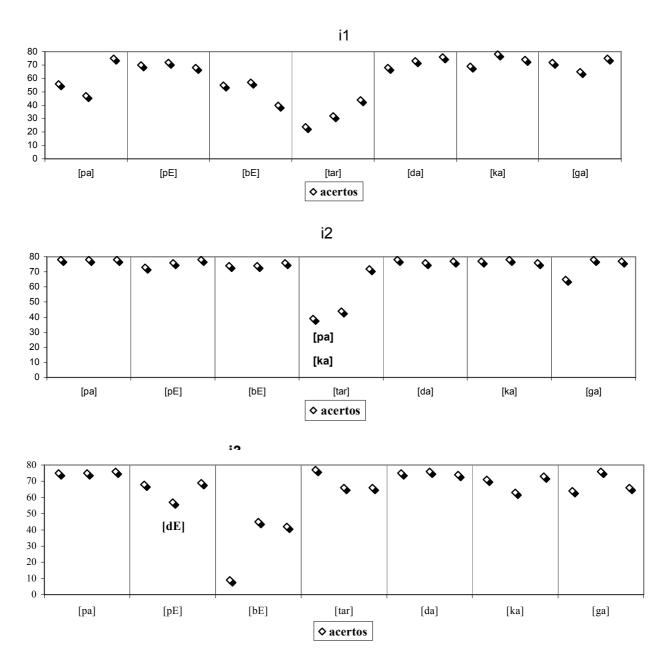
CATEGORIA	AJUSTES	GRAU	GRAUS DE ESCALA					
		neutro	1	2	3	4	5	6
SUPRALARÍNGEOS	LARÍNGEOS							
LONGITUDINAIS	Laringe alta				0			
	Laringe baixa			$\Diamond$				
	LABIAIS							
	Labiodentalização			0				
	Protrusão							
SUPRALARÍNGEOS TRANSVERSAIS	LABIAIS							
IRANSVERSAIS	Arredondados							
	Estirados							
	MANDIBULARES							
	Fechada					lack		
	Aberta							
	LINGUAIS - PONTA/							
	LÂMINA							
	Avançada							
	Recuada							
	LINGUAIS - CORPO							
	Avançado							
	Retraído							
	Elevado							
	Abaixado							
	LINGUAIS – BASE							
	Constrição faríngea							
	Expansão faríngea							
VELOFARÍNGEOS	ACOPLAGEM VELAR							
	Nasal							
	Denasal							
FONATÓRIOS	Modal	<b>A</b> 0						
	Falsete		Г					
	Escape de ar		Т				$\rangle$	
	Voz soprosa		Г				_	
	Vocal fry (crepitância)					i(◊)		
	Voz crepitante							
	Voz áspera					00		
	Hiperfunção				0			
	Hipofunção							
	inlofonia ( )quebras ( ) instabilida		1	L	1	ı	1	Щ

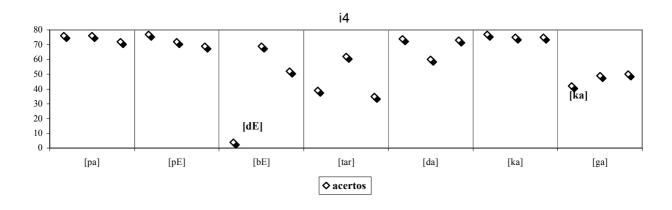
Ocorrências em curto termo: ( **a**)diplofonia ( )quebras ( ) instabilidades Legenda:

i1=  $\triangle$  i2=  $\bigcirc$  i3=  $\square$  i4=  $\diamondsuit$  iR=  $\square$  i= ajuste intermitente

# 7.2.2. Inteligibilidade de fala (sílabas)

Os dados referentes à inteligibilidade de sílabas são apresentados na figura 40 com relação ao número de acertos para um total de 78 julgamentos por sílaba. Quando os erros gerados concentraram-se na referência predominante a um mesmo estímulo, o mesmo aparece transcrito no gráfico.





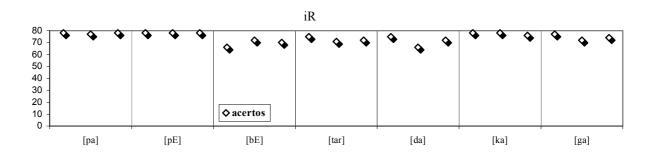


Figura 40. Total de acertos em julgamentos de inteligibilidade para as três emissões das sílabas [pa], [pɛ], [bɛ], [tar], [da], [ka] e [ga] dos indivíduos do grupo estudado (i1 a i4) e indivíduo referência (iR).

Diferenças estatisticamente significativas (p<0,05) foram identificadas para as sílabas [pa] [pɛ] [bɛ] e [tar] de i1, [pa] [pɛ] [bɛ] [ka] de i3 e [pɛ] [tar] [ga] de i4, todos comparativamente aos valores de iR.

## 7.3. Análise da dimensão fisiológica das amostras de fala

A exposição dos resultados passa a ser abordada enquanto informações provenientes da inspeção da eletroglotografia, complementadas por dados de exame otorrinolaringológico, mais especificamente pelos procedimentos de telelaringoestroboscopia e videoquimografia. Tais esferas de avaliação representam as duas seções deste tópico final da apresentação de resultados.

## 7.3.1 Eletroglotografia (EGG)

As informações da eletroglotografia são abordadas enquanto observações da onda Lx, incluindo sua avaliação semi-automática, além de medidas extraídas de seu processamento automático e, finalmente, de índices provenientes da inspeção das ondas Lx e acústica.

## 7.3.1.1. Inspeção dos traçados da onda Lx

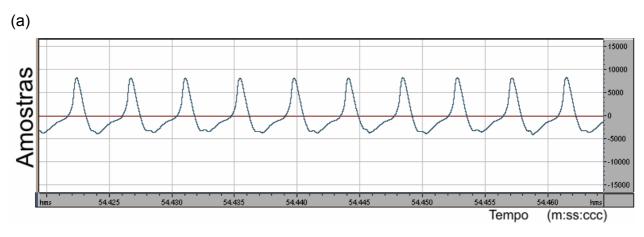
Os dados obtidos da inspeção da onda Lx passam a ser apresentados de forma diferenciada para cada indivíduo do grupo estudado (i1 a i4) e indivíduo referência (iR), segundo parâmetros de avaliação abordados em 6.3.1.

- i1

Em avaliação semi-automática o traçado recebeu a avaliação 4, entretanto a amplitude encontra-se diminuída (figura 41a). Os traçados das várias emissões revelaram o predomínio de deslocamento do pico da onda para direita, com alongamento do ramo correspondente à fase de aumento de contato e abaulamento do rodapé em algumas situações. Verificou-se adição de ruído leve e esporádica, com amplitude constante entre as emissões, denotando manutenção do padrão de contato atingido.

O traçado modificou-se apenas nas emissões em freqüências graves e agudas, quando o pico mostrou-se achatado e o deslocamento do pico acima referido não foi revelado. Nas emissões em forte intensidade, tal característica também não foi registrada.

Não foram evidenciadas mudanças no traçado de onda em função da pressão digital sobre os eletrodos durante as várias etapas de registro. O traçado de eletroglotograma diferenciado (DEGG) apresentou definição de pontos correspondentes aos momentos do fechamento e abertura glótica (Figura 41b).



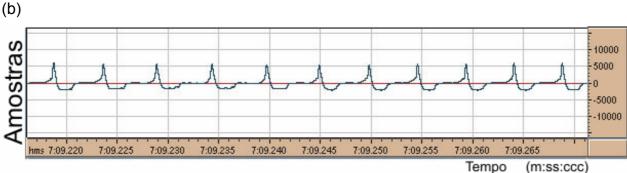


Figura 41. Traçados da onda Lx (a) para a emissão de vogal [a] central de [aaa] e DEGG (b) para a emissão vogal [a] tônica na emissão "a arara" (b) de i1.

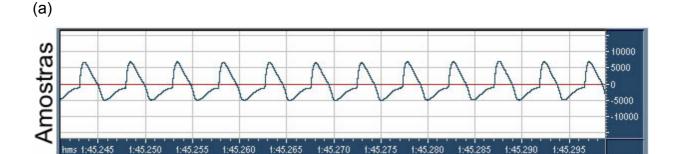
- i2

Em avaliação semi-automática o traçado da onda Lx recebeu a avaliação entre 3, com trechos de maior amplitude e alguns indícios de ruído e irregularidades (figura 42a), e 2, quando a amplitude do sinal diminui (figura 42c).

As observações ao longo das várias emissões revelaram o deslocamento do pico da onda Lx para direita, com alongamento e freqüentemente, arredondamento do ramo de fase de aumento de contato. Em emissão em forte intensidade houve registro de aumento de sua amplitude com manutenção da geometria da onda.

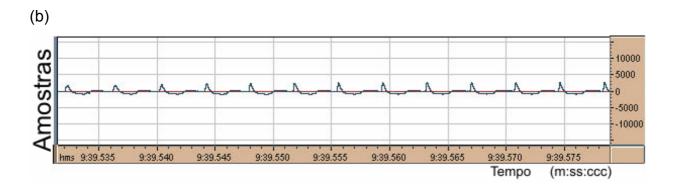
No caso de emissões em freqüências agudas, apresentou séries de f0 (figura 42c), identificadas por pequenos ciclos entre aqueles característicos dos demais registros, numa correspondência às bifurcações referidas do ponto de vista acústico.

As emissões de i2 caracterizam-se também, por apresentar maiores irregularidades no sinal ao início de vocalizações sustentadas e entrecortadas. Observa-se, entretanto, que, nestas situações, houve variação do traçado e da amplitude da onda, porém a duração do ciclo foi constante. Os traçados DEGG foram visíveis apenas no registro de emissão encadeada, com amplitude reduzida (figura 42b).



Tempo

(m:ss:ccc)



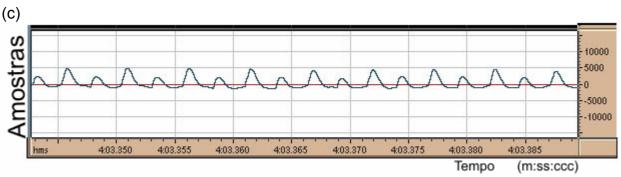


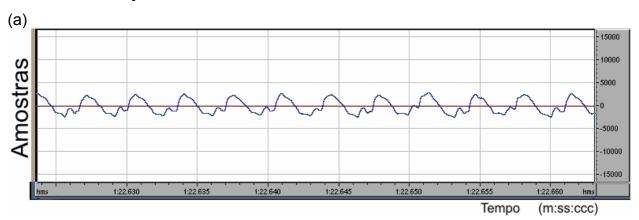
Figura 42. Traçados da onda Lx (a) e DEGG (b) para a emissão vogal [a] tônica em "a arara" e traçado da onda Lx (c) para emissão em freqüência aguda de i2.

- i3

Em avaliação semi-automática o traçado da onda Lx variou entre 2, com séries de f0 presentes de forma consistente e 1, com importante aperiodicidade e diminuição da amplitude (figura 43a e 43b). Os traçados mostraram-se portanto variáveis ao longo de várias emissões, incluindo adição de ruído, achatamento de pico e alongamento do pulsos e presença de séries de f0, conforme revela a figura 43a.

No caso de emissões vocálicas em freqüências graves e agudas e intensidade fraca praticamente não houve registro de onda, apresentando-se como um esboço de traçado de amplitude mínima. Outro achado freqüente foi a alteração de traçado em trechos iniciais de vogais, com grande variação de formato dos pulsos na onda Lx.

O traçado DEGG não foi registrado, denotando mínima amplitude e predomínio de ruído. As medidas automáticas de avaliação da onda Lx não foram coletadas em função dos achados descritos.



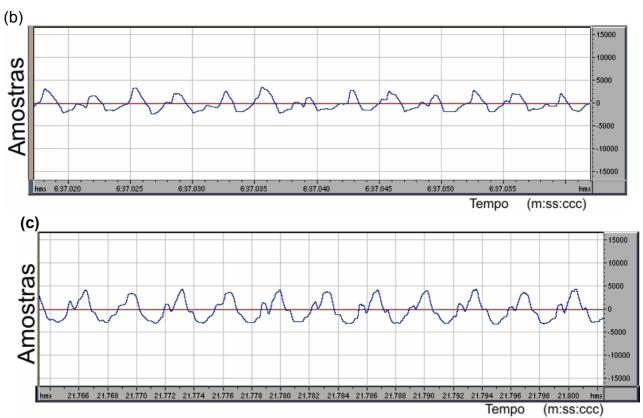


Figura 43. Traçados da onda Lx (a) para emissões de vogal [a] em "a arara" e (b) vogal [a] central da seqüência [aaa] (c) de i3.

- i4

Em avaliação semi-automática o traçado da onda Lx recebeu a avaliação 1, com grande variabilidade detectada, desde trechos sem registro de pulsação, até momentos esporádicos de maior amplitude e detecção de indícios de contato (figura 44).

O formato da onda não se manteve ao longo da gravação registrada, denotando aperiodicidade. Nas emissões de vogais sustentadas o traçado foi praticamente ausente. Na emissão em agudos, apesar da mínima amplitude, o traçado guardou uma característica de ramo ascendente com menor duração em relação ao restante (figura 44). Em forte intensidade não houve registro de contato, apenas de maiores irregularidades no traçado.

Na onda acústica, observou-se a grande variabilidade de amplitude ao longo do sinal, dificultando inclusive a demarcação do trecho necessário à estabilização do sinal, para cômputo de tempo de subida de contato.

No traçado DEGG não houve registro de qualquer atividade, mesmo com recursos de aumento de resolução da tela. As medidas automáticas de avaliação da onda Lx não foram coletadas em função das irregularidades descritas.

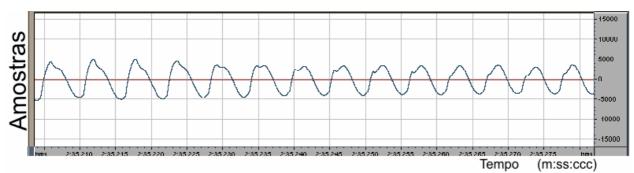
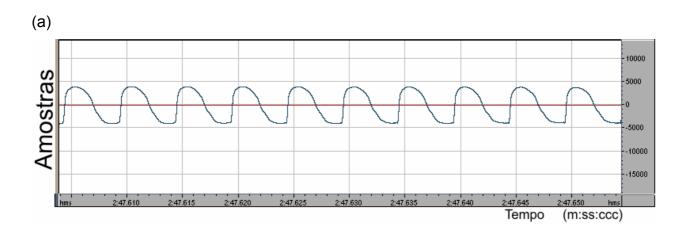


Figura 44. Traçado da onda Lx para a emissão de [i] em posição tônica, representativa da maior amplitude detectada nas amostras de i4.

- iR

A avaliação semi-automática atribuída foi 4, constatando-se periodicidade do sinal (figura 45a). Não foram observadas variações do traçado em função da pressão digital exercida sobre os eletrodos. Traçado DEGG revelou características da possibilidade de demarcação de momentos de abertura e de fechamento, porém de baixa amplitude (figura 45b).

Nas emissões em tons graves, registro de onda mais achatada e nos agudos, onda de mínima amplitude com diminuição do contato. Houve pouca variação da traçado especialmente em termos de amplitude para a emissão forte em relação aos padrões habituais detectados. Em emissões de fraca intensidade, o declínio de amplitude foi detectado.



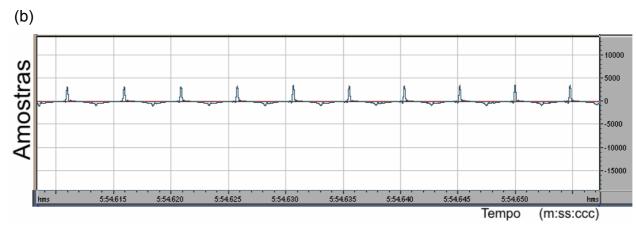


Figura 45. Traçados da onda Lx (a) para as emissões de vogal [a] em posição tônica em "a arara" e traçado DEGG (b) para a emissão vogal [a] em posição tônica em "a arara" de iR.

#### 7. 3.1.2. Medidas extraídas a partir do processamento automático da onda LX

As medidas de f0, de perturbação (jitter e shimmer), coeficiente de contato e índice de velocidade referentes às emissões da vogal [a] em padrões entrecortado [aaa] e encadeado, na seqüência "a arara" para i1, i2 e iR são apresentadas na tabela 5 e figura 46. Foram respeitados os critérios de avaliação semiautomática previamente expostos no item anterior, correspondentes à possibilidade de extração a partir do processamento automático das ondas Lx.

TABELA 5. MEDIDAS DE f0 (Hz), JITTER (%), SHIMMER (%), COEFICIENTE DE CONTATO (%) E ÍNDICE DE VELOCIDADE PARA TRÊS EMISSÕES DA VOGAL CENTRAL DE TRÊS SEQÜÊNCIAS DE EMISSÕES [aaa] E DUAS DA VOGAL TÔNICA EM "A ARARA" PARA DOIS INDIVÍDUOS DO GRUPO ESTUDADO (i1 E i2) E INDIVÍDUO REFERÊNCIA (iR)

	Medidas EGG										
Indivíduos	f0 (Hz)		Jitter (%)		Shimmer RMS (%)		Coeficiente de contato (% )		Índice de velocidade		
	[aaa]	[arara	[aaa]	arara	[aaa]	arara	[aaa]	arara	[aaa]	arara	
i1a	228	204	0,79	1,24	5,66	3,82	41	41	0,30	0,19	
i1b	237	199	0,79	0,94	5,94	5,63	38	39	0,20	0,30	
i1c	210		0,70		6,61		40		0,16		
i2a	252	230	1,35	1,27	10,28	3,85	61	58	0,19	0,13	
i2b	254	231	0,71	1,64	3,71	5,14	64	61	0,22	0,11	
i2c	249		0,99		5,07		65		0,20		
iRa	217	191	0,49	1,27	3,77	6,13	64	59	-0,49	-0,55	
iRb	193		0,78		5,78		65		-0,53		
iRc	189		0,77		6,31		64		-0,49		
eferência <sup>3</sup> :			0,1 a 0,3%		0,4 a 1,40	0,4 a 1,40%		51 a 57%		-1,00 a -0,71	

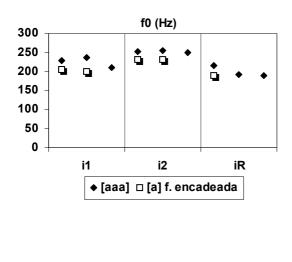
Legenda:

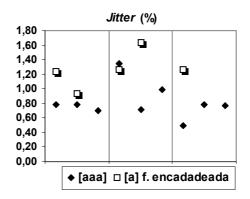
a = 1<sup>a</sup> emissão b = 2<sup>a</sup> emissão

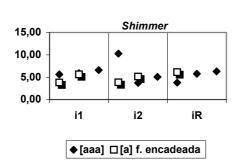
c = 3<sup>a</sup> emissão

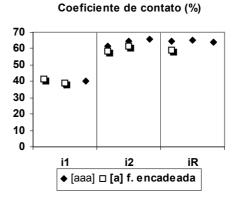
--- = amostra não disponível

<sup>&</sup>lt;sup>3</sup> Valores sugeridos como de maior probabilidade para indivíduos não disfônicos em estudo de VIEIRA (1997).









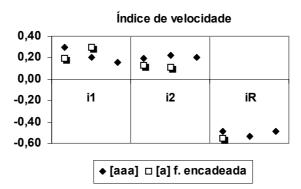


Figura 46. Distribuição de medidas de f0 (Hz), Jitter (%), Shimmer (%), coeficiente de contato (%) e índice de velocidade para três emissões da vogal central das seqüências de emissões [aaa] e duas emissões da vogal tônica das emissões "a arara" para dois indivíduos do grupo estudado (i1 e i2) e indivíduo referência (iR).

#### 7.3.1.3. Medidas extraídas a partir do processamento simultâneo das ondas Lx e acústica

As medidas, extraídas pelos recursos de inspeção simultânea de ondas Lx (f0 e coeficiente de contato) e acústica (tempo de subida de contato), foram realizadas na vogal em posição central nas emissões entrecortadas [aaa] e na sílaba tônica das três emissões "a arara". Na tabela 6 e na figura 47 estão expostas as medidas para todos os falantes do grupo estudado (i1 a i4) e para o indivíduo referência (iR).

TABELA 6. MEDIDAS ELETROGLOTOGRÁFICAS DE F0 (HZ), QUOCIENTE DE CONTATO (CC EM %) E ACÚSTICA DE TEMPO DE SUBIDA DE CONTATO (TSUB EM MS) PARA AS TRÊS EMÍSSÕES DA VOGAL CENTRAL DA SEQÜÊNCIA [AAA] DOS INDIVÍDUOS DO GRUPO ESTUDADO (I1 A I4) E DO INDIVÍDUO REFERÊNCIA (IR).

	N	/ledida	s EGG	<del></del>			
					Medida acústica		
La alla dala a a	f	0	Coefi	ciente			
Indivíduos	(Hz)			ontato	Tsub	(ms)	
			<del></del>	%)			
	[aaa]	arara	[aaa]	arara	[aaa]	arara	
i1a	233	207	34	37	065	026	
i1b	238	197	34	33	062	041	
i1c	213		37		093		
i2a	259	222	39	41	083	037	
i2b	262	240	37	40	077	036	
i2c	268		37		106		
i3a	283	298	48	40	042	065	
i3b	290	298	54	40	066	028	
i3c	275		48		084		
i4a	308	198	*	*	156	012	
i4b	265	180	*	*	139	036	
i4c	286		*		016		
iRa	216	197	50	52	079	058	
iRb	191	200	40		060	055	
iRc	180		60		051		

Legenda:

As medidas apresentadas acima foram agrupadas em gráficos revelando a sua distribuição para a da análise da vogal [a] em amostras de emissões entrecortadas e de fala encadeada (trecho "a arara") para todos os falantes estudados (i1 a i4) e indivíduo

a= 1<sup>a</sup> emissão b= 2<sup>a</sup> emissão

c= 3<sup>a</sup> emissão

<sup>--- =</sup> amostra não disponível

<sup>\* =</sup> sem condições de registro

referência (iR), com relação aos valores de f0 e coeficiente de contato, além da medida acústica de tempo de subida (Tsub).

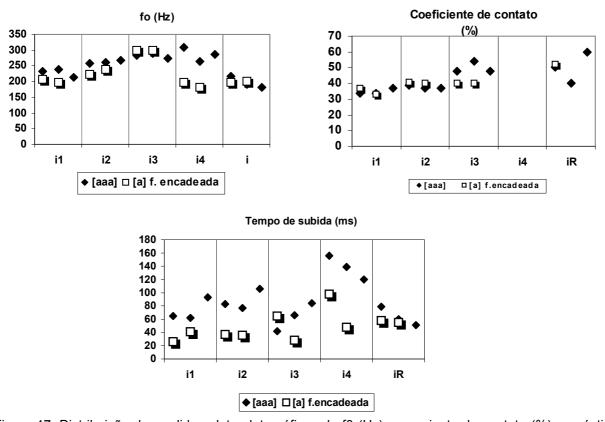


Figura 47. Distribuição de medidas eletroglotográficas de f0 (Hz) e quociente de contato (%) e acústica de tempo de subida (s) da vogal central das seqüências de emissões [aaa] e da vogal tônica das emissões "a arara" para os indivíduos do grupo estudado (i1 a i4) e indivíduo referência (iR)

As medidas de VOT também foram extraídas em função da possibilidade de conjugar a inspeção dos traçados de ondas acústica e Lx. Os valores são apresentados na tabela 7.

TABELA 7. MEDIDAS DE *VOT* (EM SEGUNDOS) PARA AS TRÊS EMISSÕES DAS SÍLABAS TÔNICAS DOS SETE VOCÁBULOS (**PA**RA, ES**PER**TO, **BE**LA, ES**TAR**, HUMIL**DA**DE, APLI**CA**DA, LU**GA**RES) DOS INDIVÍDUOS DO GRUPO ESTUDADO (i1 A i4) E DO INDIVÍDUO REFERÊNCIA (iR) EXTRAÍDAS POR INSPEÇÃO SIMULTÂNEA DAS ONDAS LX E ACÚSTICA.

Indivíduos	VOT (s)											
	[pa]	[ρει]	[bε] [tar]	[da]	[ka]	[ga]						
i1a	0,012	0,008	-0,068 *	-0,073	0,032	-0,049						
i1b	0,012	0,010	-0,062 *	-0,051	0,030	-0,034						
i1c	0,011	0,009	-0,064 0,004	-0,050	0,038	0,019						
i2a	0,002	0,004	-0,150 0,030	-0,079	0,024	-0,069						
i2b	0,012	0,004	-0,14 0,023	-0,095	0,037	-0,093						
i2c	0,004	0,003	-0,156 0,026	-0,078	0,024	-0,087						
i3a	0,016	0,006	-0,101 0,028	-0,075	0,023	0,023						
i3b	0,013	0,007	-0,107 0,02	-0,146	0,019	0,026						
i3c	0,006	0,000	-0,099 0,023	-0,075	0,020	0,023						
i4a	0,036	0,031	0,028 0,029	-0,014	0,036	0,017						
i4b	0,019	0,019	0,085 0,02	-0,008	0,04	0,039						
i4c	*	*	* *	*	*	*						
iRa	0,009	0,005	-0,071 0,014	-0,074	0,022	-0,085						
iRb	0,010	0,009	-0,097 0,017	-0,101	0,023	-0,077						
iRc	0,010	0,009	-0,084 0,024	-0,081	0,037	-0,077						

Legenda

As medidas de VOT extraídas por EGG foram comparadas à extração acústica e são apresentadas em forma de gráficos no anexo 12 para cada indivíduo do grupo estudado (i1 a i4) e referência (iR).

a= primeira leitura

b= segunda leitura

c= terceira leitura

<sup>(\*)</sup> sem condições de registro

## 7.3.2. Complementação da descrição por meios fisiológicos - dados de exames otorrinolaringológico (telelaringoestroboscopia e videoquimografia)

Os resultados de exame otorrinolaringológico passam a ser descritos para cada indivíduo do grupo estudado, complementados por fotos indicativas dos ajustes mais freqüentes identificados no momento da avaliação. Ao lado da identificação do indivíduo avaliado, é indicado o tipo de procedimento efetuado, uma vez que um dos indivíduos examinados apresentou menor tolerância ao endoscópio rígido, dado acentuado reflexo nauseoso.

- i1 Dados de telescopia laríngea, associada à videoestroboscopia e videoquimografia
  - Condições gerais de exame: um pouco de secreção em hipofaringe e vestíbulo laríngeo
  - Aspectos de integridade estrutural e mobilidade: paralisia de prega vocal direita em posição paramediana, com prega vocal levemente arqueada (Figura 48a)
  - Fonação dinâmica vibratória:
    - Videoestroboscopia (Figura 48b)
      - fenda triangular posterior
      - assimetria de vibração
      - leve assimetria de fase
      - onda mucosa diminuída em amplitude do lado paralisado
      - mobilização de prega vestibular esquerda em grau leve
    - Videoquimografia (Figura 49)
      - periodicidade
      - assimetria de fase discreta

onda mucosa diminuída do lado paralisado (não se evidencia a borda inferior de prega vocal).

 abertura lenta e fechamento atrasado do lado paralisado em relação à prega vocal contralateral.





Figura 48. Imagens de laringe durante respiração (a) e fonação (b), proveniente de exame telescópico de laringe de i1. Em (b) associação de videoestroboscopia.

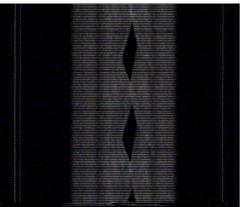




Figura 49. Imagens de videoquimografia da laringe durante a fonação de i1.

- i2 Dados de nasofibrolaringestroboscopia
- Condições gerais de exame: sem presença de secreção
- Aspectos de integridade estrutural e mobilidade: paralisia de prega vocal esquerda em posição paramediana a intermediária, com aparente desnível de prega vocal (figura 50a)
- Fonação dinâmica vibratória (figura 50b):
  - o fechamento completo
  - periodicidade
  - o assimetria de fase discreta
  - o presença de onda mucosa
  - assimetria de movimentação de aritenóides
  - o constrição supraglótica (ântero-posterior e medial) em grau moderado



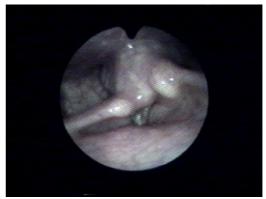


Figura 50. Imagens de laringe durante respiração (a) e fonação (b), proveniente de exame nasolaringoscópico de laringe de i2. Em (b) associação de videoestroboscopia.

- i3 Dados de telescopia laríngea, associada à videoestroboscopia e videoquimografia
  - Condições gerais de exame: prótese dentária móvel
  - Aspectos de integridade estrutural e mobilidade: estenose posterior fibrose de prega vocal bilateral (Figura 51a)
  - Fonação dinâmica vibratória:
    - Videoestroboscopia (figura 51b)
      - fechamento incompleto
      - diferença de fase entre as pregas vocais e na mesma prega vocal (esquerda) ao longo de sua extensão
      - diminuição de amplitude bilateral
      - · diminuição de onda bilateral acentuada
      - fase fechada incompleta
      - constrição supraglótica (ântero-posterior) em grau leve
    - Videoquimografia (Figura 52)
      - o aperiodicidade
      - o assimetria de fase
      - o diminuição da onda mucosa
      - o fase fechada de curta duração

 emissão em freqüência aguda: porção posterior de pregas vocais tem amplitude de vibração semelhante à direita e à esquerda

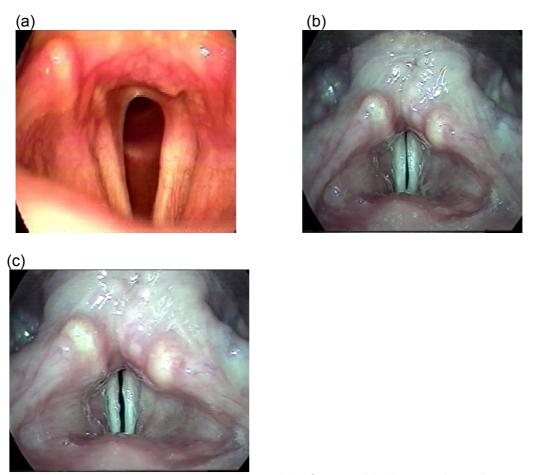
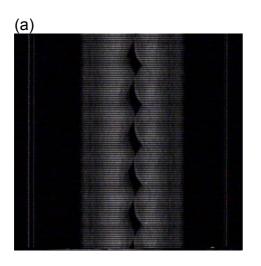
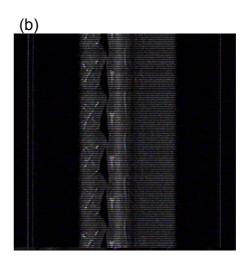


Figura 51. Imagens de laringe durante respiração (a) e fonação (b) (c), proveniente de exame telescópico de laringe de i3. Em (b) (c) associação de videoestroboscopia.





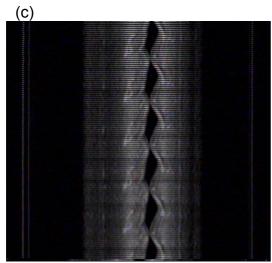


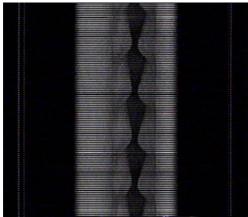
Figura 52. Imagens de videoquimografia da laringe durante a fonação de i3.

- i4 Dados de telescopia laríngea, associada à videoestroboscopia e videoquimografia
  - Condições gerais de exame: prótese dentária
  - Aspectos de integridade estrutural e mobilidade: Paralisia de prega vocal direita em posição paramediana (figura 53a)
  - o Fonação dinâmica vibratória (figura 53b):
    - Estroboscopia:
      - o fechamento incompleto
      - periodicidade
      - assimetria de amplitude (maior à esquerda)
      - o onda mucosa praticamente ausente
      - constrição supraglótica (medial), com padrão variável de movimentação de pregas vestibulares e vibração esporádica
    - Videoquimografia (Figura 54)
      - fechamento incompleto (ausente)
      - periodicidade
      - o assimetria de fase discreta
      - o amplitude maior à esquerda
      - o diminuição bilateral de onda mucosa





Figura 53. Imagens de laringe durante respiração (a) e fonação (b), proveniente de exame telescópico de laringe de i4. Em (b) associação de videoestroboscopia.



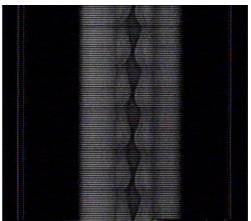


Figura 54. Imagens de videoquimografia da laringe durante a fonação de i4.

## 8. DA PRODUÇÃO À PERCEPÇÃO SONORA: DETERMINAÇÃO DA QUALIDADE VOCAL DE INDIVÍDUOS DISFÔNICOS SOB O OLHAR DA PERSPECTIVA FONÉTICA

No universo em que se inscreve a interface da produção e da percepção do sinal acústico de fala, os resultados expostos no capítulo anterior representam tão somente um primeiro passo a colaborar para a caracterização de uma das informações nele impressas, a saber, a qualidade vocal, enfocada em termos de confluência de parâmetros acústicos, perceptivos e fisiológicos. Tais considerações poderão colaborar para o avanço do almejado conhecimento das características de uma situação peculiar: a disfonia.

A decomposição do sinal acústico de fala de indivíduos disfônicos revelou diversas particularidades de atividade do aparelho fonador na situação de compensação das ações alteradas pelas manifestações de paralisia unilateral de prega vocal, em posição paramediana, por três dos indivíduos estudados (i1, i2 e i4) e de fibrose de prega vocal e de comissura posterior decorrentes de intubação orotraqueal prolongada por i3. Os achados apresentados possibilitaram o estabelecimento de interrelação entre alguns aspectos do julgamento perceptivo da qualidade vocal, explicitados pela combinação de ajustes laríngeos e supralaríngeos (modelo fonético), de índices acústicos a revelarem tais mobilizações (análises de longo e curto termo) e de dados da fisiologia laríngea, especialmente provenientes da eletroglotografia (conformação da onda Lx e medidas correspondentes) e do exame laringológico (telelaringovideoestroboscopia e videoquimografia). Tais indicativos da produção sonora foram correspondidos a estimativas de inteligibilidade de fala para estímulos de sílabas, com enfoque aos sons consonantais plosivos.

Diante da correspondência de parâmetros estudados, restou-nos a tarefa de relacioná-los à luz das reflexões que permeiam os campos de estudos da Fonética Acústica, em crescente evolução ao longo dos últimos cinqüenta anos, conforme mencionado na introdução deste trabalho. Conseqüentemente, os dados discutidos representam a resposta à tentativa de contemplar os procedimentos de varredura acústica de amostras de falantes disfônicos, valendo-nos dos vários recursos de representação atualmente disponíveis no campo da espectrografia sonora.

Estruturada desta forma, esta pesquisa enfoca a qualidade vocal numa abordagem que se afasta do estabelecimento da dicotomia normalidade e alteração vocal. Ao contrário, para avançar no enfoque da disfonia enquanto manifestação de caráter adaptativo, defendemos a posição de que o quadro a que ela se refere implica necessariamente diversos graus de mobilização para a tentativa de manutenção do produto sonoro, mobilização essa que revela a imensa plasticidade do aparelho fonador em face de alterações em sua estrutura, representadas por comprometimentos de mecanismos glóticos de características anteriormente delineadas. Tal concepção justifica a adoção de um grupo de quatro falantes e um indivíduo referência, no sentido de contemplar a variabilidade inerente à função estudada.

Neste particular, o estudo de população cede lugar ao enfoque de diversos eventos sonoros e da complexa relação entre eles com base no mesmo falante, em que produção e percepção encontram-se entrelaçadas pela dimensão acústica da estrutura sonora do estímulo em questão. Tal nível de correspondência é manifestado ao longo do fluxo da fala, enquanto características invariantes do falante, numa menção à abordagem de longo termo da emissão. Neste ponto, emerge a necessidade de correspondermos os achados sob a perspectiva de um referencial teórico, de acordo com os modelos anteriormente expostos.

Nesse universo, a retomada de estudos clássicos na área de análise acústica em quadros de disfonia revelou a necessidade de adoção procedimentos de cuidadosa inspeção do sinal, diante das referências à natureza do estímulo sonoro e às variações de intensidade por faixas de freqüência do sinal.

Hammarberg, Gauffin (1995) num importante estudo de caracterização acústica de alterações vocais ressaltaram que não há clara demarcação entre parâmetros de qualidade vocal para normalidade e alteração, sendo importante caracterizar os fatores acústicos que variam no tempo. Outras referências foram fundamentais para a constituição das bases de inspeção do sinal adotadas no presente estudo (Yanagihara, 1967; Rontal et al, 1975; Hirano et al, 1978; Kim et al, 1982; Dejonckere, Villarosa, 1986; Bless, Baken, 1992; Dejonckere, Wienecke, 1994; Titze, 1994, 1995; Ananthapadmanabha, 1995; Hammarberg, Gauffin, 1995; Holmberg et al, 1995).

O cuidado necessário à análise do sinal vocal alterado reflete-se no fato de conter irregularidades que dificultam o seu processamento e que, pela mesma razão, acabam por expor a sua maior característica estrutural: a aperiodicidade. Neste ponto, pode-se afirmar que a perturbação do sinal é justamente a dimensão que interfere de forma significativa na extração de medidas acústicas. Tais considerações tornam-se marcantes na vigência de disfonias decorrentes de comprometimentos de mecanismos glóticos, como o grupo eleito para este estudo.

Para minimizar as interferências nos procedimentos de extração de valores de f0, necessárias ao cômputo de medidas de perturbação, é necessário o refinamento do método de demarcação do sinal, porém, no momento atual, a implementação desse refinamento ainda se constitui em um desafio. Trabalhos como os de Vieira et al (1996a, 1996b), Lindsey, Vieira (1997), Vieira et al (1997) e Vieira et al (2002), cujos passos de análise da onda Lx e algoritmos para extração de medidas dela derivadas foram utilizados na presente pesquisa, particularmente no âmbito da eletroglotografia, representam avanços no sentido dessa implementação, primando pelo rigor científico voltado à obtenção de estimativas de alterações laríngeas.

No futuro, tais estimativas poderão sinalizar a necessidade de investigação mais detalhada de estruturas do aparelho fonador e dos mecanismos fisiológicos presentes. Entretanto, dificilmente poderão corresponder à qualidade vocal, a qual engloba aspectos além da irregularidade anteriormente referida. Representam, portanto, um passo na descrição de alguns aspectos dos eventos laríngeos, mais precisamente glóticos e acentuam a validade de estudos que se ocupam da criteriosa coleta de amostras de dados para a finalidade de análises acústica, perceptiva e eletroglotográfica.

Há que se destacar, também, o cuidado no delineamento e condução dos procedimentos de coleta e análise de dados apresentados, os quais permitem que se façam reflexões sobre a produção sonora, especialmente pela garantia da qualidade das amostras coletadas. Tais observações ganham especial destaque ao se enfocar o campo das disfonias em que, a falta de precisão na coleta de amostras poderia sinalizar irregularidades que não corresponderiam à realidade da função vocal do indivíduo disfônico.

Segundo a premissa básica do presente estudo, os achados são discutidos enquanto caracterização das várias dimensões da produção sonora, para cada falante e de sua apreciação entre os vários componentes do grupo. Dessa forma, os indivíduos disfônicos (i1 a i4) e referência (iR) não se opõem simplesmente em termos dos extremos de presença ou ausência de parâmetros vocais, porém partilham algumas características entre si, diferenciam-se em outras e assim sucessivamente.

A análise estatística por agrupamentos expõe tal particularidade, em que os aspectos de longo e curto termo revelaram a diferenciação das várias emissões para cada um dos falantes do grupo estudado (i1 a i4) e do indivíduo referência (iR). Nesse sentido, o aumento da população estudada apenas continuaria a contemplar novos agrupamentos em função das características de produção sonora presentes, porém não implicaria nova modalidade de abordagem das informações, especialmente diante das bases teóricas que respaldam esta investigação.

Referindo-se especificamente à dimensão de qualidade vocal, o grande desafio imposto refere-se à tentativa de mapear os fatores que seriam importantes à delimitação do referido espaço perceptivo em que se dispersaram as amostras do grupo estudado. Dessa maneira, a integração dos dois campos teóricos eleitos em nosso estudo, e expostos ao longo da revisão teórica, representa a possibilidade de enfocarmos as relações entre percepção e produção sonora, de forma a não nos limitarmos a um aspecto do sinal de fala registrado, conforme sugerem Colton, Estill (1981).

Em nosso estudo, a adoção do modelo fonético de qualidade vocal propiciou a condição de conhecimento da complexa relação entre qualidade vocal e segmentos fonéticos. No centro desta abordagem, integrando tais níveis, figura a unidade analítica adotada, o ajuste, servindo de base para a descrição da qualidade vocal e para a reflexão a respeito do impacto promovido pelas alterações no aparelho fonador, propiciando, inclusive, a consideração de aspectos de inteligibilidade do sinal.

Nossa opção foi conjugar as propostas dos modelos teóricos adotados, de forma que as ocorrências em termos de mecanismos da fonte e do filtro, explicitadas no modelo fonte-filtro (Fant, 1970) pudessem ser correspondidas à dimensão de ajustes recorrentes durante a fala, em planos laríngeo e supralaríngeo do

aparelho fonador, enquanto enfoque do modelo fonético de descrição da qualidade vocal (Laver, 1980), os quais revelariam a qualidade sonora detectada.

Na tentativa de desvendar a complexidade da relação da produção vocal e alteração da estrutura do aparelho fonador, o indivíduo sem queixas de disfonia, nomeado como referência, pode ser incluído no referido complexo de estímulos de produção sonora, junto ao restante do grupo de indivíduos portadores de disfonia. Em estudos da Fonética Acústica, a adoção de parâmetros de referência representa o ponto de ancoramento para que se considerem as possibilidades e limitações de análise diante de alterações da função em questão. Concebido desta maneira, não representa tentativa de normatização de função, situação esta impraticável ao se enfocar a produção vocal.

Na seqüência desta reflexão, por muitas ocasiões, os dados do indivíduo referência partilharam algumas semelhanças com aqueles de i1, i2, i3 e i4, como, por exemplo, a ocorrência de ataques vocais bruscos e instabilidades da estrutura harmônica em vogais sustentadas.

Ao retomarmos as bases teóricas que respaldam nossa análise das amostras de fala, deparamo-nos com as observações de Fant (1962b), ao atribuir à inspeção da onda acústica a maior possibilidade de oferecer esclarecimentos sobre a produção sonora, quando comparada a qualquer outra manifestação da emissão humana. Nesse sentido, a Fonética Acústica traria à tona a possibilidade de se relacionar a onda sonora a todo e qualquer aspecto observável do ato da fala. Segundo a referência original, a descrição completa da amostra de fala poderia ser alcançada somente ao preço de uma representação muito detalhada das emissões de cada falante.

No campo da Fonoaudiologia, encontramos uma limitação, no que se refere ao aprofundamento de bases teóricas para a composição de instrumentos de avaliação de qualidade vocal, traduzida em pouca tradição no uso de escalas de avaliação perceptiva ou roteiros de aplicação em maior proporção no território nacional. Tal situação gera, por muitas vezes, a necessidade de adaptação de roteiros originais em outras línguas, os quais envolvem a delicada demanda de tradução de termos, em sua maioria referentes a atributos da percepção, o que acaba por aumentar as

controvérsias na área.

A adoção do enfoque de qualidade vocal com motivação fonética utilizada como instrumento de avaliação perceptiva da qualidade vocal revelou-se da maior importância na caracterização das emissões estudadas. Apesar de dificuldades iniciais relatadas pelos juízes experientes em avaliação vocal, especialmente nos julgamentos relativos ao grupo dos ajustes supralaríngeos (articulatórios), em relação aos laríngeos (fonatórios), com os quais estão familiarizados, houve consenso entre os examinadores.

Tais achados sinalizam a possibilidade de implementação de propostas de avaliação da qualidade vocal que contemplem as variadas situações de atuação fonoaudiológica, cobrindo o território da clínica dos distúrbios vocais e da assessoria vocal. Nesta última situação, em que a alteração do sinal vocal geralmente não se faz presente, a dimensão de apreciação da qualidade vocal é, por muitas vezes, negligenciada.

Nesse ponto, na realidade brasileira, deve-se registrar a proposta de avaliação da qualidade vocal implementada como parte integrante das escalas Brandi de avaliação da voz falada (Brandi, 1996). A concepção de qualidade vocal adotada pela autora também incorporou recentemente alguns aspectos do modelo fonético adotado no presente estudo (Brandi, 2002).

No plano da investigação acústica, expandimos da tradição também vigente em nosso meio, traduzida em concentração de estudos numéricos, muitos deles sem o respaldo do devido tratamento e da inspeção prévia do sinal, para garantir a qualidade das informações obtidas.

Em nosso enfoque, a qualidade vocal é adotada como pano de fundo para tecer considerações a respeito das características inerentes do sinal sonoro. Caracteriza-se como uma tentativa de compreender as particularidades da emissão que, no futuro, poderão ser ampliadas para outras populações de falantes disfônicos, levando em conta o amplo espectro de alterações de qualidade sonora que repousam sob o manto do termo disfonia.

Nessa linha de pesquisa, a dicotomia que emerge em estudos de

grupos, cede lugar à tentativa de compreensão da disfonia em si, sem contraponto a um suposto ideal de qualidade vocal (Hollien, 2000) ou de normalidade, o qual não contempla a realidade de que a variabilidade é a marca da fala.

Na mesma linha de investigação, o refinamento da descrição da dimensão fisiológica centra-se na criteriosa descrição de eventos laríngeos. Aqueles de ordem supralaríngea não foram contemplados por dificuldades de ordem metodológica. Conseqüentemente, optamos por caracterizar a particularidade das mobilizações implementadas pelo aparelho fonador na tentativa de resgatar a condição de sonoridade, as quais pudessem revelar a magnitude do fenômeno ciclo-a-ciclo de atividade glótica. Vale ressaltar que a qualidade das informações registradas é fundamental para a apreciação dos resultados, especialmente por expressarem apenas uma parte da totalidade das mobilizações laríngeas.

Nesse sentido, as técnicas de eletroglotografia (EGG), telescopia laríngea com videoquimografia foram eleitas por possibilitarem a investigação da fala em tempo real, com registros da pulsação ciclo-a-ciclo. A necessidade de uso de várias técnicas reside nas limitações inerentes a cada uma delas. A eletroglotografia não registra exclusivamente o comportamento de pregas vocais, pode registrar contato glótico a partir de um único ponto na extensão das pregas vocais, o qual pode inclusive ocorrer por presença de secreção, estabelecendo uma via de contato entre as pregas vocais (Smith, 1981; Baken, 1987; Colton, Conture, 1990; Vieira, 1997; Orlikoff, 1998; Baken, Orlikoff, 2000; Fourcin, 2000). Por sua vez, a videoquimografia representa a análise de uma linha da imagem disponível do exame de telescopia laríngea, ou seja, o comportamento de um ponto das pregas vocais ao longo do tempo (Švec, Schutte, 1996; Tsuji, Sennes, 1998; Larsson et al, 1999, 2000; Tsuji et al, 2001).

Além disso, a menção ou estruturação de avaliação fisiológica para contemplar a noção de qualidade vocal adotada envolveria um número de técnicas e aparato incompatíveis com nossa realidade e pouco viável em termos experimentais, principalmente quando se leva em conta a quantidade de técnicas desenvolvidas nos últimos anos, e ainda em franca evolução. Grande parte dessa instrumentação para investigação da produção de fala concentra-se especialmente no campo das Ciências da Fala (Borden et al, 1994; Kent, 1997; Baken, Orlikoff, 2000; Kent, Ball, 2000).

Outro fator destacado na concepção de método refere-se à opção por modalidades menos invasivas ao aparelho fonador, excetuando-se apenas as técnicas de endoscopia laríngea, consideradas em nosso meio como procedimento de rotina no diagnóstico das disfonias. Desta constatação decorre a opção por também conjugar as observações da estroboscopia.

No enfoque da variação dos parâmetros acústicos ao longo da emissão, optamos pela inspeção da conformação geral dos traçados de espectros de longo termo e pela verificação das variações de amplitude ao longo de 128 pontos de freqüência, no intervalo entre 0-11 kHz. Nesta dimensão encontrou-se importante correlação com o referencial do modelo fonético de qualidade vocal, a qual é inclusive respaldada por estudos que aplicaram a análise de longo termo às gravações que acompanham a publicação original do modelo fonético de descrição da qualidade vocal (Nolan, 1983; Ananthapadmanabha, 1995).

Dessa forma, na análise dos traçados dos espectros de longo termo e medidas correspondentes às médias de concentração de energia por faixas de freqüência, iR apresentou declínio gradual entre os picos espectrais, com pouca concentração em freqüências altas. Tal descrição é característica de fonação modal e do equilíbrio de atividade de pregas vocais.

Em contrapartida, a observação dos espectros de longo termo das três leituras realizadas por i1 e i2 denotaram declínio abrupto de energia a partir da faixa de 1 kHz, com níveis de intensidade muito próximos nas faixas superiores. Para i3, o declínio foi acentuado entre os intervalos 0-1 kHz e 1-3 kHz, a partir do qual registra energia até 8 kHz. A falante i4, por sua vez, praticamente não revelou declínio espectral, com suave inclinação entre 0-1 kHz e 1-5 kHz e aumento de amplitude em torno dos componentes espectrais de 4 kHz, levando ao aumento da média da faixa. Há também concentração de energia entre 6,5-8 kHz, cujo declínio é apenas registrado a partir de 9 kHz, ao contrário do restante do grupo, em que esta faixa apresentou predominante declínio de energia. Além disso, i4 representou a menor amplitude registrada no grupo, valendo recordar que o máximo de amplitude registrada na faixa compreendida entre 0-2 kHz corresponde ao nível de intensidade geral do sinal (Hammarberg et al, 1980, 1986).

A descrição das concentrações espectrais por médias seguiu tendências da literatura, como tentativa de reproduzir o declínio ou inclinação observados nos traçados. Nesse campo, muitas procedimentos de cômputo de médias e de valores máximos de energia por intervalos de freqüência foram descritos, na seqüência de proposta do parâmetro alfa (α) por Frkøjær-Jensen, Prytz (1976), baseado na proporção entre nível de energia espectral acima e abaixo de 1 kHz.

As várias propostas subseqüentes envolveram a tentativa de expor tais características de declínio e inclinação espectral, passando a incorporar diferenças entre nível de intensidade máximo registrado entre 0 -1 kHz ou 2 kHz e de várias faixas de freqüência superiores, variando em seus limites máximos em média entre 8-10 kHz, atingindo o máximo de 16 kHz (Dejonckere, Cuvelier, 1983; Dejonckere, 1983, 1984; Dejonckere, Villarosa, 1986; Hammarberg et al, 1986; Löfqvist, Mandersson, 1987; Shoji et al, 1991).

Neste ponto, salientamos a particularidade da situação de alteração vocal, em contraponto aos cantores (Cleveland et al, 20001), nos quais não foi necessária a observação de faixas além de 4000 Hz, tanto para voz falada como cantada.

Conforme assinalado na revisão teórica referente aos estudos de longo termo, além da influência da atividade glótica no declínio espectral, há considerações de interferências da esfera supraglótica, em que se destacam as ressonâncias do trato vocal (Hammarberg, Gauffin, 1995; Linville, Rens, 2001; Cleveland et al, 2001; White, 2001). Pode-se então supor a possibilidade de interpretação de ajustes de ambas as esferas nas representações de longo termo, em função de ações correspondentes à fonte e ao filtro do ponto de vista acústico, tornando-se a análise mais complexa por acentuar a importância de compreensão da sobreposição de eventos referentes aos dois níveis de atividade.

Pode-se, igualmente, hipotetizar que a ineficiência do mecanismo de pregas vocais favorece a acoplagem de outras cavidades que terão influência na resposta de ressonância e inclusive na fonte, como é o caso das cavidades subglóticas. Tal aspecto será apreciado ao longo da discussão de outros padrões acústicos, no que se refere às ressonâncias subglóticas.

Apesar das restrições à apreciação de relações estreitas e do estabelecimento de leis para unir os aspectos acústicos e perceptivos da qualidade vocal, alguns estudos relevantes para o campo das alterações vocais sinalizam achados consistentes, justamente por revelarem correspondências entre achados acústicos, perceptivos e fisiológicos (Hammarberg et al, 1980; Dejonckere, 1984; Dejonckere, Vilarosa, 1986; Hirano et al, 1986; Klatt, Klatt, 1990; Shoji et al, 1991; Hillenbrand et al, 1994; Naranjo et al, 1994; Hammarberg, Gauffin, 1995; Dejonckere et al, 1996). Nesse campo, figuram as associações entre soprosidade e diminuição de fechamento glótico com menor amplitude detectada na faixa de 2-5 kHz, em relação ao primeiro pico espectral, denotando portanto declínio espectral abrupto. Tal aspecto foi detectado nos casos i1, i2 e i3, com variações de amplitude em faixas mais altas do sinal. A inspeção das amostras de fala de i4 revelou menor o declínio em comparação aos demais integrantes do grupo e indivíduo referência, porém também registrou diminuição de intensidade na referida faixa de frequência, sendo inclusive a característica que o diferenciou perante os demais falantes avaliados pela análise de agrupamentos.

Nesse sentido, os fatores de influência na composição dos agrupamentos pela análise fatorial representaram uma forma de compreender os parâmetros que distinguem os falantes em questão entre si. Neste particular, não foram encontrados estudos semelhantes para a situação de disfonias. No caso das análises de longo termo, o fator subjacente de forte de influência na composição dos agrupamentos (88,9%) referiu-se à amplitude registrada nas faixas de freqüências para emissões de i1, i2, i3, i4 e iR, de forma que i1 e i2 apresentaram aumento de amplitude nas faixas críticas relativas à sua distinção perante o grupo, enquanto os demais revelaram diminuição.

Os agrupamentos referidos revelaram a possibilidade de resgatar as três emissões de um mesmo falante, sendo que apenas i2 dividiu-se em dois grupos, os quais se mantiveram próximos. Quando excluímos iR do grupo ou fizemos o cômputo por médias de concentração em faixas de freqüência (0-1 kHz, 1-3 kHz, 3-5 kHz, 5-11 kHz), a análise fatorial resultou em sobreposição dos pontos referentes às emissões de i1 e i2. Dessa forma, foi possível detectar que, mesmo diante de instabilidades e irregularidades tão características das alterações vocais por

incompetências glóticas, há padrões recorrentes de emissão que distinguem os falantes entre si.

Nesse referencial, vale destacar que i4 representou a menor concentração espectral justamente em regiões nas quais os demais falantes apresentaram maior amplitude, coincidindo com as freqüências de ressonância. Além disso, apresentou aumento de energia a partir de 6 kHz, quando os demais apresentaram declinação, em concordância com Löfqvist, Manderson (1987) ao relatarem aumento de concentração em freqüências agudas em casos de paralisia de prega vocal com soprosidade e hipofunção. Nesse tocante, i4 diferiu substancialmente dos traçados de i1, i2 e iR, guardando certa similaridade com i3. Neste último caso, revelou-se concentração de energia aperiódica dominante, apresentando mais harmônicos em relação a i4, porém com maior grau de irregularidade em relação a i1 e i2.

Nos casos de i1 e i2 a diferenciação em longo termo pela maior concentração em faixas de freqüências mais altas sinalizou a adição de ruído no sinal acústico de indivíduos que apresentam registro de harmônicos. Dessa forma, pudemos supor que o procedimento detectou aspectos relativos à adição de ruído ao sinal de base, numa alusão aos componentes que destoam da energia harmônica presente. Tal situação pode estar relacionada ao declínio abrupto do espectro da fonte, em que harmônicos não se mantêm com amplitude detectada por muitas faixas de freqüência, sendo progressivamente substituídos por ruído. Gradativamente, de i1 a i4, pudemos detectar diminuição da energia espectral harmônica e a crescente adição de componentes não harmônicos, refletindo-se na tendência à substituição dos formantes por ruído, numa caracterização típica da esfera de correspondência ao parâmetro perceptivo de rouguidão (Yanagihara, 1967; Dejonckere, Villarosa, 1986).

Dessa maneira, i3 e i4 parecem ter se diferenciado na análise fatorial perante o grupo, e inclusive entre si, pela falta de definição de estrutura harmônica. As emissões da primeira revelaram mais harmônicos, porém mesclados às bifurcações, e a segunda praticamente não registrou harmônicos.

Diante de tais observações das características de longo termo complementadas por aspectos de dimensão de curto termo, no âmbito da

espectrografia de banda estreita, torna-se importante enfocar o julgamento perceptivo atribuído às amostras de fala, a partir do roteiro com motivação fonética inspirado no modelo de Laver (1980) e dos roteiros para sua aplicação clínica – *VPAS* (Laver et al, 1981; Beck, 1988; Laver, 1994, 2000). Tais considerações permitirão a correlação das três esferas enfocadas neste capitulo (acústica, percepção e fisiologia).

A contribuição do enfoque fonético da qualidade vocal revelou-se promissora, na medida em que na avaliação realizada inicialmente pelos juízes nas etapas 1A e 1B, praticamente não houve menção a aspectos supralaríngeos, exceto pelas atribuições de hiponasalidade e ressonância laringofaríngea. Com a aplicação do roteiro foneticamente motivado, pudemos registrar ajustes de laringe (na dimensão vertical), lábios, mandíbula, língua e faringe, revelando aspectos que não seriam salientados numa proposta tradicional de abordagem de qualidade vocal. Tal situação foi especialmente importante para a apreciação do perfil da qualidade vocal de i1, com atribuição centrada em torno de sinônimos de ajuste modal, porém nenhuma menção prévia aos ajustes de mandíbula e língua apontados nas etapas 2A e 2B.

Para os falantes i2, i3 e i4 juntamente a vários graus de mobilizações laríngeas, acrescentaram-se mobilizações supralaríngeas que ajudaram inclusive a diferenciá-los entre si. Para as amostras de fala de i2 predominaram referências a ajustes de laringe alta, voz áspera e hiperfunção. Para i3 foram apontados ajustes de constrição faríngea, escape de ar, voz áspera, falsete e hiperfunção. Para i4 predominaram julgamentos de laringe baixa, escape de ar e voz áspera. Finalmente, iR teve seu perfil de qualidade vocal traçado a partir da combinação de ajustes de labiodentalização e modal.

Neste ponto de exploração da possibilidade de corresponder os aspectos auditivos à fisiologia, Flanagan (1958) destacava a importância dos componentes de ruído originados de fluxo turbulento resultante de fechamento incompleto ou vibração irregular de pregas vocais. Quatro décadas mais tarde, Hammarberg; Gauffin (1995) salientaram a relação entre a distribuição de freqüência do ruído e os parâmetros perceptivos correspondentes, de maneira que a concentração em altas freqüências encontraria correspondência com a percepção de aspereza, e, nas faixas graves, com a rouquidão.

Nos casos avaliados, a menção aos ajustes de voz áspera e soprosa ou escape de ar e ainda combinados ao *vocal fry* (crepitância) sinalizaram a possibilidade de referência à rouquidão, em conformidade com o modelo em questão (Laver, 1980; Laver et al, 1982; Beck, 1988). Vale ressaltar que tais emissões registraram menores índices de agradabilidade, bem como maiores graus de alteração. Achado importante refere-se ao fato da predominância de ajuste de voz áspera (i3 e i4) coincidir com as maiores taxas referidas de alteração e menores de agradabilidade. Esses dois indivíduos apresentaram sinais com os menores índices de amplitude registrados.

Nesse contexto, Zwicker, Fastl (1999) postulam que a percepção da agradabilidade é influenciada por sensações como a aspereza, tonalidade, *loudness* e acuidade e depende tanto dos parâmetros físicos do som (freqüência, amplitude e duração) quanto da relação subjetiva que o ouvinte estabelece com o padrão sonoro. Estudos visando a proposição de parâmetros para a modelagem da percepção, especialmente no tocante à aspereza, continuam em franca evolução (Bergan, Titze, 2001).

No que tange a associação de padrões espectrais de longo termo da qualidade vocal, vale retomar a descrição daqueles da dimensão acústica de curto termo, como é o caso das emissões de i3, em que os harmônicos aparecem ao fundo do sinal predominante de ruído e registro de bifurcações, que chegam a se apresentar em número de três entre os harmônicos. Tais achados de bifurcações também foram encontrados em menor proporção em i2.

Tais componentes são enfocados na literatura tanto nos casos de alterações vocais (Titze, 1995; Švec et al, 1996; Cavalli, Hirson, 1999), quanto dos efeitos sonoros no canto (Fuks et al, 1998; Fuks, 1999; Bergan, Titze, 2001), destacando-se que as interpretações são conflitantes no sentido de que algumas apontam para a caracterização de aspereza enquanto outras não indicam alterações no plano da percepção.

Nesta altura de discussão, torna-se importante reforçar que, apesar de alto grau de influência de ajuste voz áspera, i4 não apresentou registro de bifurcações, dentre outras irregularidades presentes, em que se destaca o predomínio

de ruído e registro de harmônicos restritos a poucos componentes. Poderíamos então diferenciar mecanismos da ordem da irregularidade vibratória para i2 e, especialmente i3, enquanto para i4, o indício mais marcante da alteração reside na escassa ação de sonorização da coluna aérea expiratória.

No presente estudo, tais bifurcações, também descritas como subhamônicos, influenciaram a análise cepstral, com o registro de dois picos de mesma amplitude, em correspondência ao harmônico e à bifurcação para i3. Na maioria das vezes, foi registrada a bifurcação em maior intensidade que o próprio componente fundamental. Tais dimensões foram observadas nos espectrogramas de banda estreita de i3, nos quais o aspecto de irregularidade vibratória correspondeu à característica marcante de ajuste de voz áspera.

O termo bifurcação, empregado ao longo da exposição dos achados acústicos, refere-se à linguagem da dinâmica não linear, enquanto uma mudança qualitativa de comportamento (Titze, 1995; Fuks et al, 1998; Fuks, 1999). Outro termo também utilizado na esfera das não linearidades refere-se a séries de f0 e foi utilizado na descrição dos achados da eletroglotografia, dada a possibilidade de demarcar as divisões do ciclo (Fuks et al, 1998; Fuks, 1999), em que o padrão sonoro detectado equivale à multiplofonia, diferenciando-se inclusive as dimensões de diplofonia, triplofonia e quadriplofonia na descrição da qualidade sonora correspondente. A ocorrência de diplofonia foi apontada de forma esporádica pelos juízes nas etapas 2A e 2B para as emissões de i3 e i4.

O que desperta atenção do pesquisador nesta situação é a característica de relativa regularidade mantida, em intervalos de freqüência, entre os elementos que caracterizam a bifurcação, além do fato de não estarem presentes de forma constante nas emissões. No grupo estudado, foram especialmente detectadas na emissão sustentada de vogais, consideradas como não representativas da emissão habitual do falante. Tais achados conduzem-nos à indagação do impacto real das bifurcações na situação de fala encadeada e portanto nos padrões que caracterizam a qualidade vocal.

Continuando o enfoque da abordagem integrativa dos achados, encontramos correspondência entre aspectos espectrais indicativos da presença de

maior número componentes harmônicos para i1 e i2 em relação a i3 e i4, com correspondência à eficiência de mecanismos para geração do sinal laríngeo, em que a análise perceptiva revela a magnitude de desvios de ajustes em graus progressivamente maiores em relação ao ajuste modal, correspondendo à diminuição da agradabilidade e aumento de grau de alteração. Tal aspecto tem impacto direto na dimensão da queixa presente em âmbito clínico, em que passa a ser um fator relevante à situação de comunicação oral, juntamente com a inteligibilidade, a ser explorada em próximos comentários sobre as análises de curto termo.

Em nosso estudo, detectamos correspondências entre variações de amplitude ao longo das faixas de freqüência do sinal a outras representações acústicas e à dimensão perceptiva. A confiabilidade dos achados intra-falantes reforça a possibilidade de uma abordagem interpretativa no tratamento dos dados provenientes de tais modalidades de análise, justamente por revelarem ajustes compensatórios consistentes ao longo das emissões de um mesmo falante. Dessa forma, caracterizase como uma proposta promissora para o campo das disfonias, justamente por registrar os ajustes recorrentes e não situações isoladas de emissão, as quais poderiam apontar para flutuações ou instabilidades não características do perfil de qualidade vocal do falante em questão.

Em contrapartida, Ananthapadmanabha (1995), num estudo dos fatores acústicos determinantes da percepção de qualidade vocal, não referiu correspondência entre a conformação do espectro de longo termo e a qualidade vocal para amostras manipuladas por meio de procedimentos de síntese de fala. Vale ressaltar que considerou apenas a dimensão de conformação do gráfico.

Ao atingirmos tal nível de discussão a respeito dos mecanismos subjacentes às irregularidades vibratórias apontadas, deparamo-nos com a necessidade de apreciação de componentes da dimensão fisiológica da produção sonora, representados pelas informações da eletroglotografia e do exame laringológico. De forma sintética, descreveríamos que a regularidade vibratória manifesta por iR, no caso da eletroglotografia, evolui gradativamente da adição de pequenas irregularidades, como i1, para a situação de seu aumento, i2 e o ponto em que o mesmo passa a predominar no sinal, i3 até atingir a situação em que atividade vibratória praticamente não se sustenta, i4. Nesse campo, podemos descrever detalhes

de atividade glótica na confluência de achados da eletroglotografia e da videoquimografia para a compreensão do fenômeno ciclo-a-ciclo que representa o substrato fisiológico das distribuições de energia anteriormente referidas.

Para cada um dos falantes, diferenciaram-se registros de atividade laríngea à eletroglotografia, ressaltando sua validade para a situação em questão, mesmo diante de referências às limitações para seu uso em mulheres e em casos de paralisias de pregas vocais (Baken, 1987; Colton, Conture, 1990; Orlikoff, 1998; Baken, Orlikoff, 2000). Certamente, a adoção de critérios de inspeção do traçado da onda Lx, selecionando aquelas que podem fornecer estimativas de f0, perturbação, coeficiente de contato e índice de velocidade foram essenciais para a adequada interpretação de achados, numa importante correspondência aos dados laringológicos, especialmente aos índices de atividade vibratória, representados pelos parâmetros de estroboscopia e videoquimografia.

Vale ressaltar que os procedimentos de filtragem aos quais as ondas registradas foram submetidas (Vieira et al, 1996a; Vieira, 1997) garantiram tais condições de análise no âmbito da eletroglotografia, novamente reforçando a importância de procedimentos de registro e tratamento das amostras para o âmbito de análise acústica e eletroglotográfica. Na abordagem apresentada, primou-se pela qualidade dos registros, de forma que a leitura da onda eletroglotográfica passou a ser dado relevante para análise, inclusive a sinalização da ausência de contato ao longo de tempo ou de suas flutuações (Vieira, 1997; Orlikoff, 1998; Baken, Orlikoff, 2000; Fourcin, 2000). Sua contribuição estende-se inclusive para a possibilidade de confirmação de medidas correlatas na esfera acústica, como é caso do VOT em que houve consenso entre medidas extraídas pelas duas técnicas, especialmente nas condições de avaliação semi-automática situadas entre 3 e 4.

Nesse universo de análise, partimos para a comparação de dados de ordem fisiológica, com ajustes anteriormente delineados na esfera perceptiva e de seus correlatos acústicos. Nesse plano, iR apresentou onda Lx de maior amplitude, delimitando fases do ciclo vibratório compatíveis com ajuste modal, regularidade na manutenção de intervalo entre os ciclos, bem como de sua amplitude. As falantes i1, i2 e i4, com quadro de paralisia unilateral de prega vocal em posição paramediana, revelaram padrões variáveis a serem considerados a seguir.

O traçado da onda Lx de i1 correspondeu à escassa mobilização de borda inferior de prega vocal e assimetria de fase de vibração, traduzidas em valores de índice de velocidade compatíveis com fechamento lento de pregas vocais, medidas de perturbação aumentadas, além de diminuição de coeficiente de contato.

O estudo dos registros de i2 revelou achados de fechamento glótico completo, apesar de sugestivo desnível de pregas vocais pela assimetria de movimento de aritenóides e discreta assimetria de fase de atividade glótica, além de mobilização supraglótica, representada por constrição medial e ântero-posterior de vestíbulo, com sugestivo aumento da fase de subida de contato e valores de índice de velocidade igualmente compatíveis com a diminuição deste nível de atividade glótica. A atividade vibratória revelou discreta assimetria de fase, vibração periódica com presença de onda mucosa bilateralmente e diminuída. As medidas de perturbação encontraram-se elevadas, compatíveis com as alterações descritas, enquanto o coeficiente de contato revelou valores próximos a iR. Entretanto as medidas de coeficiente de contato extraídas manualmente da forma de onda revelaram valores diminuídos, provavelmente pelo fator de perturbação do sinal e séries de f0 registrados (Švec et al, 1996).

Apesar de relativa mobilização para fechamento glótico evidenciada nos exames laringológicos, em grau parcial em i1 e completo em i2, o detalhamento da onda Lx revelou particularidades dos mecanismos compensatórios, de forma que i1 e i2 apresentaram similaridades que levaram ao seu agrupamento em algumas das modalidades de análise fatorial. Apesar de certa distância em termos de julgamentos perceptivos de qualidade vocal, nas esferas laríngea e supralaríngea, revelaram padrões acústicos e fisiológicos com várias similaridades. Tais aspectos contribuem para o campo das discussões das complexas relações envolvendo a percepção de qualidade vocal e os fatores subjacentes a tais julgamentos, exemplificando a dificuldade em se eleger um parâmetro para investigar tal provável nível de implicação.

A partir dos registros de i3, não foi possível realizar a extração automática de medidas, devido às irregularidades e diminuição da amplitude das ondas. Tais eventos foram compatíveis com achados de exame laringoscópico de fechamento incompleto à fonação, diminuição de amplitude vibratória, diferença de fase de vibração entre as pregas vocais e inclusive em diferentes porções da mesma (esquerda). A onda

mucosa mostrou-se diminuída bilateralmente. Medidas de extração manual revelaram coeficiente de contato diminuído e valores acústicos de subida de contato ligeiramente aumentado em relação a iR, denotando maiores instabilidades de controle da atividade glótica.

Na situação extrema do grupo, i4 apresentou trechos de emissão em que onda Lx não chegou a ser registrada. Os achados laringoscópicos revelaram fechamento incompleto à fonação, com onda mucosa oscilando entre ausente e diminuída bilateralmente. As instabilidades dos traçados Lx não permitiram a extração de valores manuais de coeficiente de contato. Em termos acústicos, revelou valores aumentados de tempo de subida de contato em relação ao restante do grupo e ao iR, denotando a dificuldade de manutenção de padrão vibratório.

Pode-se deduzir que i3 e i4 apresentaram as maiores alterações de mecanismos glóticos e, igualmente, maior tendência a falhas de contato de pregas vocais. As falantes i1 e i2 revelaram inconsistências nos mecanismos glóticos, os quais certamente influenciaram os resultados sonoros traduzidos nos vários aspectos acústicos anteriormente salientados.

Em suma, os achados de harmônicos e suas concentrações na emissão por faixas de freqüências encontrou correspondência com várias representações do declínio da energia espectral, conhecida como inclinação ou declínio espectral na dimensão de longo termo e de atividade laríngea. As falantes i1 e i2 revelaram condições de fechamento glótico e índices de vibração que garantem alguma condição de sonoridade à emissão.

A falante i3, por sua vez, apesar de não apresentar paralisia de prega vocal, denota aumento dos componentes de rigidez e fechamento incompleto que colabora para a diminuição de amplitude espectral e prejuízo do padrão de harmônicos. Há dificuldades na detecção de onda mucosa, enquanto i4 representa a situação em que o escape aéreo supera a atividade vibratória, com escasso registro de onda mucosa, levando-a a apresentar diminuição de energia exatamente nas faixas nas quais os demais falantes apresentavam picos, numa possível alusão à participação da cavidade subglótica na emissão.

Nesse aspecto, detecta-se provável influência da alteração da fonte glótica e conseqüentemente de todos os outros registros que dependam de tais contornos, como é o caso de análises pelo processamento pelos algoritmos *LPC* e *FFT*. Tais procedimentos de análise também revelam os efeitos da interação das cavidades subglóticas, as quais, em contrapartida àquelas da esfera supraglótica, apresentam dimensões fixas, com efeitos correspondentes presumíveis (Stevens, 1998).

Nesse âmbito, novamente o parâmetro de declínio espectral ganha destaque, de forma a influenciar a variação de amplitude do sinal em suas variantes de análise. Tal aspecto coloca em evidência o grau de inter-relação entre eventos da fonte e do filtro, reforçada por vários autores ao revisarem os princípios da teoria acústica (Borden et al, 1984; Kent, Read, 1992; Ladefoged, 1996; Kent, 1997).

Detectamos a tendência de manutenção de amplitude em regiões correspondentes a F1 e F2, levando inclusive à dificuldade de distinção do traçado dos formantes nos espectrogramas, a qual poderia ser influenciada por mecanismos de aumento de banda de F1 (Klatt, Klatt, 1990; Hanson, Stevens, 1995). Além disso, as estimativas espectrais de curto e de longo termo abordadas por análise fatorial revelaram a diferenciação de caso de maior comprometimento de mecanismos de fechamento glótico (i4) em função da diminuição de energia em faixas espectrais que correspondem justamente às duas primeiras supostas ressonâncias subglóticas.

Nesse ponto, especial dificuldade foi registrada para a detecção dos picos de ressonância por processamento *LPC*, embora tenha sido observada a adequação de ordem do filtro à freqüência de amostragem e tenham sido introduzidas variações nas modalidades de análise, incluindo a redução da freqüência de amostragem (11050 Hz) e do limite superior de freqüência nas análises (5 kHz), conforme proposta de Vieira (1997). A extração de medidas relativas à estrutura formântica foi complexa, no sentido que as características de ruído e as possibilidades delineadas de adição de picos extras (formantes) ou de antirressonâncias (Klatt, Klatt, 1990; Cranen, Boves, 1991; Hanson, Stevens, 1995; Stevens, 1998) colaboraram para este aspecto, denotando necessidade de cuidadoso delineamento de método para abordagem da estrutura formântica.

Dadas as características descritas, optamos por não extrair os valores de sua banda, mantendo a abordagem do padrão de formantes, enquanto freqüência e amplitude dos picos de ressonância. Quanto à influência dos eventos de ordem subglótica, também são destacados os achados de descontinuidade ou rupturas na trajetória dos formantes, novamente interferindo nas condições de análise da amostras, especialmente visíveis nos casos de i3 e i4.

Tal aspecto de dificuldade de obtenção de estimativas confiáveis por análise pelo algoritmo *LPC*, em que o traçado isolado do envelope espectral não demarcou todas as ressonâncias para a maioria dos casos estudados, requereu o uso auxiliar do algoritmo *FFT* para demarcação das freqüências centrais dos formantes.

Outros parâmetros neste âmbito envolvem as diferenças de amplitude entre primeiro e segundo harmônicos, reforçadas como correlatos fisiológicos do fechamento de pregas vocais, no sentido de que seu aumento seria relacionado à condição de soprosidade (Holmberg et al, 1995; Mendonza et al, 1996; Decoster, Debruyne, 1997). Tal situação é referida como achado comum em vozes femininas, especialmente em estudos baseados em comparações de vozes entre os sexos (Klatt, Klatt, 1990; Decoster, Debruyne, 1997), destacando-se também o ruído na altura de F3, novamente para a população feminina.

A análise por FFT revelou valores de f0 mais estáveis (intra-falantes) entre as várias emissões analisadas para i1, i2, i3 i4 e iR no que tange à emissão encadeada, compatíveis com eletroglotografia e demais modalidades de análise acústica.

Nos casos estudados, i1 apresentou níveis de intensidade de H1 e H2 próximos, enquanto os demais tenderam a apresentar valores maiores de intensidade para H1, sendo a diferença maior encontrada em i4 (H1 média de 32 dB e H2 média de 23 dB), justamente o falante com as maiores alterações em termos de fechamento glótico.

À procura de estabelecimento de correlatos acústicos com a velocidade de fechamento de pregas vocais, o trabalho de Holmberg et al (1995) investiga, além da intensidade relativa dos harmônicos, a intensidade relativa dos

formantes em comparações como H1-F1, H1-F3 e F1-F3. Tais estimativas não foram exploradas no presente estudo devido às dificuldades expostas por ocasião do processamento das amostras e pelo fato de que alguns trabalhos estabelecem relações de tais estimativas com medidas de fluxo, não utilizadas no presente estudo.

As medidas baseadas em diferenças de declínio H1- H2 poderiam figurar com uma alternativa para os casos em que não fosse possível extrair os índices de velocidade correspondentes, justamente pelas perturbações da onda. Novamente o grande problema é o fato de não ser a amplitude de tais componentes facilmente aferida.

Nesse âmbito de detalhamento da atividade glótica, encontramos correspondência a vários achados, em que efeitos da fenda glótica corresponderam a irregularidades na fase de fechamento (Båvegård, Fant, 1995), como no caso de i1 e i2, detectados na eletroglotografia como deslocamento dos picos da onda Lx e aumento do índice de velocidade.

Outra consideração a respeito da interferência da acoplagem subglótica refere-se ao escape de ar não sonorizado ao início e/ou final das emissões para i2, i3 e i4, cuja concentração registrou-se por volta de 1500 Hz, justamente aquela região espectral descrita com ganho de amplitude nesta situação, por provável influência de cavidades subglóticas. Stevens (1998) também salientou o grau de influência da região subglótica em função do tamanho da fenda glótica. Em nossa casuística, os achados mais consistentes foram encontrados em maior proporção em i3 e i4, porém em menores dimensões também em i1 e i2.

Tais achados denotam o avanço em relação à construção do conhecimento sobre os mecanismos compensatórios e que poderá, inclusive, vir a causar impacto clínico, especialmente por considerar um sistema que gera muitas polêmicas em termos do campo da reabilitação das disfonias: a esfera respiratória e a modalidade de enfoque da dinâmica respiratória.

Desconsiderando-se as divergências de abordagem clínica, a maior parte dos fonoaudiólogos reconhece a magnitude dos eventos de ordem aerodinâmica, porém carece de compreensão mais detalhada desse nível. Como exemplo, podemos

citar as tentativas de obtenção de estimativas mais precisas de valores de freqüência, amplitude e banda de ressonâncias subglóticas, os quais possivelmente serão estendidos a populações com alterações vocais em futuro próximo (Spencer, Titze, 2001).

Com o avanço da tecnologia e do refinamento das análises, surgem as possibilidades tidas como as estimativas acústicas, as quais buscam a leitura de parte das informações anteriormente delineadas. Circunscritas no âmbito de curto termo, revelam-se como a base de muitas abordagens de medidas acústicas na aplicação ao estudo da voz, destacando-se análise por *FFT*, *cepstrum* e *LPC* (Feijoo, Hernandez, 1990; De Krom, 1993; Dejonckere, Wienecke, 1994; Dejonckere, 1998; Awan, 2002).

No âmbito de medidas acústicas aplicadas aos estudos das disfonias, torna-se inevitável a menção às referidas medidas de perturbação, conduzindo-nos ao ponto em que, justamente o parâmetro a ser aferido é aquele que sofre maior influência do nível de irregularidade que ele mesmo representa, numa alusão ao quão paradoxais podem se tornar as medidas acústicas de perturbação (Bielamowicz et al, 1996).

Tais constatações revelam o grande fator limitante na avaliação acústica e mesmo eletroglotográfica voltada à extração de medidas: a dificuldade em usar parâmetros cujos processos de medição pressupõem a regularidade do sinal, quando essa é justamente a primeira dimensão afetada na disfonia subseqüente a comprometimentos de mobilização glótica. Neste trabalho, garantidos os aspectos de tratamento e aplicação de procedimentos de análise anteriormente detalhados, os valores de *jitter* e *shimmer*, mostraram-se aumentados em proporção à indicação de manifestações de alterações de mecanismos laríngeos, com base em estudo de Vieira (1997).

Neste ponto da discussão, retornamos a questão realçada no início deste trabalho, em que a formalização de medidas depende da cuidadosa captura do sinal, sua prévia inspeção e aplicação de procedimentos cuidadosamente delineados para tal finalidade. Deve-se ressaltar que a ausência de registro ou impossibilidade de

extração de valores é um dado confiável, pois indica que a atividade de pregas vocais está perturbada.

Torna-se importante estender nossas observações a outras medidas disponíveis, como é o caso das estimativas relativas à análise cepstral, propostas como alternativas aos algoritmos de demarcação de ciclos. Sua inspeção é referida na avaliação de alterações vocais, atribuindo-lhe a possibilidade de traçar estimativas da estrutura harmônica do sinal (Dejonckere, Wienecke, 1994; Heman-Ackah et al, 2002) e de extrair de medidas de curto termo, incluindo os valores de f0 e outros dele dependentes, tais como as medidas de perturbação, ruído e formantes (De Krom, 1993; Dejonckere, Wienecke, 1994; Dejonckere, 1998; Awan, 2002).

Em nosso estudo, apesar de i1, i2, i4 e iR revelarem concordância entre as estimativas de f0 por *cepstrum* e por outras técnicas, os valores de f0 obtidos para i3 foram discrepantes. Ao buscarmos pelo fator de interferência, encontramos nos espectrogramas de banda estreita uma bifurcação presente na região abaixo de f0, de intensidade fraca, porém ao inspecionarmos o traçado cepstral, constatamos a existência de nível de amplitude alto, indicando ser fator recorrente no tempo.

Vale recordar que as dimensões de amplitude são distintas para ambas as representações acústicas. Na representação cepstral, a amplitude está relacionada à recorrência do período ao longo de várias oscilações. Dessa forma o pico cepstral de maior amplitude não correspondeu ao fundamental detectado nas demais modalidades de análise, mas ao subharmônico representado como a bifurcação na altura de f0 e que expreessa exatamente a metade do valor de f0 estimado por outras modalidades de análise. Para tal falante, ao adotarmos o valor de T0 correspondente ao segundo pico cepstral em termos de nível de amplitude, o valor de f0 correspondente coincidiu com os achados de outras modalidades de extração.

Este tipo de consideração revela a limitação do método em questão para a extração da estimativas de f0 e medidas dela dependentes em situação de perturbação da estrutura harmônica. Além disso, acentua o cuidado em sua adoção na rotina de avaliação, inclusive pela necessidade de cuidadosa descrição dos aspectos acústicos e eletroglotográficos ao se implementar medidas de respectiva ordem. Entretanto os traçados cepstrais revelaram a importante informação da atividade

harmônica e a possibilidade de se estabelecer correspondência aos aspectos de estrutura harmônica e padrões de longo termo delineados anteriormente nesta discussão.

Avançando para além do nível de descrição dos eventos glóticos e da influência daqueles da esfera subglótica, partimos para a busca de indicativos que pudessem revelar a dimensão de mobilizações da ordem dos articuladores e ressoadores. Como primeira menção, retomamos a abordagem dos formantes da emissão, em que, superadas as dificuldades na sua extração, os falantes avaliados, apresentaram valores de freqüência de F1 e F2 próximos entre si e dentro dos limites previstos para a vogal [a] no português e mantiveram suas diferenças estáveis entre suas próprias repetições de emissões, garantindo a manutenção de aspectos de informações relativas ao plano do segmento. A maior variabilidade interfalantes ocorreu para freqüências de F3 e, especialmente F4, a qual sinaliza a interferência de alguns ajustes individuais.

Diante de tal possibilidade de enfoque de emissões em padrão encadeado, intentamos a busca de parâmetros que continuassem a revelar a relação e a implicação entre esferas de atividade relativas à fonte e ao filtro. Nesse âmbito, encontramos respaldo para a abordagem de *VOT*, na medida em que, além de contemplar os aspectos de interação fonte-filtro, permite uma incursão ao campo segmental, de forma a refletir inclusive a influência dos ajustes na produção da fala. Destacamos as medidas de *VOT* obtidas por meios acústicos e inclusive confirmadas por eletroglotografia.

O *VOT* destacou-se como índice promissor para compreensão das limitações impostas pela disfonia. Curiosamente o referido índice não é explorado no campo da voz, talvez por não contemplar somente aspectos da dinâmica laríngea.

A apreciação de valores obtidos para emissões de sete sílabas ([pa] [pε] [bε] [tar] [da] [ka] [da]) em três repetições para cada falante e sua distribuição para o grupo dos sons consonantais surdos e sonoros revelaram que a dinâmica de acoplamento de ações glóticas e supraglóticas no tempo perturba-se de forma mais importante para sons posteriores nos casos de maiores alterações de atividade laríngea (i3 e i4). Para i1 e i2 as alterações ocorreram na esfera da diferenciação entre

pontos bilabiais e dentais/alveolares, portanto em relação ao ponto de articulação das consoantes, as quais encontram correspondência com a dimensão de inteligibilidade do sinal (com significância estatística).

Na emissão de sons surdos, a maioria dos indivíduos do grupo estudado (i1 a i4) não separou nitidamente os pontos de articulação a partir das médias de VOT em três emissões, situação em que se esperariam valores de [p] e [t] mais próximos, porém distintos, e menores em relação a [k]. Nesse âmbito, as emissões de i1 revelaram valores de VOT de [t] menores que de [p], com inversão das faixas, (estatisticamente significativos) das gerando erros de percepção sílabas correspondentes aos sons consonantais anteriores [p] e [t]. Para i2, os valores de [t] e [k] foram próximos e, novamente, geraram erros em esfera perceptiva. O indivíduo i3 revelou inconsistências nos julgamentos de sons e valores de VOT. O indivíduo i4 revelou valores registrados de VOT muito próximos para os três pontos articulatórios, não delimitando fatores característicos da língua, com relação aos critérios de classificação articulatória dos sons da fala. Erros na esfera de inteligibilidade foram gerados para os sons anteriores, tais como [t] apontado como [d].

Em termos de estimativas do grupo, os valores de VOT para [p] e [t] praticamente não se diferenciaram no grupo e a inteligibilidade para tais segmentos fonéticos esteve comprometida na maioria dos casos.

Para o grupo de sons consonantais plosivos sonoros, outras ocorrências peculiares foram detectadas no grupo estudado. Em emissões de i1 e i2 houve mínima dispersão de valores com relação aos pontos de articulação, quase uma repetição, enquanto i3 e i4 apresentaram-na em níveis extremos, a maior parte deles acima de zero. iR demonstrou certa variabilidade, mas dentro de uma faixa que não caracteriza a dispersão, nem concentração extremas. Por ser influenciado por aspectos suprassegmentais (Shimizu, 1996; Higgins et al, 1998), espera-se variabilidade nos valores, especialmente para a situação de fala encadeada, em comparação à emissão de sílabas isoladas ou daquelas inseridas em sentença-veículo. Nesse sentido, podemos supor que valores muito similares ou muito dispersos para uma mesmo falante denotam alguma dificuldade em relação à sincronização entre o gesto glótico e o gesto articulatório nos falantes do grupo estudado.

Outro aspecto a ser destacado é a detecção de valores positivos para [g] de I3 e I4, além de falhas de inteligibilidade no contraste surdo/sonoro para i4 em todos os pontos articulatórios estudados (bilabial, dental/alveolar e velar), com significância estatística. Tais falantes revelam, do ponto de vista fisiológico, falha de fechamento glótico e diminuição das condições vibratórias, traduzindo-se em onda mucosa diminuída (i3) ou praticamente ausente (i4), em contraposição a i1 e i2 que apresentaram menor alteração laríngea, com fechamento completo, presença de onda mucosa e contraste surdo/ sonoro preservados no julgamento auditivo. Tais achados reforçam dados de estudo anterior conduzido no LIAAC-PUC/SP, em que os maiores desvios de valores de *VOT* para plosivos sonoros na emissão de falantes submetidos a laringectomias parciais verticais ocorreram para os casos em que se detectou ausência de onda mucosa (Camargo, 1998).

Na consideração dos erros gerados para cada falante, optamos por estabelecer associações aos ajustes referentes à qualidade vocal e que pudessem ter influenciado o plano segmental, novamente com valores estatisticamente significativos.

Neste sentido, i1 revelou mobilizações de longo termo com tendência a fechamento de mandíbula e retração de corpo de língua, cujas emissões geraram muitos erros para sons bilabiais, apontados como labiodentais e dentais/alveolares. A falante i2, com mobilização característica de laringe alta, hiperfunção e voz áspera apresentou valores de VOT para sons sonoros na mesma faixa de iR, devendo-se recordar que a laringe em posição alta tende a favorecer a força adutora, condição primordial para desencadear a sonorização neste contexto, de forma que não promoveu erros no julgamento dos contrastes surdo/sonoro. Outro dado interessante refere-se ao fato de apresentar os melhores índices de julgamento de inteligibilidade, mesmo com maior grau de disfonia atribuído e de ajuste predominante em laringe referente à voz áspera, a qual denota irregularidades vibratórias. i3, dentre vários ajustes, apresenta tendências laríngeas de escape de ar, voz áspera e supralaríngea de constrição faríngea, denotando falha de sonoridade em [g], justamente aquele segmento com ponto de articulação mais próximo à região de predomínio de alterações do aparelho fonador e, consegüentemente, de caracterização de mobilização compensatória de ajustes.

Na falante i4, os ajustes de laringe baixa e escape de ar sugerem a diminuição da força adutora glótica, gerando os maiores erros em termos de contraste de sonoridade, em que os fones sonoros, em seus vários pontos de articulação foram referidos como surdos pelos ouvintes, com maior significância estatística para aqueles posteriores. i4, acima da faixa de 60 anos, apresentou similarmente ao estudo de Decoster, Debruyne (1996) os maiores valores de *VOT* para [p] e [t] no grupo. Nessa caso, há possibilidade de aumento de *VOT* para sons surdos em função da idade. Uma hipótese para tal referência poderia residir em achados de alongamento do trato vocal feminino com o avanço da idade, levando a laringe a ocupar posição mais baixa, a referida *ptose* laríngea (Linville, Rens, 2001), a qual poderia induzir a menor força adutora e favorecer maior duração para implementar a coaptação glótica necessária à sonorização. Tais hipóteses, ainda necessitam de cuidadosa investigação e são traçadas como considerações preliminares no enfoque do *VOT* na situação de disfonia.

Tais achados sinalizam a possibilidade de se assumir a influência do(s) ajuste(s) sobre os segmentos e portanto da qualidade vocal ter influenciado a inteligibilidade de fala, especialmente para populações com alterações em aparelho fonador, como aquelas portadoras de limitações da ação glótica. Afirmações deste porte caracterizam-se como um provável indicativo de supremacia do ajuste em relação ao segmento, ao contrário das observações de Laver (1980) e em concordância à opinião expressa por Albano (1988), dimensionando o impacto das alterações de qualidade vocal na comunicação oral.

Outro aspecto a incrementar tal argumentação refere-se a achados de estudos anteriores em que foi possível destacar alterações de inteligibilidade para vogais em contexto silábico, com confusões de julgamento geradas para os contrastes [e] x [ɛ] e [o] x [ɔ] em falantes submetidos às laringectomias parciais verticais (Camargo, 1996). Tais alterações do plano da inteligibilidade foram atribuídas à perda dos ajustes finos que regulam as pequenas variações de conformação de cavidades supraglóticas e suas relações com atividade glótica. Novamente, numa nova indicação da implicação dos eventos da fonte e do filtro para casos de alterações vocais. No presente estudo, mantendo a linha de pesquisa anteriormente adotada, partimos para a investigação da inteligibilidade das consoantes e novamente encontramos implicações entre as esferas glótica e supraglótica. Neste aspecto somam-se as influências da esfera respiratória,

no sentido de que também podem refletir aspectos relativos ao volume pulmonar. Nesse sentido, muitos achados sinalizam a alteração de coordenação com esta esfera, ampliando a noção de que o *VOT* poderia inclusive refletir a esfera de coordenação de eventos subglóticos, glóticos e supraglóticos.

Nesse sentido, os dados apreciados permitem-nos apontar que a maior dificuldade parece residir na sincronização dos gestos glótico e oral. Os grandes desvios de *VOT* para sons surdos, especialmente quando comparados a dados do indivíduo referência e de falantes de português (Madureira et al, 2002), revelam a dificuldade em se iniciar o processo de sonorização com refinamento, dados estes confirmados pela descrição fisiológica de atividade glótica. Os dados do referido estudo também foram adotados, junto com aqueles de iR, como valores de referência para o português.

Outro achado interessante refere-se à identificação de valores de VOT zero para as sílabas [pɛ][bɛ] e [da] de i4, ocorrência pouco usual para o português e ainda para situação de sons surdos. Todos os sons foram referidos pelos ouvintes como surdos, levando-nos a estabelecer correspondência à dimensão da alteração de ordem vibratória.

A importância do índice de tempo de início de sonorização na Fonoaudiologia reside na possibilidade de enfocar algumas das complexas relações anteriormente expostas, as quais não foram devidamente aplicadas aos estudos das disfonias. Nesse âmbito, pudemos encontrar relatos de aplicação a outras situações (Gandour, Dardarananda, 1984; Ryalls et al, 1999; Madureira et al, 2002). Alguns estudos prévios desenvolvidos dentro da mesma linha de pesquisa no LIAAC-PUC/SP, encontraram importante correspondência para a situação de disfonia (Camargo, 1998; Gregio, 2000), inclusive em associação com disfagia (Andrade, 2000), porém não localizamos enfoque semelhante na literatura. Do ponto de vista de métodos, devemos citar Higgins et al (1998), associando técnica de eletroglotografia ao estudo de fluxo oral para extração das medidas, porém não para a situação de disfonias.

Estudos que podem colaborar para a futuro abordagem deste índice na situação de disfonia referem-se àqueles relacionados às interferências da idade (Decoster, Debruyne, 1997) e do sexo dos falantes, e, especialmente, de influência de mecanismo respiratórios, como é caso do aumento do volume pulmonar, com efeitos nos movimentos laríngeos, resultando num mecanismo conhecido como *tracheal pull* (Hoit et al, 1993; Higgins et al, 1998). Tais esferas não foram devidamente associadas no estudo da população portadora de alterações vocais, sinalizando um futuro promissor, especialmente diante das variadas adaptações de dinâmica respiratória à fonação e de adaptações como aquelas descritas neste estudo, especialmente com relação à fonação inspiratória, à sua duração, em algumas situações próximas ao do trecho sonorizado e ao escape de ar não sonorizado. Somam-se a estes comentários, os achados de comprometimento de inteligibilidade, em proporções equivalentes ao comprometimento dos valores de *VOT* e também valores de tempo de subida de emissão, revelando que o fechamento atingido ainda não é plenamente suficiente e atividade vibratória também está comprometida, como indicado pela diminuição ou ausência de onda mucosa.

Diante das observações delineadas ao longo das páginas anteriores, podemos afirmar que nossa contribuição ao diagnóstico e reabilitação de disfonias reside na possibilidade de superarmos o limite do julgamento perceptivo baseado apenas em escalas de nível de alteração vocal, especialmente centradas em mobilizações da região glótica, para ajustes de longo termo e, portanto, recorrentes na produção sonora, tanto em nível glótico quanto supraglótico, segundo proposta de modelo fonético (Laver, 1980). Nossa abordagem de avaliação perceptiva do sinal gerado incluiu o avanço ao estudo da inteligibilidade de fala e portanto às relações entre ajustes e segmentos. Tal nível de correlação também pode ser estabelecido a partir do ponto em que a qualidade vocal emerge como uma das propriedades prosódicas que têm o papel fundamental de veicular sentido.

Diante da consistência dos ajustes desenvolvidos pelos indivíduos portadores de alterações vocais estudados, resta-nos a tarefa de investigar aspectos realmente indicativos da dimensão sonora, sugerindo-se a investigação por análise acústica em modalidades de análise de longo termo ede curto termo, com destaque para estrutura harmônica e formântica e *VOT*, além da possibilidade de complementação por análise eletroglotográfica, com inspeção da onda Lx e indicativos de contato e de velocidade de fechamento de pregas vocais.

A partir das implicações descritas entre os vários indicativos de produção e de percepção do sinal de fala, resta-nos o desafio de relacionar as dimensões de curto e longo termo, procurando depreender os aspectos invariantes da emissão e o caráter das irregularidades detectadas. Tal seria o ponto de partida para futuramente abordarmos as flutuações que se estabelecem no sinal. Entendendo-se esta complexa sobreposição, das flutuações sobre os ajustes de base, teríamos a condição de realizar uma incursão mais apropriada ao campo das disfonias, no sentido de procurar reconhecer as características de adaptação, enquanto compensações, que tendem a se manter ao longo do fluxo da fala.

Nessa abordagem, diríamos que a evolução dos estudos em nossa área ocorreu no sentido inverso, caracterizando-se, em primeira instância, as variações momentâneas, sem que se tivesse a noção mais precisa das mobilizações de base, das quais os falantes partem para sobrepor as oscilações. Deve-se inclusive recordar que em alguns falantes a tão referida variabilidade foi substituída por uma dimensão de equivalência e igualdade que desperta a atenção, especificamente no caso dos valores de VOT para sons sonoros. Tal achado revela que há adaptações que apresentam um caráter oposto ao da tradicional descrição de irregularidade, que também denota alteração no mecanismo natural do fluxo da fala, na qual certas variações são esperadas.

Além disso, um ponto que merece destaque nesta etapa de articulação dos achados refere-se especialmente à almejada correspondência das esferas de produção e percepção. Nesse campo, a abordagem da qualidade vocal adotada, do ponto de vista fonético, mais especificamente do modelo fonético de descrição da qualidade vocal, propiciou o detalhamento de sua avaliação auditiva, superando a limitação de julgamentos com base apenas em graus de alteração glótica, nos quais ajustes apontados em porções supraglóticas do aparelho fonador não seriam contemplados.

Além disso, conforme assinalado no capítulo referente à estruturação do roteiro de avaliação adotado, utilizamos alguns tópicos relativos à descrição de escalas como *GRBAS* e *SVEA* que guardam semelhanças com os ajustes na proposta de Laver (1980). Tal proximidade é inclusive referida em alguns estudos, como os de Hammarberg et al (1986) ao traçarem correspondência entre ajuste originalmente

designado como *whispery voice* no *VPAS* (Laver, 1980) com o parâmetro perceptivo de voz soprosa presente na proposta *SVEA*.

Os vários achados de qualidade vocal apresentados revelam a dificuldade de adoção de procedimento de avaliação que não contemple o conjunto de mobilizações supralaríngeas presentes em todos os indivíduos avaliados. Tais ocorrências foram fundamentais para articulação dos achados e para traçarmos hipóteses sobre a implicação dos componentes de atividade glótica (laríngeos em sentido mais amplo) e de região supraglótica. A implicação entre tais níveis de produção sonora foi especialmente revelada em avaliação de longo termo, enquanto indicativos de declínio espectral, além de estimativas de amplitude de harmônicos e índice de velocidade e, finalmente, de apreciação dos índices de *VOT*. Como um exemplo dessa dimensão, podemos citar achados de i1, em que apesar de revelar ajuste modal em plano glótico, apresentou mobilizações supraglóticas que culminaram em dificuldades de distinção entre sons consonantais anteriores, mais especificamente bilabiais e dentais/alveolares.

Em relação a i1, a análise mais detalhada de seus dados sinaliza que o alto grau de agradabilidade atribuído à sua qualidade vocal ocorreu independentemente de ajustes que prejudicaram a inteligibilidade de sua fala (teste de avaliação da inteligibilidade das sílabas) e às condições fisiológicas reveladas, na medida em que, apesar do fechamento completo detectado, há alterações no ciclo referentes ao impacto da paralisia de prega vocal. Nesse respeito, destacamos que a refinada análise dos mecanismos laríngeos, especialmente no que se refere ao ciclo vibratório, detalhado pela eletroglotografia e pela videoquimografia, pode nos ajudar a dimensionar o impacto de algumas alterações na qualidade final do sinal sonoro. Tal possibilidade pode nos conduzir à investigação e à qualificação de compensações realmente provedoras de melhores condições para a comunicação oral.

Tal achado também desperta a atenção no sentido de que não podemos traçar uma relação biunívoca entre um ajuste no plano fisiológico e uma determinada qualidade sonora. Deve existir um limiar acústico que as mobilizações atinjam para causar o impacto na percepção auditiva. Tal reflexão somente acentua a necessidade de abordagem integrada das vertentes acústica, perceptiva e fisiológica do sinal sonoro.

Neste caso, especificamente utilizando como base desta argumentação i1, a qualidade vocal não foi descrita como desagradável ou alterada, porém as sincronizações glótica e oral (supraglótica) mostram-se claramente comprometidas. O grau mais leve de alteração e de maior agradabilidade não exclui alteração de fala, especialmente no âmbito da inteligibilidade, de forma que classificar graus de alteração baseados especialmente na aperiodicidade do sinal sonoro não necessariamente revela aquelas propriedades de maior proeminência à dimensão de sensação auditiva.

Outro ponto que merece destaque na consideração da modalidade de avaliação perceptiva considerada é o fato de que os ajustes implementados pelo falante por ocasião de alterações vocais podem também sinalizar um efeito estilístico ou mesmo mobilizações prévias à instalação da alteração (Shimizu, 1986).

Dessa forma, procuramos identificar mobilizações presentes que interfiram na qualidade vocal, tendo a noção de que algumas podem fazer parte do inventário de ajustes do falante, anteriormente à instalação dos quadros de paralisia e/ou fibrose laríngea para a população estudada. Diante de tais constatações, podemos afirmar que não necessariamente o grau isolado de desvio do ajuste em relação ao neutro traduz o mecanismo compensatório específico para a situação de alteração vocal, mas certamente revela a constelação de ações no aparelho fonador que culminam no efeito sonoro.

Ressaltamos que o objetivo do presente trabalho não foi propor ou validar uma adaptação do instrumento em questão, mas responder à necessidade de uso de material compatível com o referencial teórico adotado, o qual permitiria o estabelecimento de correspondências traçadas. Reconhecemos, inclusive, que a adaptação implementada do instrumento desta pesquisa representa um primeiro passo para tal finalidade. Diante dos achados, algumas adaptações podem ser propostas, com vistas ao uso clínico.

Nesse nível, abordando tópicos específicos do roteiro utilizado, destacamos a necessidade de aprofundamento na descrição dos ajustes labiais, em que as dimensões longitudinal e transversal praticamente se sobrepõem, do ponto de

vista da descrição clínica. Além disso, como primeira observação, apesar das modalidades recentes do *VPAS* não mais incluírem a avaliação em duas passadas (Laver, 1994; Laver, 2000), optamos por mantê-la neste estudo conforme mencionado anteriormente, pois a atribuição do grau após a identificação do ajuste fica facilitada.

Vale relembrar que as amostras foram apresentadas nas duas etapas em ordem aleatória e com repetição de dois falantes distintos em cada uma delas, procurando-se pelo julgamento confiável dos juízes. A maior parte dos julgamentos foi consistente, denotando que o modelo é aplicável em nossa realidade.

Outro aspecto a facilitar futuras incursões neste campo em nossa realidade, envolve a possibilidade de análise em duas etapas, em que a primeira poderia envolver a identificação de ajustes laríngeos, dada a maior familiaridade dos juízes com esta esfera da qualidade vocal, e, numa segunda, o julgamento dos aspectos supralaríngeos da qualidade vocal. Registramos tal possibilidade como proposta para adequação de instrumento na formação dos fonoaudiólogos em nosso meio.

Outro ponto a ser salientado neste momento da discussão concerne às combinações adotadas pelos juízes para referirem a rouquidão, parâmetro de primeira eleição diante das manifestações de prega vocal. No modelo em questão, podem ser referidas enquanto combinações de ajustes referentes a voz áspera e voz soprosa ou ao escape de ar, em estreita relação com propostas de que a rouquidão seja entendida como a combinação das dimensões auditivas de soprosidade e aspereza (Laver et al, 1982; Titze, 1995)

Além disso, os juízes revelaram consistência entre seus julgamentos no sentido de que alguns apontaram ajustes de laringe alta, enquanto outros de falsete para i3. No caso de i4 referiram laringe baixa ou *vocal fry*, numa correspondência à esfera de julgamento perceptivo referente ao *pitch* da emissão. Tais observações traçadas são preliminares e envolverão futuras reflexões e incursões na área, as quais poderão contar com apoio da literatura e de profissionais atuantes nas áreas de Fonaoudiologia, Fonética Acústica e Ciências da Fala, numa concepção interdisciplinar da qualidade vocal em indivíduos portadores de disfonias e, especialmente em

pesquisas pautadas na implementação do *VPAS* (Laver et al, 1981; Beck, 1988; Laver, 1994; Laver, 2000).

Neste ponto, destaca-se a confluência de áreas de conhecimento voltadas à exploração das particularidades do sinal sonoro exposta no início deste trabalho. Poderíamos, por assim considerar, o interesse comum que a qualidade vocal desperta desperta em diversas áreas de conhecimento. A abordagem interdisciplinar da qualidade vocal caracteriza-se como um passo importante no conhecimento mais amplo dessa dimensão da manifestação da sonoridade humana, a qual tem grande impacto na qualidade de vida do falante. Diante da possibilidade de comprometimento das condições de comunicação oral em vigência de alteração do sinal vocal, expressas enquanto queixa ou em variados graus de manifestações clínicas, pudemos reconhecer a complexidade do quadro que pode decorrer de alterações do aparelho fonador como aquela eleita para estudo.

Conseqüentemente, a exploração da qualidade vocal enquanto um traço recorrente da característica da emissão do falante deve ser cuidadosamente dimensionada na Fonoaudiologia, no sentido de se traduzir, no futuro, em abordagens mais fidedignas de avaliação, a revelarem os fatores subjacentes que caracterizam o sofrimento dos indivíduos disfônicos. Tal aspecto inclui a necessidade de estabelecimento de consistente treinamento auditivo na formação do fonoaudiólogo e das nossas possibilidades de atuação em avaliação e assistência dos distúrbios da voz, com importantes contribuições aos campos experimental e clínico e, inclusive para adaptação da formação do profissional fonoaudiólogo.

A justificativa para que não adotemos roteiro de avaliação de qualidade vocal é um tanto contraditória, no sentido de que a referida necessidade de consideração da individualidade de falante na situação de comunicação serve justamente de argumento para a adoção de um modelo de avaliação perceptiva que considere a qualidade recorrente ao longo da fala, unindo a concepção de investigação acústica e fisiológica adotada. Tal abordagem favoreceu a caracterização dos ajustes particulares dos falantes estudados, diferenciando-os inclusive entre si, de forma a salientar a possibilidade de avaliação refinada e individualizada.

Diante da dimensão dos achados discutidos, como contribuição à

clínica dos distúrbios da voz, salientamos a possibilidade de se caracterizar os mecanismos compensatórios presentes nos portadores de alterações vocais, como tentativa de resgate do processo de sonorização. Ao considerarmos a plasticidade do aparelho fonador, devemos, como clínicos e pesquisadores, reconhecer as particularidades da emissão, a fim de direcionar nossa intervenção para o que é peculiar do falante e, num segundo momento, contrapor tais achados às estimativas baseadas em estudos de populações, caminho este adotado por muitos estudos na área.

Em nosso ponto de vista, a trajetória dos estudos deveria ter sido direcionada a num primeiro momento, entender a riqueza dos mecanismos compensatórios, detectando ajustes relevantes, e, posteriormente, partir para uma abordagem quantitativa e estudos de população.

O grupo estudado revelou a fascinante condição de adaptação do aparelho fonador diante de limitações de ordem de mobilização e vibração dos tecidos da região laríngea e as variadas resultantes sonoras possíveis a partir de gestos particularizados e integrados para cada falante, como forma de cumprir a necessidade crucial de sonorização da coluna aérea e transmissão da energia sonora, como um importante requisito da comunicação oral.

Este exercício teórico, a partir de amostras de fala de um grupo particular de indivíduos disfônicos, ressalta, especialmente a necessidade de embasamento que permeia a incursão do fonoaudiólogo na avaliação, reabilitação e assessoria vocal, como forma de realmente contemplar a característica fundamental de plasticidade do aparelho fonador humano no desempenho de uma função adaptada, enquanto desempenhada por estruturas primariamente destinadas ao exercício de funções vitais de respiração e deglutição.

Sinaliza, também, a necessidade de aprofundarmos o campo de descrição sonora transcendendo a clássica divisão fala/voz implementada na clínica fonoaudiológica como forma de ajudar a dimensionar o real impacto da disfonia nas condições de comunicação.

Por fim, aponta a relevância de se aprofundar as descrições das relações entre eventos glóticos e supraglóticos. Nesse sentido, as análises de *VOT* e de longo termo revelaram-se particularmente importantes.

Podemos então considerar que, diante de manifestações de base comum, enquanto quadro clínico, desvendamos ações diferenciadas com maior ou menor efetividade em termos da adaptação da função vocal, condição visada pela intervenção terapêutica.

### 9. CONSIDERAÇÕES FINAIS

A investigação, empreendida nesta tese, sobre a qualidade vocal de um grupo de indivíduos disfônicos seguiu uma abordagem interpretativa e integrativa, por considerar a esfera acústica como intermediadora entre a perceptiva e a fisiológica e por se pautar pela correlação de dados obtidos por variados instrumentais advindos dessas três esferas.

Tal nível de correlação foi possível graças à adoção de referencial teórico do campo da Fonética, na conjunção de propostas de Fant (1970) e Laver (1980), possibilitando-nos romper a clássica dicotomia estabelecida entre normalidade e alteração vocal em nossa área, a Fonoaudiologia, ao enfocar o impacto que a disfonia promove na dimensão de produção e percepção do sinal sonoro, com conseqüências mais amplas na comunicação oral.

Numa abordagem pouco usual em nossa realidade, a consideração da qualidade vocal sob o ponto de vista fonético revelou-nos a possibilidade de enfocála a partir de uma noção básica da teoria fonética: a plasticidade do aparelho fonador.

Ao estudarmos a qualidade vocal de quatro falantes que partilham de condições estruturais semelhantes, registramos uma vasta gama de mobilizações, traduzidas em variados registros acústicos, perceptivos e fisiológicos. Pudemos identificar as informações impressas no sinal acústico em modalidades de análise de longo termo e seus correlatos na dimensão de curto termo. Os resultados dessas análises permitiram o estabelecimento de correspondências em nível fisiológico com eventos relativos à ação da fonte e à formação de acoplamentos com as cavidades da via aérea inferior. Tais achados encontraram correspondência com os ajustes laríngeos e supralaríngeos na proposta de avaliação da qualidade vocal com motivação fonética e, finalmente, na inteligibilidade de fala.

Nesse campo, a adoção de roteiro que congrega mobilizações do plano laríngeo (fonatório) e supralaríngeo (articulatório) conduziu-nos à possibilidade de transcender as limitações impostas por procedimentos de análise acústica pautados exclusivamente em estimativas numéricas e pelas descrições de ordem perceptiva essencialmente impressionísticas, além de pouco consistentes para aplicação em

diferentes contextos da prática clínica fonoaudiológica. Livres dessas tendências, fomos capazes de avançar em relação à tentativa de classificar os falantes em termos da simples dicotomia exposta, ao priorizarmos os ajustes altamente individualizados, parte deles, pelo menos, impostos pela situação de limitação de ação glótica.

Diante deste quadro, adotar a clássica oposição entre normalidade e alteração vocal limitaria a compreensão do que se encontra na intermediação dos extremos, ao ponto de perdermos a possibilidade de aprofundar a noção do que há de peculiar na situação de disfonia.

Destacamos a necessidade de uma abordagem altamente integrada e cooperativa entre a diversas áreas de conhecimento voltadas ao enfoque da produção de fala humana, com o intuito de contribuir para a formação de conhecimento sobre os complexos níveis de atividade que permeiam a interface da produção e da percepção sonora.

Estudos desta natureza concretizam-se como uma proposta de colaboração entre profissionais da área das Ciências da Fala, entre elas a Fonética. Em face desta perspectiva, esperamos que o presente estudo desperte o estímulo para que o fonoaudiólogo busque nessas ciências subsídios que o levem a refletir sobre a descrição dos eventos fônicos, como forma de exercer seu papel de terapeuta no auxílio à reabilitação de faculdade tão preciosa ao ser humano: a fala.

# 10. ANEXOS

# Anexo 1 - Roteiro de avaliação da qualidade vocal segundo modelo fonético

1.1. Versão Original do *Vocal Profile Analysis – VPAS* (Laver, 2000)

CATEGORY	SETTING	SCALAR DEGREE			S			
		neutral	1	2	3	4	5	6
	LARYNGEAL							
LONGITUDINAL	raised larynx							
	lowered larynx							
	LABIAL							
	labiodentalization							
	labial protrusion							
	LABIAL							
CROSS	lip-rounded							
SECTIONAL	lip-spread							
	MANDIBULAR							
	close jaw							
	open jaw							
	LINGUAL TIP/							
	BLADE							
	advanced tip/blade							
	retracted tip/blade	1						
	LINGUAL BODY							
	advanced body							
	retracted body							
	raised body							
	lowered body							
	LINGUAL ROOT							
	constricted pharynx							
	expanded pharynx							
VELOPHARYNGEAL	VELIC COUPLING							
	nasal							
	denasal							
PHONATORY	modal voice							
	falsetto							
	whisper							
	whispery voice							
	creak							
	creaky voice							
	harsh voice							
	breathy voice							

# 1.2. Versão traduzida do Vocal Profile Analysis – VPAS (Laver, 2000)

CATEGORIA	AJUSTES	GRAUS DE ESCALA					
		NEUTRO	1 2	2 3	4	5	6
SUPRALARÍNGEOS	LARÍNGEOS						
LONGITUDINAIS	laringe alta						
	laringe baixa						
	LABIAIS						
	labiodentalização						
	protrusão						
SUPRALARÍNGEOS	LABIAIS						
<b>TRANSVERSAIS</b>	arredondados						
	estirados						
	MANDIBULARES						
	fechada						
	aberta						
	LINGUAIS - PONTA/ LÂMINA		_			_	_
	avançada						
	recuada						
	LINGUAIS - CORPO						
	avançado				П		
	retraído						
	elevado						
	abaixado						
	LINGUAIS – BASE						
	constrição faríngea						
	expansão faríngea						
VELOFARÍNGEOS	ACOPLAGEM VELAR						
	nasal						
	denasal						
FONATÓRIOS	voz modal						
	falsete						
	cochicho						
	voz cochichada						
	vocal fry (crepitância)						
	voz crepitante						
	voz áspera						
	voz soprosa						

# Anexo 2. Guia de consulta ao prontuário médico-hospitalar

Iniciais:	Idade:	DN:
Sexo:	Profissão:	
Natural de :		
Instituição:	Nº de registro:	
Patologia: paralisia da pre	ga vocal em posição	
atribuída a	(duração: m)	
Procedimento de tratamer	nto complementar:	
Queixa/ duração:		
Sintomas vocais:		
Sintomas à deglutição:		
Fonoterapia ( ) s ( ) n		
Período:	Enfoque:	
Resultados:		
Exames complementares	(especificar resultados):	
Programação de tratamen	ito:	
Seguimento:		

#### Anexo 3 - Termo de Solicitação e Informação Para Participação em Pesquisa

Venho por meio deste solicitar a sua participação em pesquisa aprovada pelo Comitê de Ética deste hospital. Passo a informar abaixo detalhes da pesquisa intitulada "Avaliação integrativa e interpretativa da qualidade vocal em um grupo de indivíduos disfônicos" para sua informação e livre manifestação quanto ao consentimento de sua participação:

-Objetivo da pesquisa: avaliar os vários aspectos da qualidade vocal de indivíduos portadores de disfonia, como forma de propiciar um atendimento mais amplo às suas necessidades no âmbito da reabilitação fonoaudiológica

#### - Procedimentos, riscos e condições gerais de participação:

As pessoas que fizerem parte do grupo estudado passarão por uma avaliação detalhada da produção de sua voz: gravação utilizando microfones na cabeça e no pescoço e exame detalhado da sua laringe. Os procedimentos serão realizados por uma fonoaudióloga especialista em Voz e um médico especialista em Otorrinolaringologia, fora das instalações do HSPM em duas etapas num mesmo dia. As pessoas participantes não terão gastos com transporte e exames, bem como não serão remuneradas para tal finalidade.

Os dados da avaliação realizada serão posteriormente encaminhados ao SAME (Serviço de Arquivo Médico) para arquivamento em prontuário.

Os exames que serão realizados não causam efeitos colaterais ou problemas de saúde e não requerem preparo prévio. Os participantes não receberão medicação, nem deverão interromper os tratamentos que estejam realizando. Esses exames foram autorizados pelo Comitê de Ética deste hospital, justamente porque não causarem riscos à saúde. Não haverá necessidade de repouso após o procedimento.

Os dados dos exames serão apenas utilizados para pesquisa e eventos científicos, sendo mantido sigilo de nome, histórico e demais informações que venham a ferir os direitos dos participantes.

#### - Esclarecimento de dúvidas:

As dúvidas podem ser esclarecidas com a pesquisadora **Zuleica Camargo** no telefone **36708316**.

Como considerações finais, gostaria de ressaltar que os resultados desta pesquisa poderão ajudar na reabilitação dos problemas de voz que afetam nossa população. Sua participação é de livre escolha, de forma que sua recusa não causará prejuízo em termos dos tratamentos recebidos nesta instituição.

Tendo compreendido os aspectos anteriormente delineados
( ) aceito participar desta pesquisa- nome
( ) não aceito participar desta pesquisa
Data:

Fonoaudióloga Zuleica Camargo- CRFa 4588/SP Hospital do Servidor Público Municipal (HSPM) Programa de Estudos Pós Graduados em Lingüística Aplicada e Estudos da Linguagem (LAEL) – Pontifícia Universidade Católica de São Paulo (PUC/SP)

# Anexo 4-Termo de Consentimento Pós Informado do Sujeito do Estudo ou do Responsável Legal

# HOSPITAL DO SERVIDOR PÚBLICO MUNICIPAL

Eu,
RG, residente
responsável
legal por
natureza (grau de parentesco, tutor, curador, etc.)
consinto que sejam utilizados os meus dados clínicos de história e/ou exame físico,
e/ou exames laboratoriais, e/ou fotográficos, e/ou gravações em áudio e vídeo para
pesquisas científicas e publicações em revistas de circulação unicamente do meio
científico.
Estou ciente da liberdade de retirar meu consentimento a qualquer
momento e de deixar de participar do estudo, sem que isto traga prejuízo à
continuidade da assistência nesta instituição. Receberei uma cópia assinada deste
termo.
São Paulo,
(assinatura do sujeito ou responsável legal)
(assinatura do pesquisador)

#### Anexo 5- Roteiro de preparo do instrumental e indivíduo avaliado

#### Etapa 1

#### Preparo do estúdio de gravação

- Regulagem de altura do suporte de cabeça da cadeira e distância da mesa e do eletroglotógrafo
- Posicionamento de tapetes de borracha
- Fixação do texto em suporte, colocados sobre mesa situada em frente ao local que o falante virá a ocupar
- Checagem da melhor posição a ser ocupada pelo posição indivíduo avaliado, equipamento de eletroglotografia e texto
- Carga prévia de bateria do EGG (não utilizar alimentação direta da eletricidade no momento da avaliação)
- Troca de pilhas do microfone, revisando opção de curva de entrada linear

#### Etapa 2

#### Preparo dos materiais na sala anexa

- Fitas de gravação DAT: avanço completo, retorno ao segundo minuto, demarcando o ponto inicial de gravação
- Revisão das conexões de entrada de sinais de microfone e eletrodos e seus processamento pelos recursos computacionais e de gravador digital
- Checagem do nível de ruído presente na gravação e da qualidade do sinal registrado (teste prévio com examinadores presentes).
- Montagem de tripé com câmera de vídeo, posicionando-a de forma a focalizar o painel frontal do equipamento EG2 e o indivíduo testado, situados no estúdio de gravação

#### Etapa 3

#### Preparo do indivíduo avaliado

- Acomodação em cadeira adaptada, com apoio de cabeça ajustado à sua altura, observando a posição de cabeça e proeminência laríngea, além de apoio de pés em tapete de borracha
- Solicitação para remoção de colares e brincos, demais adornos na região de pescoço
- Limpeza da pele do pescoço com lenço de papel, seguida de algodão embebido com álcool etílico hidratado anti-séptico tópico
- Palpação cervical e localização das lâminas da cartilagem tireóide
- adaptação dos eletrodos de superfície do equipamento EG2, cujas placas metálicas receberam uma camada de gel para eletrodos hipoalergênico Spectra 360, fixação dos eletrodos com fita de velcro ao redor do pescoço
- Checagem do nível de fixação de eletrodos (indicador de potência do sinal o display do EG2 e equipamento de registro na sala anexa).
- Acoplagem de microfone à cabeça distância da lateral direita da comissura labial em 3 cm.

#### Anexo 6 -Roteiro de gravação digital e eletroglotográfica das amostras de fala

- Sem apoio manual nos eletrodos (VFCA, LF 40 Hz, ganho baixo) –
   Emissão de:
  - vogal [a] em padrão entrecortado
     – seqüências de três estímulos:
     cinco tentativas;
  - vogal [a] em padrão sustentado de duração mínima de dois segundos em padrão habitual e confortável: duas tentativas;
  - leitura do texto: uma tentativa.
- Com apoio digital nos eletrodos (VFCA, LF 40 Hz, ganho baixo) –
   Emissão de:
  - vogal [a] em padrão habitual e confortável: uma tentativa;
  - vogal [a] grave: uma tentativa;
  - vogal [a] mais grave possível: uma tentativa;
  - · vogal [a] aguda: uma tentativa;
  - vogal [a] mais aguda possível: uma tentativa;
  - vogal [a] em padrão habitual e confortável: uma tentativa;
  - vogal [a] em fraca intensidade: uma tentativa;
  - vogal [a] mais fraca possível: uma tentativa;
  - vogal [a] forte: uma tentativa;
  - vogal [a] mais forte possível: uma tentativa;
  - vogal [a] em padrão habitual e confortável: uma tentativa;
  - vogal [a] em padrão entrecortado
     – seqüências de três estímulos:
     cinco tentativas;
  - leitura do texto: uma tentativa.
- Com apoio digital nos eletrodos (DEGG, LF 40 Hz, ganho baixo) –
   Emissão de:
  - vogal [a] em padrão entrecortado
     – seqüências de três estímulos:
     cinco tentativas;

- vogal [a] em padrão sustentado de duração mínima de dois segundos em padrão habitual e confortável: uma tentativa;
- leitura do texto: uma tentativa.

# Anexo 7- Roteiro de avaliação funcional da laringe

1.	Condições gerais do exame
	a. Procedimento ( ) telescopia ( ) nasolaringospia
	( ) videoestroboscopia ( ) videoquimografia
	b. Observações:
2.	Aspectos estruturais e de mobilidade
	Aspecto geral e integridade das estruturas
	b. Mobilidade de pregas vocais
•	December de la como 7 de la como 10 de de
3.	Presença de secreção – quantidade
4.	Fonação – dinâmica vibratória
	a. Fechamento glótico
	( ) completo ( ) fenda anterior ( ) irregular ( ) fenda posterior
	( ) completo ( ) fenda anterior ( ) irregular ( ) fenda posterior ( ) fenda em ampulheta ( ) fechamento incompleto
	( ) fenda em ampulheta ( ) fechamento incompleto
	( ) fenda em ampulheta
	( ) fenda em ampulheta ( ) fechamento incompleto b. Periodicidade c. Fechamento em fase
	( ) fenda em ampulheta ( ) fechamento incompleto b. Periodicidade c. Fechamento em fase ( ) fase aberta ( ) fase fechada
	<ul> <li>( ) fenda em ampulheta</li> <li>( ) fechamento incompleto</li> <li>b. Periodicidade</li> <li>c. Fechamento em fase</li> <li>( ) fase aberta</li> <li>( ) fase fechada</li> <li>d. nível vertical de aproximação de pregas vocais</li> </ul>
	<ul> <li>( ) fenda em ampulheta</li> <li>( ) fechamento incompleto</li> <li>b. Periodicidade</li> <li>c. Fechamento em fase</li> <li>( ) fase aberta</li> <li>( ) fase fechada</li> <li>d. nível vertical de aproximação de pregas vocais</li> <li>e. amplitude de vibração</li> </ul>
	<ul> <li>( ) fenda em ampulheta</li> <li>( ) fechamento incompleto</li> <li>b. Periodicidade</li> <li>c. Fechamento em fase</li> <li>( ) fase aberta</li> <li>( ) fase fechada</li> <li>d. nível vertical de aproximação de pregas vocais</li> <li>e. amplitude de vibração</li> <li>f. simetria de fase</li> </ul>
	<ul> <li>( ) fenda em ampulheta</li> <li>( ) fechamento incompleto</li> <li>b. Periodicidade</li> <li>c. Fechamento em fase</li> <li>( ) fase aberta</li> <li>( ) fase fechada</li> <li>d. nível vertical de aproximação de pregas vocais</li> <li>e. amplitude de vibração</li> <li>f. simetria de fase</li> <li>g. onda mucosa</li> </ul>

Anexo 8. Roteiro de Avaliação Perceptiva Auditiva da Qualidade Vocal com Motivação Fonética

8a.

Roteiro do juiz participante da análise de dados do projeto de doutor	ado
desenvolvido por Zuleica A. Camargo sob orientação de Prof.ª Dra Sar	ıdra
Madureira (Programa de Estudos Pós-Graduados em Lingüística Aplicad	ае
Estudos da Linguagem: LAEL- PUC/SP)	
CD – Faixa 1	
Data:	
Iniciais:	
Formação (graduação):	
Titulação (mais recente):	
Tempo de atuação em avaliação vocal:	
Usa alguma escala ou roteiro de avaliação?	
( ) não ( ) sim	
A partir deste ponto são expostos os itens completos referentes a um indiví	duo
avaliado:	
<b>8b.</b> ETAPA 1A. Nível geral de <b>Alteração vocal</b> , Grau de <b>Agradabilidade</b> para emiss	ကိုဝင
de vogais	,UE3
CD – Faixa 2	
(0= ausente; 5= máximo) e Julgamento de qualidade vocal:	
Grau geral de Alteração Vocal	
0 1 2 3 4 5	
Grau de Agradabilidade	

3

2

Qualidade Vocal:\_

ETAPA 1B. Nível geral de Alteração vocal, Grau de Agradabilidade para emissões de sentenças

#### CD - Faixa 3

(0= ausente; 5= máximo) e julgamento de qualidade vocal:

1	2	3	4	5
		•	-	
dabilidade				
1	2	3	4	5
	dabilidade 1	dabilidade 1 2	dabilidade 1 2 3	dabilidade 1 2 3 4

Vocal:

#### 8d.

#### **CD – Faixas 4 e 5**

Faixa 4 Considerações sobre o modelo teórico – ajustes supralaríngeos Faixa 5 - Considerações sobre o modelo teórico - ajustes laríngeos

#### Referência: Ajuste Neutro

- •Lábios: nem projetados, nem estirados
- •Laringe: nem abaixada, nem elevada
- Diâmetro do trato supralaríngeo: constante
- •Articulações orais anteriores: lâmina da língua
- •Base de língua: nem avançada, nem recuada
- Trato vocal sem constrições
- •Mandíbula: nem acentuadamente aberta, nem fechada
- •Nasalidade presente apenas nos segmentos característicos
- Vibração de pregas vocais periódica: com tensão longitudinal e adutora moderada, idem para compressão medial

- •Uso eficiente da corrente de ar, sem ruídos
- •Tensão muscular no aparelho fonador: nem alta, nem baixa

8e.

#### Faixa 6

ETAPA 2A. Identificação dos ajustes presentes

CATEGORIA	AJUSTES	PRESENTES
SUPRALARÍNGEOS	LARÍNGEOS	
LONGITUDINAIS	Laringe alta	
	Laringe baixa	
	LABIAIS	
	Labiodentalização	
	Protrusão	
SUPRALARÍNGEOS	LABIAIS	
TRANSVERSAIS	Arredondados	
	Estirados	
	MANDIBULARES	
	Fechada	
	Aberta	
	LINGUAIS - PONTA/ LÂMINA	
	Avançada	
	Recuada	
	LINGUAIS – CORPO	
	Avançado	
	Retraído	
	Elevado	
	Abaixado	
	LINGUAIS – BASE	
	Constrição faríngea	
	Expansão faríngea	
VELOFARÍNGEOS	ACOPLAGEM VELAR	
	Nasal	
	Denasal	
FONATÓRIOS	Modal	
	Falsete	
	Escape de ar	
	Voz soprosa	
	Vocal fry/ crepitância	
	Voz crepitante	
	Voz áspera	
	Hiperfunção	
	Hipofunção	

Ocorrências em curto termo: ( )diplofonia ( )quebras ( ) instabilidades

Para ajustes de ocorrência intermitente assinalar (i).

8f.

# Faixa 7

Etapa 2B. Atribuição de graus na escala de avaliação dos ajustes

CATEGORIA	AJUSTES	GRAUS DE ESCA						
		Neutro	1	2	3	4	5	6
SUPRALARÍNGEOS	LARÍNGEOS							
LONGITUDINAIS	Laringe alta							
	Laringe baixa							
	LABIAIS							
	Labiodentalização							
	Protrusão							
SUPRALARÍNGEOS	LABIAIS							
TRANSVERSAIS	Arredondados							
	Estirados							
	MANDIBULARES							
	Fechada							
	Aberta							
	LINGUAIS - PONTA/ LÂMINA							
	Avançada							
	Recuada							
	LINGUAIS - CORPO							
	Avançado							
	Retraído							
	Elevado							
	Abaixado							
	LINGUAIS – BASE							
	Constrição faríngea							
	Expansão faríngea							
VELOFARÍNGEOS	ACOPLAGEM VELAR							
	Nasal							
	Denasal							
FONATÓRIOS	Modal							
	Falsete							
	Escape de ar							
	Voz soprosa							
	Vocal fry/crepitânncia							
	Voz crepitante							
	Voz áspera							
	Hiperfunção							
	Hipofunção							

Ocorrências em curto termo: ( )diplofonia ( )quebras ( ) instabilidades

Para ajustes de ocorrência intermitente assinalar (i).

#### Graus:

- 1-3: faixa dos ajustes presentes na fala normal
  - •1 mínima diferença em relação ao ajuste neutro
  - •2 leve diferença em relação ao ajuste neutro
  - •3 moderada diferença em relação ao ajuste neutro
- 4-6: patologia ou extremos paralingüísticos da expressão da emoção
  - •4 notável/ importante diferença em relação ao ajuste neutro
  - •5 severa diferença em relação ao ajuste neutro
  - •6 extrema diferença em relação ao ajuste neutro

8g.

#### Faixa 8

Caro juiz, este anexo pode ser destacado. Trata-se de uma síntese dos ajustes descritos por Laver (1980) na publicação original do modelo fonético de descrição da qualidade vocal e da proposta para *SVEA* (*Swedish Voice Evaluation Approach*) desenvolvido desde a década de 80 e presente na publicações de Hammarberg, Gauffin (1995) e Hammarberg (2000).

Lista de ajustes laríngeos e supralaríngeos descritos na publicação original do modelo fonético de descrição da qualidade vocal (Laver,1980).

#### AJUSTES SUPRALARÍNGEOS

#### Longitudinais

#### Modificações no eixo longitudinal do trato vocal

- Laringe
- Alta
- Baixa
- Lábios
- Protrusão
- Labiodentalização

#### Transversais (latitudinais)

Tendências quase permanentes a manter um efeito particular constritivo ou expansivo na área transversal de todo o trato vocal

•l ábios

- Expansão horizontal do espaço interlabial
- Expansão vertical
- Constrição horizontal
- Constrição vertical
- Expansão horizontal e expansão vertical
- Constrição horizontal e constrição vertical
- Expansão horizontal e constrição vertical
- Constrição horizontal e expansão vertical

#### •Língua

#### -Ponta/lâmina

- Articulação de ponta
- Articulação de lâmina
- Articulação retroflexa

#### -Corpo

- Dentalizada
- Alveolarizada
- Palatoalveolarizada
- Palatalizada
- Velarizada
- Uvularizada
- Faringealizada
- Faringolaringealizada

#### -Base

- Base de língua avançada
- Base de língua retraída

#### Pilares

De fauces ou de pilares

#### Faringe

Faringealizada

#### •Mandíbula

- Posição aberta
- Posição fechada
- Posição protruída
- Sistema Velo-faríngeo
  - Nasal
  - o denasal

#### **AJUSTES LARÍNGEOS**

- Tipos simples
  - Modal
  - Falsete
  - Cochicho
  - vocal fry/crepitância
- Tipos compostos
  - Vocal fry cochichado
  - Voz cochichada
  - Falsete cochichado
  - Voz crepitante
  - Falsete crepitante
  - Voz cochichada crepitante
  - Falsete cochichado crepitante
  - Voz soprosa
  - Voz áspera
  - Falsete áspero
  - Vocal fry áspero
  - Voz áspera cochichada

- Falsete áspero cochichado
- Voz áspera crepitante
- Falsete áspero crepitante
- Voz áspera cochichada crepitante
- Falsete áspero cochichado crepitante

Descrição dos parâmetros de qualidade vocal propostos para avaliação perceptiva de alterações vocais na clínica - SVEA (Swedish Voice Evaluation Approach):

PARÂMETRO DE QUALIDADE VOCAL	DEFINIÇÃO
Afônica Afônica	Voz perde a sonoridade de forma constante ou intermitente (há momentos de escape de ar ou
Soprosa	perda da voz)  Voz produzida com fechamento glótico insuficiente, pregas vocais vibram porém um tanto abduzidas,
Hiperfuncional/ Tensa	criando um ruído audível turbulento na glote Sons sonoros com características de tensão devido à compressão de pregas vocais e laringe durante a fonação com fluxo insuficiente
Hipofuncional/ relaxada	Oposto da hiperfuncional, tensão de pregas vocais e de musculatura laríngea insuficiente, resultanto em voz fraca e com escape
V. fry/ creaky	Vibração de baixa freqüência periódica/ aperiódica, na qual as pregas vocais estão bem próximas e apenas uma pequena porção delas encontra-se livre para vibrar. Também conhecido como registro pulsátil
Rouca	Aperiodicidade de baixa freqüência, presumivelmente relacionada a algum grau de vibração irregular de pregas vocais
Rangido/ áspero	Aperiodicidade de alta freqüência, presumivelmente relacionada a algum grau de vibração irregular de pregas vocais
Qualidade Vocal/ pitch instável	Voz flutua em qualidade ou <i>pitch</i> ao longo do tempo
Quebras Vocais	Quebras intermitentes entre os registros modal e falsete
Diplofônica	Dois <i>pitches</i> diferentes são identificados simultaneamente
Registros modal e falsete	Modal é o padrão normal (padrão) de fonação com fechamento glótico em oposição ao falsete, que usualmente se encontra no extremo superior da extensão de f0, resultado da conformação alongada e fina de pregas vocais com pequeno fechamento

	glótico
Pitch	Principal correlato auditivo de f0
Loudness	Principal correlato auditivo do nível de pressão
	sonora do som

(Hammarberg, Gauffin, 1995; Hammarberg 2000)

\_

### Anexo 9. Roteiro de avaliação perceptivo-auditiva de inteligibilidade de sílabas

Indique as sílabas que você identifica na gravação apresentada. Para confirmação de sua resposta, aguarde a repetição que ocorre em seguida à primeira apresentação:

1.						
2.						
3.						
4.						
5.						
6.						
7.						
8.						
9.						
10.						
11.						
12.						
13.						
14.						
15.						

# Anexo 10. Textos utilizados em teste voltado à validação da amostra para análise de longo termo

# 10.1. Texto de composição exclusiva por sons sonoros (145 palavras, não balanceado foneticamente):

Li uma alegoria de um jovem nobre, abordando a viagem de uma ave na India. A líder das aves é uma bela e animada arara. É a rainha e ninguém duvida da realeza dela. Reina, venerada na região do Vale dourado. A arara exibiu os dons maravilhosos a ela doados e originados de um dom divino. Aos brados diz: "Vamos dominar o ódio e ganhar o reino do Deus amado". Divulga às demais aves as verdades e misérias da vida e julga a maioria dos inimigos dos valores morais de maneira ardorosa. Diz às aves da árdua jornada aos vales do amor, da humildade e da nobreza e ganha aliados valorosos. Ela ainda verbalizava e já formas aladas voavam mirando o ideal de vida a elas divulgado. Baseadas nos novos valores, ajudaram umas as outras de verdade e ganharam o galardão real, banhado em ouro.

#### 10.2. Texto de composição mista (159 palavras, não balanceado foneticamente):

Sei uma estória de um sábio, versando sobre a caminhada de um grupo de pássaros. O chefe dos pássaros é um astuto papagaio. É o mais falante e todos acreditam em seus dotes poderosos. Preside a selva, querido e respeitado. Não esconde suas qualidades e as atribui ao poder Alto. Aos gritos, proclama: "Temos de vender o rancor para alcançar a proteção do ser superior". Discursa sobre a solidariedade e os sentimentos a serem preservados. Reflete sobre os fatos da vida e atribui o fracasso ao não cultivo de valores corretos. Aconselha os pássaros sobre a postura a ser adotada na caminhada pelos vales da fé, da esperança e da caridade e é aplaudido calorosamente. Ele ainda discursava e já cortavam o céu seres alados que alçaram vôo com uma missão bem definida, buscando a felicidade através do espírito de vida conclamado a eles. Fundamentados na ética, comprometeram-se com os que os acompanharam e obtiveram a recompensa da felicidade.

Anexo 11. Distribuição de valores de *VOT* para emissão de sílabas destacadas dos vocábulos "para", "esperto", "bela", "estar", "humildade", "aplicada" e "lugares" para três emissões de cada indivíduo do grupo estudado (I1 a I4) e indivíduo referência (IR).

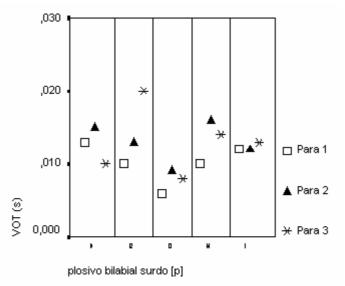


Figura 55. Distribuição das medidas de VOT em segundos para as três emissões de consoante plosiva bilabial surda [p] em **pa**ra dos indivíduos do grupo estudado (i1 a i4) e do indivíduo referência (iR)

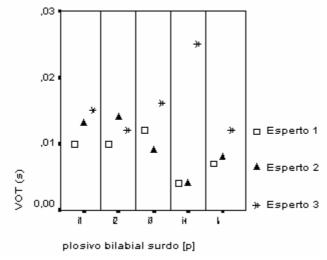


Figura 56. Distribuição das medidas de VOT em segundos para as três emissões de consoante plosiva bilabial surda [p] em es**per**to dos indivíduos do grupo estudado (i1 a i4) e do indivíduo referência (iR)

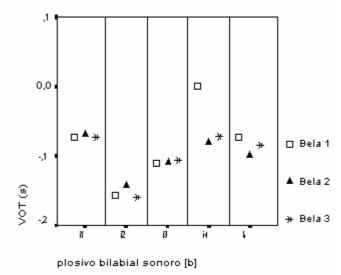


Figura 57. Distribuição das medidas de VOT em segundos para as três emissões de consoante plosiva bilabial sonora [b] em **be**la dos indivíduos do grupo estudado (i1 a i4) e do indivíduo referência (iR)

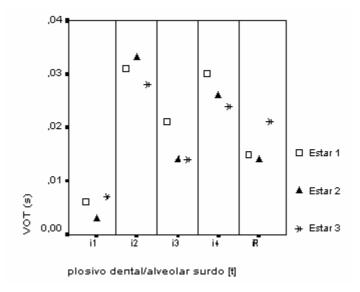


Figura 58. Distribuição das medidas de VOT em segundos para as três emissões de consoante plosiva alveolar surda [t] em estar dos indivíduos do grupo estudado (i1 a i4) e do indivíduo referência (iR)

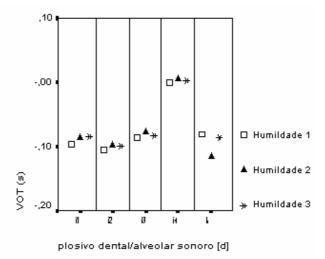


Figura 59. Distribuição das medidas de VOT em segundos para as três emissões de consoante plosiva alveolar sonora [d] em humil**da**de dos indivíduos do grupo estudado (i1 a i4) e do indivíduo referência (iR)

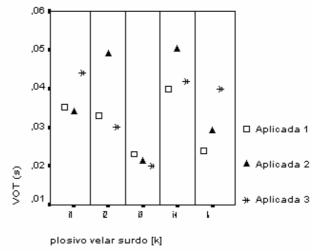


Figura 60. Distribuição das medidas de VOT em segundos para as três emissões de consoante plosiva velar surda [k] em apli**ca**da dos indivíduos do grupo estudado (i1 a i4) e do indivíduo referência (iR)

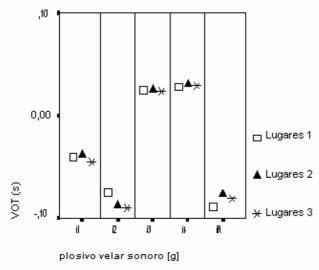
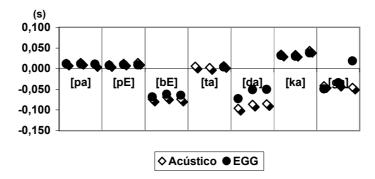


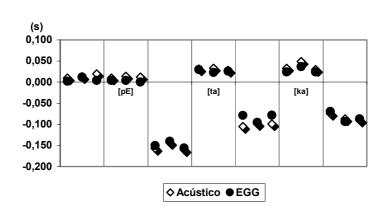
Figura 61. Distribuição das medidas de VOT em segundos para as três emissões de consoante plosiva velar sonora [g] em lu**ga**res dos indivíduos do grupo estudado (i1 a i4) e do indivíduo referência (iR)

Anexo 12. Medidas de *VOT* extraídas por modalidades de análise acústica (inspeção da onda acústica e espectrogramas de banda larga) e eletroglotográfica (inspeção da onda Lx conjugada à onda acústica) para i1, i2, i3, i4 e iR.

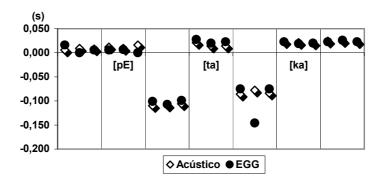
i1



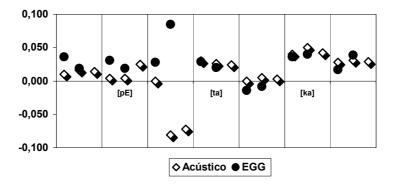
i2



i3



i4



iR

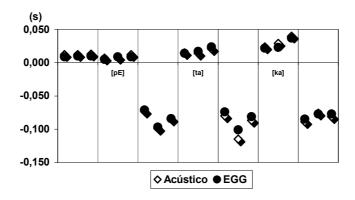


Figura 62. Valores comparativos de VOT acústico e eletroglotográfico (EGG) para três emissões das sílabas [pa], [pɛ], [bɛ], [tar], [da], [ka] e [ga] pelos indivíduos do grupo estudado (i1 a i4) e do indivíduo referência (iR)

#### 11.REFERÊNCIAS

Abercrombie. Elements of general phonetics. Edinburgh: Edinburgh University Press; 1967.

Albano E. Fazendo sentido do som. Ilha do Desterro 1988; 19:11-26.

Ananthapadmanabha TV. Acoustic factors determining perceived voice quality. In: Fujimura O, Hirano M. Vocal fold physiology. San Diego: Singular Publishing Group Inc; 1995. p.113-26.

Andrade LGC. Estudo preliminar da relação entre qualidade vocal e disfagia: uma abordagem acústica [trabalho de conclusão de curso]. São Paulo: Pontifícia Universidade Católica de São Paulo; 2000.

Askenfalt AG, Hammarberg B. Speech waveform perturbation analysis: a perceptual-acoustical comparison of seven measures. J Speech Hear Res 1986; 29:50-64.

Attar FU. A conferência dos pássaros. 9<sup>a</sup> ed. São Paulo: Cultrix; 1993.

Awan SN, Frenkel ML. Improvements in estimating the harmonics-to-noise ratio of the voice. J Voice 1994; 8:255-62.

Awan SN. Automatic estimation of vocal harmonics-to-noise using cesptral analysis. In: Windsor F, Kelly ML, Hewlett N. Investigations in clinical phonetics and linguistics. Mahwah: Lawrence Erlbaum Associates Inc Publishers; 2002. p.475-82.

Baken RJ. Clinical measurement of speech and voice. Boston: College-Hill Press; 1987.

Baken RJ. Eletroglottography. J Voice 1992; 6:98-110.

Baken RJ, Orlikoff RF. Clinical measurement of speech and voice. 2<sup>nd</sup> ed. San Diego: Singular Publishing Group Inc; 2000.

Barrichelo VMO, Heuer RJ, Dean CM, Sataloff R. Comparison of singer's formant, speaker's ring, and LTA spectrum among classical singers and untrained normal speakers. J Voice 2001; 15:344-50.

Båvegård M, Fant G. Interative voice source modeling. In: 13<sup>th</sup> International Congress of Phonetic Sciences; 1995; Stockholm. Proceedings. Stockholm: ICPhS; 1995; 2:634-7.

Beck JM. Organic variation and voice quality [PhD thesis]. Edinburgh: University of Edinburgh; 1988.

Behlau M, Pontes P. Avaliação global da voz. São Paulo: Lovise; 1995.

Bergan CC, Titze IR. Perception of pitch and roughness in vocal signals with subharmonics. J Voice 2001; 15:165-75.

Bickley CA, Stevens KN. Effects of a vocal tract constriction on the glottal source: Data from voice consonants. In: Baer T, Sasaki C, Harris KS. Vocal fold physiology: laryngeal function in phonation and respiration. San Diego: Singular Publishing Group Inc; 1991.

Bielamowicz S, Kreiman J, Gerratt BR, Dauer MS, Berke GS. Comparison of voice analysis systems for perturbation measurement. J Speech Hear Res 1996; 39:126-34.

Bless DM, Hirano M, Feder RJ. Videostroboscopy evaluation of the larynx. Ear Nose Throat J 1987; 66:289-96.

Bless DM, Baken RJ. International Association of Logopedics and Phoniatrics (IALP) Voice Committee discussion of assessment topics. J Voice 1992; 6:194-210.

Borden GJ, Harris KS, Raphael LJ. Speech science primer-physiology, acoustics and perception of speech. 3<sup>rd</sup> ed. Baltimore: Williams & Wilkins; 1994.

Brandi E. Escalas Brandi de avaliação da voz falada. 2ª ed. São Paulo: Atheneu; 1996.

Brandi E. Educação da voz falada. 4ª ed. São Paulo: Atheneu; 2002.

Camargo ZA. Parâmetros vocais e configurações laríngeas na fonação de indivíduos submetidos às laringectomias parciais verticais [dissertação]. São Paulo: Pontifícia Universidade Católica de São Paulo; 1996.

Camargo ZA, Alonso K, Ferreira S. Estudo da prevalência de sintomas de desordem temporomandibular (DTM) num grupo de indivíduos disfônicos [resumo]. Anais; 1996:97. [Apresentado no 4º Congresso Brasileiro de Fonoaudiologia; 1996; Goiânia].

Camargo ZA. Tempo de início de sonorização - VOT - como importante parâmetro de análise de incompetência glótica e suas implicações articulatórias [resumo]. Anais INPLA 1998:32. [Apresentado no 8º Intercâmbio de Pesquisas em Lingüística Aplicada; 1998; São Paulo].

Camargo ZA, Rodrigues KA, Santos AS. Relação entre sintomatologia da desordem vocal e da alteração temporomandibular. In: Ferreira LP, Costa HO. Voz ativa. São Paulo: Roca; 2001. p.237-349.

Cassol M, Madureira S, Behlau M. Aplicação de um modelo fonético na análise da qualidade vocal de indivíduos disfônicos. In: Behlau M. A voz do especialista. Rio de Janeiro: Revinter; 2001. p.85-108.

Cavalli L, Hirson A. Diplophonia reappraised. J Voice 1999; 13:542-56.

Childers DG, Krishnamurthy AK. A critical review of electroglottography. Biomedical Engineering 1985; 12:131-61.

Childers DG, Hicks DM, Moore GP, Eskenazi L, Lalwani AL. Electroglottography and vocal fold physiology. J Speech Hear Res 1990; 33:245-54.

Chomsky N, Halle M. The sound pattern of English. New York: Harper & Row; 1968.

Cleveland TF, Sundberg J, Stone RE. Long-term-average spectrum characteristics of country singers during speaking and singing. J Voice 2001; 15:54-60.

Colton RH, Estill JA. Elements of voice quality: perceptual, acoustic and physiologic aspects. Speech and language: advances in basic research and practice. New York: Academic Press Inc; 1981. p.311-403.

Colton RH, Conture EG. Problems and pitfalls of electroglottography. J Voice 1990; 4:10-24.

Cox NB, Ito MR, Morrison MD. Data labeling and sampling effects in harmonics-to-noise ratios. J Acoust Soc Am 1989a; 85:2165-76.

Cox NB, Ito MR, Morrison MD. Technical considerations in computation of spectral harmonics-to-noise ratios for sustained vowels. J Speech Hear Res 1989b; 32:203-18.

Cranen B, Boves L. The acoustic impedance of glottis-modeling and measurements. In: Baer T, Sasaki C, Harris KS. Laryngeal function in phonation and respiration. San Diego: Singular Publishing Group Inc; 1991. p.218-39.

D'Antonio LL, Wigley TL, Zimmerman GJ. Quantitative measures of laryngeal function following Teflon injection or thyroplasty I. Laryngoscope 1995; 105:256-62.

Davis SB. Acoustic characteristics of normal and pathological voices. In: Lass NJ. Speech and language: advances in basic research and practice. New York: Academic Press; 1979. p.273-335.

Decoster W, Debruyne F. Changes in spectral measures of voice-onset time and age: across-sectional and a longitudinal study. Folia Phoniat Logop 1997; 49: 269-80.

Dejonckere PH. Recognition of hoarseness by means of LTAS-brief research report. Int J Rehab Research 1983; 6:343-5.

Dejonckere PH, Cuvelier F. Spectrographie moyennée de longue durée de la parole: discrimination entre la voix normale et la voix pathologique. Acta Oto-rhino-laryng Belg 1983; 37:784-95.

Dejonckere PH. Comparison between long-time-average-spectra and the sonagraphic degree of hoarseness according to Yanagihara's classification -brief research report. Int J Rehab Research 1984; 7:73-4.

Dejonckere PH, Villarosa D. Analyse spectrale moyenée de la voix: comparaison de voix normales et de voix altérées par différentes catégories de pathologies laryngées. Acta Otorhinolaryngol Belg 1986; 40:426-35.

Dejonckere PH, Wieneke GH. Spectral, cepstral and aperiodicity characteristics of pathological voices before and after phonosurgical treatment. Clinic Ling Phonet 1994; 8:161-9.

Dejonckere PH, Remacle M, Fresnel-Elbaz E, Woisnard V, Crevier L, Millet B. Differentiated perceptual evaluation of pathological voice quality: reliability and correlations with acoustic measurements. Rev Laryngol Otol Rhinol (Bourd) 1996; 117: 219-24.

Dejonckere PH. Cepstral voice analysis: link with perception and stroboscopy. Rev Laryngol Otol Rhinol 1998; 119:245-6.

Dejonckere PH, Lebacq J. Plasticity of voice quality: a prognostic factor for outcome of voice therapy? J Voice 2001; 15:251-6.

Delyiski DD. Acoustic model and evaluation of pathological voice production. In: 3<sup>rd</sup> Conference on Speech Communication and Technology; 1993; Berlin. Proceedings. Berlin: Eurospeech; 1993; 93:183-6.

De Krom G. Some spectral correlates of pathological breathy and rough voice quality for different types of vowel fragments. J Speech Hear Res 1993; 36:254-66.

De Krom G. Some spectral correlates of pathological breathy and rough voice quality for different types of vowel fragments. J Speech Hear Res 1995; 38:794-811.

Emanuel FW, Sansone FE. Some spectral features of "normal" and simulated "rough" vowels. Folia Phoniat 1969; 21:401-15.

Eskenazi L, Childers DG, Hicks DM. Acoustic correlates of vocal quality. J Speech Hear Res 1990; 33:298-306.

Esling JH. Crosslinguistic aspects of voice quality. In: Kent RD, Ball MJ. Voice quality measurement. San Diego: Singular Publishing Group Inc; 2000. p.25-35.

Fabre MP. Un procédé électrique percutané d'inscription de l'accolement glottique au cours del la phonation:glottographie de haute frequence. Premier résultats. Acad Natl Med 1957; 29: 66-9.

Fabre MP. La glottographie en haute fréquence, particularités de l'appareillage. Comptes Rendus des Séances de la Societé de Biologie et de ses Filiales 1959:1361-4.

Fant G. On the predictability of formant level and spectrum envelopes from formant frequencies. For Roman Jakobson, 's- Gravenhage, 1956. p.109-20.

Fant G. Descriptive analysis of the acoustic aspects of speech. Logos 1962a; 5:3-17.

Fant G. Sound spectrography. In: 4<sup>th</sup> International Congress of Phonetic; 1962; Mouton. Proceedings. Mouton: ICPhS; 1962b: 14-33.

Fant G. Acoustic theory of speech production. 2<sup>nd</sup> ed. Paris: Mouton; 1970.

Fant G. Speech sounds and features. Cambridge: MIT Press; 1973.

Fant G. Glottal flow: models and interaction. J Phonetics 1986; 14: 393-9.

Faure MA, Muller A. Stroboscopy. J Voice 1992; 6:139-48.

Flanagan JL. A difference limen of vocal formant frequency. J Acoust Soc Am 1955; 27:613-7.

Flanagan JL. A difference limen of vocal formant amplitude. J Acoust Soc Am 1957; 27:288-91.

Flanagan JL. Some properties of the glottal sound source. J Speech Hear Res 1958; 1:99-116.

Flanagan JL. Speech analysis synthesis and perception. 2<sup>nd</sup> ed. New York: Springer-Verlag; 1972.

Feijoo S, Hernandez C. Short-term stability measures for the evaluation of vocal quality. J Speech Hear Res 1990; 33:324-34.

Figueiredo RM. A eficácia de medidas extraídas do espectro de longo termo para a identificação de falantes. Cad Est Ling 1993; 25:129-60.

Fourcin AJ, Abberton E. First application of a new laryngograph. Med Biol Illust 1971; 21:172-82.

Fourcin AJ, Abberton E. First application of a new laryngograph. Volta Review 1972; 69:507-18.

Fourcin AJ, Abberton E. Laryngograph studies of vocal-fold vibration. Phonetica 1977; 34:313-15.

Fourcin A, Abberton E. Voice quality by eye, ear and measurement. In: International Congress of Phonetic Sciences; 1999; San Francisco. Proceedings. San Francisco: ICPhS; 1999; 2:1345-8.

Fourcin AJ. Voice quality and electrolaryngography. In: Kent RD, Ball MJ. Voice quality measurement. San Diego: Singular Publishing Group Inc; 2000. p.285-306.

Fritzell B, Hallén O, Sundberg J. Evaluation of Teflon injection procedures for paralytic dysphonia. Folia phoniat 1974; 26:414-21.

Fröhlich M, Michaelis D, Kruse E. The objective assessment of voice quality with the hoarseness diagram. HNO 1998; 46:684-89.

Frøkjær-Jensen B, Prytz S. Registration of voice quality. Brüel Kjaer Technical Review 1976: 3-17.

Frøkjær-Jensen B. Acoustic-statistical analysis of voice quality. Annual Report of the Institute of Phonetics of the University of Copenhagen ARIPUC 1979; 13:189-200.

Fry D. The physics of speech. Cambridge: Cambridge University Press; 1979.

Fuks L, Hammarberg B, Sundberg J. A self-sustained vocal-ventricular phonation mode: acoustical, aerodynamic and glottographic evidences. TMH-QPSR. Speech Transmission Laboratory Quarterly Progress and Status Report 1998; 3: 49-59.

Fuks L. Computed-aided analysis of extended vocal techniques for compositional applications. Entre Lugar 1999; 3:95-104.

Gandour J, Dardarananda R. Voice onset time in aphasia. Brain Lang 1984; 23:177-205.

Gauffin J, Sundberg J. Clinical applications of acoustic voice analysis - Part II: acoustical analysis, results and discussion. TMH-QPSR. Speech Transmission Laboratory Quarterly Progress and Status Report 1977;2-3:39-43.

Gauffin J, Sundberg J. Data on glottal voice source in vowel production. TMH-QPSR. Speech Transmission Laboratory Quarterly Progress and Status Report 1980;2-3:61-70.

Gerratt BR, Kreiman J, Antonanzas-Barroso N, Berke GS. Comparing internal and external standards in voice quality judgments. J Speech Hear Res 1993; 36:14-20.

Gerratt BR, Kreiman J. Theoretical and methodological development in the study of pathological voice quality. J Phonetics 2000; 28:335-42.

Gould WJ, Korovin GS. Laboratory advances for voice measurements. J Voice 1994; 8:8-17.

Gregio FN. Análise acústica da voz: pistas para avaliação da eficácia das compensações desenvolvidas num grupo de indivíduos com limitações da ação glótica [trabalho de conclusão de curso]. São Paulo: Pontifícia Universidade Católica de São Paulo; 2000.

Hammarberg B. Clinical applications of methods for acoustic voice analysis-brief research report. Int J Rehab Research 1980; 3:548-9.

Hammarberg B, Fritzell B, Gauffin J, Sundberg J, Wedin L. Perceptual and acoustic correlates of abnormal voice qualities. Acta Otolaryngol 1980; 90:441-51.

Hammarberg B, Gauffin B, Schirratzki H. Teflon injection in 16 patients with paralytic dysphonia: perceptual and acoustic evaluations. J Speech Hear Disord 1984;49:72-82.

Hammarberg B, Fritzell B, Gauffin J, Sundberg J, Wedin L. Acoustic and perceptual analysis of vocal dysfunction. J Phonet 1986; 14:533-47.

Hammarberg B, Gauffin J. - Perceptual and acoustics characteristics of quality differences in pathological voices as related to physiological aspects. In: Fujimura O, Hirano M. Vocal fold physiology. San Diego: Singular Publishing Group Inc; 1995. p.283-303.

Hammarberg B. Voice research and clinical needs. Folia Phoniat Logop 2000; 52:93-102.

Hanson HM, Stevens KN. Subglottal ressonances in female speakers and their effect on vowel spectra. In: 13<sup>th</sup> International Congress of Phonetic Sciences; 1995; Stockholm. Proceedings. Stockholm: ICPhS; 1995; 3:182-5.

Hartmann E, Cramon D. Acoustic measurements of voice quality in dysphonia after severe closed head trauma: a follow-up study. British J Commun 1984;19:253-61.

Hayward K Experimental phonetics. Edinburgh. Pearson Education Limited. 2000.

Heman-Ackah YD, Michael DD, Goding Jr GS. The relationship between cepstral peak proeminence and selected parameters of dysphonia. J Voice 2002;16:20-7.

Higgins MB, Netsell R, Schulte L. Vowel-related differences in laryngeal articulatory and phonatory function. JSLHR 1998; 41:712-24.

Hillenbrand J. A methodological study of perturbation and additive noise in synthetically generated voice signals. J Speech Hear Res 1987; 30:448-61.

Hillenbrand J, Cleveland RA, Erickson RL. Acoustic correlates of vocal quality. J Speech Hear Res 1994; 37:769-78.

Hirano M, Hiki S, Imaizzumi S, Kakita Y, Matsushita H. Acoustical analysis of pathological voice. In: 7<sup>th</sup> Symposium Care of the Professional Voice; 1978; New York City. Transcripts. New York City;1978: 5-57.

Hirano M. Clinical examination of voice. New York: Springer-Verlag Wien; 1981.

Hirano M, Hibi S, Terasawa R, Fujiu M. Relationship between aerodynamic, vibratory, acoustic and psychoacoustic correlates in dysphonia. J Phonetics 1986; 14:445-56.

Hirano M, Bless DM. Exame videoestroboscópico da laringe. São Paulo: Artes Médicas; 1993.

Hirano M, Mori K, Tanaka S, Fujita M. Vocal function in patients with unilateral vocal fold paralysis before and after silicone injection. Acta Otolaryngol (Stockh) 1995; 115:553-9.

Hirano M, Mori K. Vocal fold paralysis. In: Kent RD, Ball MJ. Voice quality measurement. San Diego: Singular Publishing Group Inc; 2000. p.385-95.

Hiraoka N, Kitazoe Y, Ueta H, Tanaka S, Tanabe M. Harmonic-intensity analysis of normal and hoarse voices. J Acoust Soc Am 1984; 76:1648-51.

Hoit JD, Solomon NP, Hixon TJ. Effect of lung voice volume on voice onset time (VOT). J Speech Hear Res 1993; 36:516-21.

Hollien H, Michel J, Doherty ET. A method for analyzing vocal shimmer jitter in sustained phonation. J Phonetics 1973; 1:85-91.

Hollien, H. The concept of ideal voice quality. In: Kent RD, Ball MJ. Voice quality measurement. San Diego: Singular Publishing Group Inc; 2000. p.13-24.

Holmberg EB, Hillman RE, Perkell JS, Guiod PC, Goldman SL. Comparisons among aerodynamic, electroglottographic and acoustic spectral measures of female voice. J Speech Hear Res 1995;38:1212-23.

Honikman B. Articulatory settings. In: Abercrombie D, Fry DB, McCarthy PAD, Scott NC, Trim JLM. In honour of Daniel Jones. Papers contributed on the occasion of his eightieth birthday. London: Longmans; 1964. p.73-84.

Horii Y. Fundamental frequency perturbation observed in sustained phonation. J Speech Hear Res 1979; 22:5-19.

Houben GB, Buekers R, Kingma H. Characterization of the electroglottographic waveform: a preliminary study to investigate vocal fold functioning. Folia Phoniat 1992; 44:269-81.

Isshiki N, Okamura H, Tanabe M, Morimoto M. Differential diagnosis of hoarseness. Folia Phoniat 1969; 21:9-19.

Isshiki N. Fisiologia de produção de fala. In: Isshiki N, Tsuji DH, Sennes LU. Tireoplastias. São Paulo: Fundação Otorrinolaringologia; 1999a. p.21-30.

Isshiki N. Fisiopatologia da produção da voz. In: Isshiki N, Tsuji DH, Sennes LU. Tireoplastias. São Paulo: Fundação Otorrinolaringologia; 1999b p.31-7.

Iwata S, Von Leden H. Phonation quotient in patients with laryngeal diseases. Folia Phoniat 1970; 22:117-28.

Karnell MP. Syncronized videostroboscopy and electroglottography. J Voice 1989; 3:68-75.

Kasuya H, Ogawa S, Masima K, Ebihara S. Normalized noise energy as an acoustic measure to evaluate pathologic voice. J Acoust Soc Am 1986; 80:1329-44.

Kelman AW. Vibratory pattern of the vocal folds. Folia Phoniat 1981; 33:73-99.

Kent RD, Read C. The acoustic analysis of speech. San Diego: Singular Publishing Group Inc; 1992.

Kent RD. Vocal tract acoustics. J Voice 1993; 7:97-117.

Kent RD. The speech sciences. San Diego: Singular Publishing Group Inc; 1997.

Kent RD, Ball MJ. Voice quality measurement. San Diego: Singular Publishing Group Inc; 2000.

Kim KM, Kakita Y, Hirano M. Sound spectrographic analysis of the voice of patients with recurrent laryngeal nerve paralysis. Folia Phoniat 1982; 34:124-33.

Kitajima K. Quantitative evaluation of the noise in the pathologic voice. Folia Phoniat 1981; 33:115-24.

Kitzing P. Stroboscopy - a pertinent laryngological examination. J Otolaryngol 1985; 14:151-7.

Kitzing P. LTAS criteria pertinent to the measurement of voice quality. J Phonetics 1986; 14:477-82.

Klatt DH, Klatt LC. Analysis, synthesis, and perception of voice quality variations among female and male talkers. J Acoust Soc Am 1990; 87:820-57.

Klingholz F, Martin F. Quantitative spectral evaluation of shimmer and jitter. J Speech Hear Res 1985; 28:169-74.

Koenig W, Dunn HK, Lacy LY. The sound spectrograph. J Acoust Soc Am 1946; 18:19-49.

Kohen EM. Impedancia acústica – puentes de medición. In: \_\_\_\_\_. Impedancia acustica. Buenos Aires: Panamericana; 1985. p.43-50.

Kohler K. Phonetics: a language science in its own right? In: 13<sup>th</sup> Congress of Phonetic Sciences; Estocolmo. ICPhS. Proceedings. Stockholm; 1995; 1:10-17.

Kojima H, Gould WJ, Lambiase A, Isshiki N. Computer analysis of hoarseness. Acta Otolaryngol 1980; 89:547-54.

Kreiman J, Gerratt BR. Perceptual reality of voice quality rating scales. ASHA 1995; 37:42.

Kreiman J, Gerratt BR. The perceptual structure of pathologic voice quality. J Acoust Soc Am 1996; 100:1787-95.

Kreiman J, Gerratt B. Measuring voice quality. In: Kent RD, Ball MJ. Voice quality

measurement. San Diego: Singular Publishing Group Inc; 2000. p.73-101.

Ladefoged P. The features of the larynx. J Phonetics 1973; 1:73-83.

Ladefoged P. Elements of acoustic phonetics. Chicago: University of Chicago Press; 1996.

Larsson H, Hertegard S, Lindestad P, Hammarberg B. Vocal fold vibrations: high-speed video imaging and acoustic analysis. TMH-QPSR. Speech Transmission Laboratory Quarterly Progress and Status Report 1999:21-9.

Larsson H, Hertegard S, Lindestad P, Hammarberg B. Vocal fold vibrations: high-speed video imaging and acoustic analysis. Laryngoscope 2000; 110: 2117-22.

Laver J. The concept of articulatory settings: an historical survey. Historiographia Linguistica 1978; 5:1-14.

Laver J. The description of voice quality in general phonetic theory. Edinburgh University Department of Linguistic work in progress 1979; 12:3-52.

Laver J. The phonetic description of voice quality. Cambridge: Cambridge University Press; 1980.

Laver J. The analysis of vocal quality: from the classical period to the twentieth century. In: Asher RE, Henderson EJA. Towards a history of phonetics. Edinburgh: Edinburgh University press; 1981. p.79-99.

Laver J, Wirs S, Mackenzie J, Hiller SM. A perceptual protocol for the analysis of vocal profiles. Edinburg University Department of Linguistics Work in Progress 1981;14:139-55.

Laver J, Hiller S, Hanson R. Comparative performance of pitch detection algorithms on dysphonic voices. In: International Conference on Acoustics, Speech an Signal Processing; 1982; Paris. IEEE .Proceedings. Paris: 1982:192-5.

Laver J. The gift of speech. Edinburgh: Edinburgh University Press; 1991.

Laver J. Principles of phonetics. New York: Cambridge University Press; 1994.

Laver J. Phonetic evaluation of voice quality. In: Kent RD, Ball MJ. Voice quality measurement. San Diego: Singular Publishing Group Inc; 2000. p.37-48.

Lecluse FLE, Brocaar MP, Verschure J. The electroglottography and its relation to glottal activity. Folia Phoniat 1975; 27:215-24.

Lieberman P. Some acoustic measures of fundamental periodicity of normal and pathologic larynges. J Acoust Soc Am 1963; 35:344-53.

Lindblom EF, Sundberg JEF. Acoustical consequences of lip, tongue, jaw, and larynx movement. J Acoust Soc Am 1971; 50:1166-79.

Lindestad PA, Persson A. Quantitative analysis pattern with laryngeal paralysis. Acta Otolaryngol (Stock) 1994; 114:91-7.

Lindsay RB. The story of acoustics. J Acoust Soc Am 1966; 39:629-44.

Lindsey LA, Vieira MN. Modulation analysis - a new method for modelling and measurement of both the production and perception of vocal fold cycle perturbation. ESCA 1997: 65-8.

Linville SE, Rens J. Vocal tract ressonance analysis of aging voice using long-term average spectrum. J Voice 2001; 15:323-30.

Lisker L, Abramson AS. A cross-language study of voicing in initial stops: acoustical measurements. Word 1964; 20:384-422.

Livesey JR, Carding PN. An analysis of vocal cord paralysis before and after Teflon injection using combined glottography. Clin Otolaryngo1995; 20:423-7.

Löfqvist A. The long-time-average spectrum as a tool in voice research. J Phonetics 1986; 14:471-5.

Löfqvist A, Mandersson B. Long-time average spectrum of speech and voice analysis. Folia Phoniat (Basel) 1987; 39:221-9.

McAllister A, Sederholm E, Ternström S, Sundberg J. Perturbation and hoarseness: a pilot study of six children's voices. J Voice 1996; 10:252-61.

Madureira S, Barzaghi L, Mendes B. Voicing contrasts and the deaf: production and perception issues. In: Windsor F, Kelly ML, Hewlett N. Investigations in clinical phonetics and linguistics. Mahwah: Lawrence Erlbaum Associates Inc; 2002. p.417-35.

Mancini PC. Estudo da oclusiva glottal e do ataque vocal brusco em falantes do português [dissertação]. Minas Gerais: Universidade Federal de Minas Gerais; 2001.

Martin D, Fitch J, Wolfe V. Pathologic voice type and the acoustic prediction of severity. J Speech Hear Res 1995; 38:765-71.

Mendonza E, Valencia N, Muñoz J, Trujillo H. Differences in voice quality between men and women: use of the long-term average spectrum (LTAS). J Voice 1996; 10:59-66.

Moore P, Thompson CL. Comments on physiology of hoarseness. Arch Olaryngol 1965; 81:97-102.

Naranjo NV, Lara EM, Rodríguez IM, Garcia GC. High-frequency components of normal and dysphonic voices. J Voice 1994; 8:157-62.

Nolan F. The phonetic bases of speaker recognition. Cambridge: Cambridge University Press; 1983.

Ondrácková J. Vocal cord activity - Its dynamics and role in speech production. Folia Phoniat 1972; 24:405-19.

Ooi LL, Chan HS, Soo KC. Color Doppler imaging for vocal cord palsy. Head Neck 1995; 17:20-3.

Orlikoff RF. Assessment of the dynamics of vocal fold contact from the electroglottogram: data from normal male subjects. J Speech Hear Res 1991; 34:1066-72.

Orlikoff RF. Vocal stability and vocal tract configuration: an acoustic and electroglottographic investigation. J Voice 1995;9:173-81.

Orlikoff RF. Scrambled EGG: the use and abuses of electroglottography. Phonoscope 1998; 1:37-53.

Osgood CE, Suci GJ, Tannenbaum PH. The measurement of meaning. In: Snider JG, Osgood CE. Semantic differential technique: a sourcebook. Chicago: Aldine Publishing Company; 1969. p56-82.

Piccirillo JF, Painter C, Fulller D, Haiduk A, Fredrickson JM. Assessment of two objective voice function indices. Ann Otol Rhinol Laryngol 1998; 107:396-400.

Pickett JM. The spectra of vowels. In: \_\_\_\_\_. The sounds of speech communication. Baltimore: University Park; 1980. p.57-78.

Pinho SR, Pontes, P, Gadelha MEC, Biasi N. Vestibular vocal fold behavior during phonation in unilateral vocal fold paralysis. J Voice 1999; 13:36-42.

Pinho SR, Pontes P. Escala de avaliação perceptiva da fonte glótica: RASAT. Vox Brasilis 2002; 3:11-3.

Pittam J. Discrimination if five voice qualities and prediction to perceptual ratings. Phonetica 1987; 44:38-49.

Pittam J, Millar JB. The long-term spectrum of voice. In: 2<sup>nd</sup> Australian international conference on speech science and technology; 1988; Sydney. Proceedings. Sydney; 1988. p.308-13.

Pittam J. Voice in social interaction: an interdisciplinary approach. London: SAGE Publications; 1994.

Qi Y. Time normalization in voice analysis. J Acoust Soc Am 1992; 92:2569-76.

Rabinovic CR, Kreiman J, Gerratt BR, Bielamowicz S. Comparing reliability of perceptual ratings and acoustic measures of jitter. J Speech Hear Res 1995; 38:26-32.

Rontal E, Rontal M, Rolnick MI. Objective evaluation of vocal pathology using voice spectrography. Ann Otol 1975; 84:662-72.

Rothenberg M, Mashie JJ. Monitoring vocal fold contact abduction through vocal fold contact area. J Speech Hear Res 1988; 31:338-51.

Rothenberg M. Correcting low-frequency phase distortion in electroglottograph waveforms. Technnical note. J Voice 2002; 16:32-6.

Ryalls J, Gustafson K, Santini C. Preliminary investigation of voice onset time production in persons with dysphagia. Dysphagia 1999; 14:169-75.

Rydell R, Schalén L, Fex S, Elner A. Voice evaluation before and after laser excision vs. radiotherapy of T1a glottic carcinoma. Acta Otolaryngol (Stockh) 1995; 115:560-5.

Sasaki Y, Okamura H, Yumoto E. Quantitative analysis of hoarseness using a digital sound spectrograph. J Voice 1991; 5:36-40.

Sawashima M, Hirose H. New laryngoscopic technique by use of fiber optics. J Acoust Soc Am 1968; 43:168-9.

Scherer RC, Vail VJ, Guo CG. Required number of tokens to determine representative voice perturbation values. J Speech Hear Res 1995; 38:1260-9.

Shadle CH. Articulatory-acoustic relations in fricative consonants. In: Hardcastle WJ, Marchal A. Speech production and speech modeling. Dordrecht: Kluwer Academic Publishiers; 1990. p.187-209.

Shewell J. The effect of perceptual training on ability to use the vocal profile analysis scheme. Int J Lang Commun Disord 1998; 33:322-6.

Shimizu K. A cross-language study of voicing contrasts of stop consonants in asian languages. Tokyo: Seibido Publishing Co; 1996.

Shoji K, Regenbogen E, Yu JD, Blaugrund SM. High-frequency components of normal voice. J Voice 1991; 5:29-35.

Smith S. Research on the principle of electroglottography. Folia Phoniat 1981; 33:105-14.

Södersten M, Lindestad PA. A comparison of vocal fold closure in rigid telecopic and flexible fiberoptic laryngostroboscopy. Acta Otolaryngol 1992; 112:144-50.

Sorensen D, Horii Y, Leonard RJW. Effects of laryngeal topical anesthesia on fundamental frequency perturbation. J Speech Hear Res 1980; 23:253-9.

Sorensen D, Horii Y. Directional perturbation factors for jitter and for shimmer. J Commun Disord 1984; 17:143-51.

Spencer ML, Titze IR. An investigation of a modal-falsetto register transition hypothesis using helox gas. J Voice 2001; 15:15-24.

Stevens KN, House AS. An acoustical theory of vowel production and some of its implications. J Speech Hear Res 1961; 4:303-20.

Stevens KN. Acoustic phonetics. Cambridge: MIT Press; 1998.

Sundberg J. Articulatory interpretation of the "singing formant". J Acoust Soc Am 1974; 55:838-44.

Sundberg J. The science of the singing voice. Dekalb: Northern Illinois University Press; 1987.

Sundberg J. Level and center frequency of the singer's formant. STL-QPSR. Speech Transmission Laboratory Quartely Progress and Status Report 1999; 3-4:87-94.

Švec JG, Schutte HK. Videokymography: high-speed line scanning of vocal fold vibration. J Voice 1996; 10:201-5.

Švec JG, Schutte HK, Miller DG. A subharmonic vibratory pattern in normal vocal folds. J Speech Hear Res 1996; 39:135-43.

Syrdal AK, Gopal HS. A perceptual model of vowel recognition based on the auditory representation of American english vowels. J Acoustic Soc Am 1996; 79:1086-100.

Takahashi H, Koike Y. Some perceptual dimension and acoustic correlates of pathologic voices. Acta Otolaryngol 1975; Suppl 338: 3-24.

Timcke R, Von Leden H, Moore P. Laryngeal vibrations of the glottic wave - part I. The normal vibratory cycle. AMA Arch Otolaryngol 1958; 68:1-19.

Titze IR. Parametrization of the glottal area, glottal flow, and vocal fold contact area. J Acoust Soc Am 1984; 75:570-80.

Titze IR, Horii Y, Scherer RC. Some technical considerations in voice perturbation measurements. J Speech Hear Res 1987; 30:252-60.

Titze IR. Physiologic and acoustic differences between male and female voices. J Acoust Soc Am 1989; 85:1699-707.

Titze IR. Toward standards in acoustic analysis of voice. J Voice 1994; 8:1-7.

Titze IR. Workshop on acoustic voice analysis. National Center for voice and speech 1995: 1-36.

Toner MA, Emanuel FW, Parker D. Relationship of spectral noise levels to psychophysical scaling of vowel roughness. J Speech Hear Res 1990; 33:238-44.

Tsuji DH, Sennes LU. Videoquimografia da laringe: novo método de avaliação da vibração cordal. Arquivos da Fundação Otorrinolaringologia 1998; 2:136-40.

Tsuji DH, Sennes LU, Figueiredo LAP, Iamamura R. Avaliação endoscópica da laringe. In: Isshiki N, Tsuji DH, Sennes LU. Tireoplastias. São Paulo: Fundação Otorrinolaringologia; 1999a. p.51-8.

Tsuji DH, Sennes LU, Pinho SMR, Figueiredo LAP, Iamamura R. Videostroboscopia e videoquimografia da Iaringe. In: Isshiki N, Tsuji DH, Sennes LU. Tireoplastias. São Paulo: Fundação Otorrinolaringologia; 1999b. p.59-67.

Tsuji DH, Sennes LU, Pinho SMR. Videoquimografia da laringe. In: Pinho SMR. Tópicos em voz. Rio de Janeiro: Guanabara-Koogan; 2001. p.45-52.

Vallencien B, Faulhaber J. Causes of error in glottography. J Fr Otorhinolaryngol Chir Maxillofac 1966; 15:383-94.

Vallencien B, Faulhaber J. Lo que debe pensarse de la glotografia. Acta Otolaryngol lber Amer 1967; 18:373-8.

Vieira MN, McInnes FG, Jack MA. Robust fo and jitter estimation in pathological voices. In: 4<sup>th</sup> International Conference on Spoken Language Processing; 1996; Philadelphia. Proceedings. Philadelphia: ICSLP; 1996a. p.745-8.

Vieira MN, Maran AG, McInnes FG, Jack MA. Detecting arytenoids cartilage misplacement through acoustic and electroglottographic jitter analysis. In: 4<sup>th</sup> International Conference on Spoken Language Processing; 1996; Philadelphia. Proceedings. Philadelphia: ICSLP; 1996b. p.741-4.

Vieira MN. Automated measures of dysphonias and the phonatory effects of asymmetries in the posterior larynx [thesis]. Edinburgh: University of Edinburgh; 1997.

Vieira MN, McInnes FR, Jack MA. Comparative assessment of electroglottographic and acoustic measures of jitter in pathological voices. J Speech Hear Res 1997; 40:170-82.

Vieira MN, McInnes FR, Jack, MA. On the influence of laryngeal pathologies on acoustic and electroglottographic jitter measures. J Acoust Soc Am 2002; 111:1045-55.

Watson C. Quality analysis of laryngography in a busy hospital ENT voice clinic. J Europ J Disord Com 1995; 30:132-9.

Wedin S, Ögren J. Analysis of the fundamental frequency of the human voice and its frequency distribution before and after a voice training program. Folia Phoniat 1982; 34:143-9.

Wendler J, Rauhut A, Krüger H. Classification of voice qualities. J Phonetics 1986; 14:483-8.

White P. Long-term average spectrum (LTAS) analysis of sex and gender-related differences in children's voices. Log Phon Vocol 2001; 26:97-101.

Wilson DK. Problemas de voz em crianças. 3ª ed. São Paulo: Manole; 1994.

Wolfe V, Martin D. Acoustic correlates of dysphonia: type and severity. J Commun Disord 1997; 30:403-15.

Wuyts FL, De Bodt M, Bruckers L, Molenberghs G. Recent work of the Belgian study group of voice disorders. Results. Acta Otorhinolaryngol Belg 1996; 50:331-41.

Yanagihara N. Hoarseness: investigation of the physiological mechanisms. Ann Otol Rhinol Laryngol 1967; 76:472-88.

Yumoto E, Gould W, Bauer T. Harmonics-to-noise ratio as an index of the degree of hoarseness. J Acoust Soc Am 1982; 71:1544-50.

Yumoto E. The quantitative evaluation of hoarseness. Arch Otolaryngol 1983; 109:48-52.

Yumoto E, Sasaki Y, Okamura H. Harmonics-to-noise ratio and psychophysical measurement of the degree of hoarseness. J Speech Hear Res 1984; 27:2-6.

Zwicker E, Fastl H. Psycho-acoustics: facts and models. 2<sup>nd</sup> ed. Berlin: Springer-Verlag Wien; 1999.

## Bibliografia Consultada

Rother ET, Braga ME. Como elaborar sua tese: estrutura e referências. São Paulo: 2001.