

**DIMENSÕES PERCEPTIVAS DAS ALTERAÇÕES DE
QUALIDADE VOCAL E SUAS CORRELAÇÕES AOS
PLANOS DA ACÚSTICA E DA FISIOLOGIA**

(Perceptual dimensions of voice disorders and their correlations to
acoustical and physiological arenas)

Zuleica Antonia de CAMARGO

(Departamento de Lingüística/ LAEL – PUC-SP;
Laboratório Integrado de Análise Acústica e Cognição (LIAAC)).

Sandra MADUREIRA

(Departamento de Lingüística/ LAEL – PUC-SP;
Laboratório Integrado de Análise Acústica e Cognição (LIAAC)).

ABSTRACT: *This study departs from the necessity of understanding compensatory mechanisms used by dysphonic individuals. The objective was to investigate dysphonic voices by taking into account acoustic analysis (long and short term analysis), physiological examination (electroglottography-EGG, laryngeal endoscopy, videostroboscopy and videokymography) and perceptual evaluation (voice quality and speech recognition) procedures. From the perceptual point of view, both the usefulness of the phonetically oriented vocal profile and the speech recognition task have been validated. Some perceptual dimensions correlated to each other, with correspondences to acoustical and physiological descriptions. From the acoustic point of view results indicated the relevance of long-term events and their correspondence to those of short term. Data from EGG showed correspondence to the results of otolaryngologic examination. The results favors focus on the multiple clinical manifestations related to dysphonia so as to instruct rehabilitation practices in clinical settings, based on acoustic-to articulatory and perceptual correlates.*
KEY-WORDS: *voice; voice disorders; acoustics, speech; auditory perception; speech production measurement.*

RESUMO: *O presente estudo fundamenta-se na necessidade de compreensão de mecanismos compensatórios usados por indivíduos portadores de alterações da qualidade vocal (disfonias). O objetivo foi investigar amostras de vozes com alterações da qualidade*

vocal por procedimentos perceptivo-auditivos (qualidade vocal e reconhecimento de fala), acústicos (análise de curto e de longo termo) e fisiológicos (eletroglotografia-EGG, videolaringostroboscopia e videoquimografia). Do ponto de vista perceptivo-auditivo, foi destacada a validade do uso de roteiro com embasamento fonético e de teste de reconhecimento de fala. Algumas dimensões perceptivas mostraram correlações, as quais encontram respaldo nas descrições acústicas e fisiológicas. Os resultados acústicos indicaram a relevância dos eventos de longo termo e sua correspondência àqueles de curto termo. Dados de EGG revelaram correspondência aos resultados do exame otorrinolaringológico. Os resultados evidenciam a multiplicidade de manifestações clínicas descritas sob a terminologia “disfonia”, de forma a revelar as correlações entre os planos perceptivo, acústico e fisiológico da qualidade vocal.

PALAVRAS-CHAVE: voz; distúrbios da voz; acústica da fala; percepção auditiva; medida da produção da fala.

Introdução

Avanços das Ciências da Fala e da Tecnologia de Fala tiveram um grande impacto nos estudos referentes às alterações da qualidade vocal (disfonias). Como resultado de tal desenvolvimento, muitas informações sobre a função vocal foram obtidas, com destaque para as investigações por meio de procedimentos acústicos e de técnicas de imagens. Um grande número de estudos investigou a atividade glótica como o principal correlato da qualidade vocal e procurou por índices acústicos em estudos de população, com destaque para o fato de que, em sua maioria, procuraram estabelecer uma relação linear entre determinada qualidade sonora e uma medida ou índice acústico (Kojima et al, 1980; Kitajima, 1981; Yumoto et al, 1982; Hiraoka et al, 1984; Klingholz, Martin, 1985; Kasuya et al, 1986; Hillenbrand, 1987; Cox et al, 1989; Sasaki et al, 1991; Qi, 1992; Deliyski, 1993; Awan, Frenkel, 1994; Kent et al, 1999; Smits et al, 2005; Uloza et al, 2005).

De forma oposta, o presente estudo centrou-se nos ajustes individuais que indivíduos disfônicos implementam em seu aparelho fonador durante a produção da fala. Como seu foco está na variabilidade individual, prioriza uma abordagem integrativa de dados de várias fontes, a fim de estabelecer correlações entre parâmetros perceptivo-auditivos, acústicos e fisiológicos. A teoria acústica da produção da fala (Fant, 1970) e o modelo

fonético de descrição da qualidade vocal (Laver, 1980) foram considerados para que tal abordagem fosse delineada, de forma a compor as correlações anteriormente apontadas.

Diante do referencial teórico adotado, o investimento em uma reflexão sobre a dimensão da alteração da qualidade vocal fez-se necessário, no sentido de demandar uma mudança de conceitualização do que é descrito como alteração de qualidade vocal (disfonia). 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 (Camargo, 2002).

A motivação para composição do presente estudo emergiu da tentativa de compreensão da individualidade das manifestações sonoras em casos de disfonia. Para realizá-lo, foram contempladas formas de análise que a Tecnologia da Fala oferece, mais especificamente a análise acústica e eletroglotográfica (EGG), e partiu-se para a reflexão sobre certos desdobramentos do conhecimento teórico construído pelas chamadas Ciências da Fala, entre elas, especialmente, a Fonética.

O 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 (Mackenzie-Beck, 2005).

Este estudo propôs-se a investigar a qualidade vocal em seus aspectos fonéticos (perceptivos, acústicos 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. 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 eletroglotográfica das amostras de fala.

Não se almejou alcançar generalizações sobre a produção sonora na situação de disfonia. Em contraposição, procurou-se ilustrar as várias facetas da qualidade vocal presentes num grupo de indivíduos disfônicos, com

fator desencadeante comum, de forma a explorar os fatores que permeiam a interface da complexa relação entre a percepção, a acústica e a fisiologia, propiciando um ensaio teórico para a reflexão das manifestações da ordem da vocalização humana.

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 e pesquisadores: identificar as dimensões auditivas salientes do sinal de fala e interpretá-las como efeitos das diversas mobilizações no aparelho fonador que afetam as situações de comunicação oral.

Para corresponder ao desafio apresentado, a abordagem acústica do sinal representa uma possibilidade de avanço, porém carrega a demanda de exploração da teoria acústica da produção da fala. O modelo fonte-filtro descrito por Fant (1970) possibilita a decomposição do sinal de fala, reverendo 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. O roteiro intitulado *Voice Profile Analysis Scheme (VPAS)* é fruto da descrição da qualidade embasada no referido modelo teórico.

O presente estudo pode 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.

Procurou-se destacar a demanda por pesquisas que considerem a variabilidade dos sintomas vocais, de forma a contemplar a atividade integrada dos diversos segmentos do aparelho fonador. Nesta concepção, torna-se crescente a necessidade de compreensão das combinações de ações implementadas pelo falante, numa abordagem que se distancie do estabelecimento de um padrão de normalidade vocal e da busca por estimativas das manifestações de disfonias exclusivamente em termos de medidas que expressem sua distância em relação à suposta condição ideal de qualidade vocal. Apesar de várias tentativas no sentido de estabelecer tais medidas (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 et al, 1996), alguns autores já apontaram limitações quanto ao estabelecimento de relações biunívocas entre tais dimensões acústica e perceptiva (Askenfelt, Hammarberg, 1986; Eskenazi et al, 1990; Feijoo, Hernandez, 1990; Bielamowicz et al, 1996; McAllister et al, 1996; Yu et al, 2001; Zhang et al, 2005).

Distantes de tais tendências, alguns estudos abordaram correlações entre as esferas perceptivo-auditiva, acústica e fisiológica para alterações e ajustes da esfera laríngea do aparelho, mais precisamente da região glótica, revelando importantes aspectos para avaliação das disfonias (Hammarberg, Gauffin, 1995; Camargo, 1996; Hammarberg, 2000; Blaj et al, 2007). Além disso, série de estudos produzida com *VPAS* na realidade nacional revelou aspectos promissores em termos de avaliação da qualidade vocal, com base em correspondências entre acústica e percepção, os quais dificilmente seriam revelados sem o devido aporte do modelo fonético de descrição da qualidade vocal (Andrade, 2004; Camargo et al, 2004; Peralta, 2005; Blaj et al, 2007; Bonfim et al, 2007).

O objetivo do presente estudo foi investigar as alterações da qualidade vocal (disfonias) tendo como base descrições da análise perceptivo-auditiva (qualidade vocal e reconhecimento de fala) e suas correspondências nas esferas acústica (de curto e de longo termo) e fisiológica (eletroglotografia-EGG, videolaringostroboscopia e videoquimografia).

1. Métodos

O grupo estudado foi composto por quatro falantes do sexo feminino, nomeados i1 a i4, com idades variadas entre 51 a 72 anos, as quais apre-

sentavam algum grau de incompetência glótica (paralisia unilateral ou fibrose de prega vocal), em acompanhamento clínico em instituição hospitalar na cidade de São Paulo. Um indivíduo sem histórico de disfonia, sexo feminino, nomeado iR, 52 anos de idade, foi adotado como referência para os parâmetros acústicos referentes aos dados do português brasileiro (PB).

Cada indivíduo do grupo estudado (i1 a i4) teve amostras de fala e dados de exame otorrinolaringológico coletados em duas sessões no mesmo dia. iR esteve presente apenas na primeira sessão, voltada à coleta das amostras de fala por meios acústico e eletroglotográfico (EGG), realizada no Laboratório de Rádio da Pontifícia Universidade Católica de São Paulo. As amostras de fala registradas incluíram ondas acústicas e EGG captadas simultaneamente (estímulo estéreo) da emissão de vogais (três emissões de [a:], três seqüências de [a/a/a] e fala encadeada (três leituras do mesmo texto). O estímulo para registro de emissão encadeada utilizado referiu-se ao seguinte texto (composto de 209 palavras), elaborado no Laboratório Integrado de Análise Acústica e Cognição (LIAAC-PUCSP):

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.

Uma frase contendo o substantivo “arara” [a8a8©] foi selecionada (sublinhada no texto acima) de forma que a frequência dos harmônicos e dos formantes da vogal [a] pudesse ser aferida. Palavras do PB contendo as consoantes plosivas em sílabas tônicas ([p], [b], [t], [d], [k], [g]) também foram destacadas (sublinhadas no texto acima) para extração da medida de tempo de ataque de vozeamento (sigla *VOT*, de *Voice Onset Time* em sua designação original).

As amostras acústicas foram registradas por meio de um microfone unidirecional com acoplagem em cabeça (*Audiotechnica* ATM 75) posicionado a 3 cm da comissura labial direita do falante. As amostras EGG foram registradas a partir da acoplagem de eletrodos de superfície do eletroglotógrafo EG2 (*Glottal Enterprises*) nas alas da cartilagem tireóide. Procedimentos prévios envolveram a remoção de colares, brincos e outros acessórios metálicos na região da cabeça e do pescoço. Os eletrodos receberam uma camada de gel hipoalergênico (*Spectra 360-Parker Lab*) e uma fita para manter os eletrodos em contato com a cartilagem tireóide. Os sinais acústicos e EGG foram registrados com auxílio da mesa de som *Mackie Microseries* 102 VLZ (12 canais), digitalizados e editados no *software Sound Forge* 4.0.

Do ponto de vista perceptivo-auditivo, a análise envolveu procedimentos voltados aos julgamentos da qualidade vocal e do reconhecimento de fala. A qualidade vocal foi avaliada em dois passos por quatro juízes. O primeiro consistiu na indicação dos graus de alteração vocal e de agradabilidade (etapa 1A- vogais e etapa 1B- fala encadeada), além do livre julgamento da qualidade vocal. As amostras de fala para avaliação incluíram emissões vocálicas e sentenças, contendo o vocábulo [$\alpha P \alpha P$], extraídas de texto para leitura.

O segundo passo da avaliação perceptivo-auditiva consistiu em julgamentos da qualidade vocal a partir de modelo fonético (Laver, 1980), com base na descrição de ajustes laríngeos e supralaríngeos, numa adaptação do *Vocal Profile Analysis Scheme - VPAS* (Laver et al., 1981; Mackenzie-Beck, 1988; Laver, 2000) para o português (Camargo, 2002).

As amostras incluíram sentenças extraídas da leitura de texto. Para este propósito, foi editado um CD de áudio (65:03 minutos, 9 faixas) contendo informações sobre o modelo fonético de descrição da qualidade vocal e instruções aos quatro juízes previamente selecionados por sua familiaridade ao modelo fonético de descrição da qualidade vocal (Laver, 1980). Os julgamentos foram submetidos a análise de componentes principais e composição de *clusters*.

O teste de reconhecimento de fala teve como estímulos sílabas contendo a consoante inicial de modo de articulação plosivo, as mesmas usadas para extração das medidas de *VOT*. A tarefa foi conduzida por 78 juízes (fonoaudiólogos, pós-graduandos (*lato sensu*) em Voz, Motricidade

Orofacial e Audiologia Clínica). Os dados foram submetidos a análise estatística (teste T de *Student*).

Do ponto de vista da análise acústica, os mesmos estímulos foram submetidos a procedimentos de análise de longo termo (traçados dos espectros de longo termo – ELT e extração de medidas de intensidade em 128 pontos de frequência entre 0 e 11025 Hz) e de curto termo (inspeção acústica de espectrogramas de banda larga, banda estreita, extração de medidas espectrais de frequência e amplitude de dois primeiros harmônicos (H1 e H2), frequência formântica (F1, F2, F3 e F4) e de *VOT*).

As medidas acústicas de longo termo (ELT) e de frequência de formantes foram submetidas a tratamento estatístico (análise de componentes principais e composição de *clusters*). As demais medidas espectrais foram analisadas estatisticamente por meio do teste *T de Student*, com nível de significância estabelecido em $p < 0,05$.

As amostras EGG foram inspecionadas simultaneamente às acústicas e também foram submetidas a procedimentos de filtragem para eliminação das flutuações de linha de base, que não são primariamente geradas pela atividade vibratória de pregas vocais. Tais técnicas incluíram procedimentos de filtragem de alta frequência (passa alta – 60 Hz) (Vieira et al., 1996; Vieira, 1997). Na sequência, as ondas filtradas (Lx) foram analisadas e, para aquelas em que o traçado se apresentava melhor definido, segundo escala de avaliação 1 a 4, em ordem crescente de melhora da qualidade do sinal (Vieira, 1997), foram extraídas medidas relativas a f0, coeficiente de contato (CC), *jitter*, *shimmer* e índice de velocidade (IV).

As amostras de dois sujeitos do grupo estudado (i1 e i2) e de iR, as quais receberam escores entre 3 e 4, foram processadas pelos referidos *softwares* por terem atingido os requisitos avaliação. As amostras dos demais sujeitos do grupo estudado, com escores entre 1 e 2, foram analisadas apenas por procedimentos não automáticos devido à aperiodicidade do sinal e a baixa fidedignidade das medidas obtidas pelo extrator automático para tal situação.

Como última etapa da coleta de dados para o grupo estudado (i1 a i4), o exame otorrinolaringológico foi realizado por um médico, especialista em otorrinolaringologia, no mesmo dia, num intervalo máximo de três horas após a gravação do corpora, em clínica privada na cidade de São

Paulo. O equipamento para avaliação incluiu endoscópio rígido 70° *LY-C530 Machida*, estroboscópio *Brüel & Kjaer type 4914* e videoquimógrafo 8900-*Kay Elemectrics Corp.* No caso de exacerbado reflexo, foi utilizado endoscópio flexível ENT 30 P3 *Machida* em associação com o referido videoestroboscópio. O exame foi realizado após a aplicação de xilocaína 10% spray. A função laríngea foi descrita em termos fechamento glótico, periodicidade, simetria de fase e amplitude, onda mucosa, indícios de atividade supraglótica e detalhamento das fases do ciclo vibratório.

Todos os participantes foram informados sobre os propósitos do estudo e a ausência de riscos, tendo consentido o uso das informações coletadas. O projeto de pesquisa foi aprovado pelo Comitê de Ética Médica do Hospital do Servidor Público Municipal sob o número 11/00.

2. Resultados e Discussão

Os resultados referentes aos julgamentos perceptivo-auditivos da qualidade vocal são apresentados nas Figuras 1 a 3. A Figura 1 apresenta os resultados do julgamento dos graus gerais de alteração vocal e de agradabilidade da voz, em que se destaca que o aumento do grau da alteração vocal coincidiu com a diminuição do grau de agradabilidade da emissão.

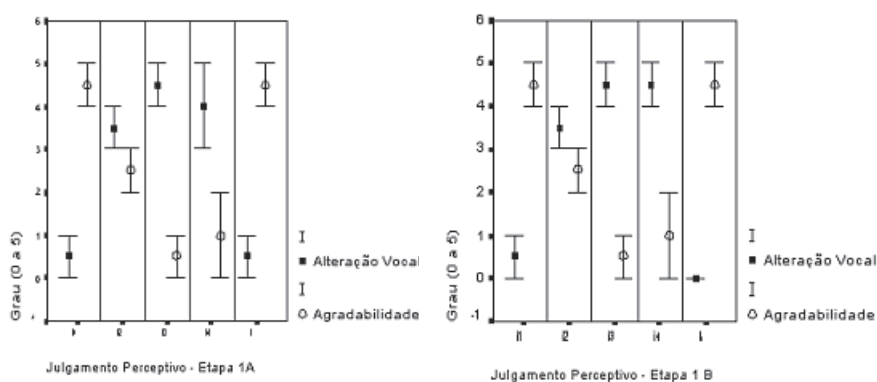


Figura 1: Julgamentos de grau geral de alteração vocal e de agradabilidade atribuídos às amostras de fala dos indivíduos do grupo estudado (I1 a I4) e indivíduo referência (IR) no julgamento da qualidade vocal para emissão vocálica (etapa 1A) e fala encadeada (etapa 1B).

A Figura 2 apresenta os resultados do julgamento livre da qualidade vocal pelo grupo de juízes, novamente para estímulos vocálicos (etapa 1A) e de fala encadeada (etapa 1B), com indicação de ocorrências do plano glótico em sua maioria e relativa coincidência de julgamentos para emissões vocálica e encadeada nos mesmos falantes.

Qualidade Vocal	i1	i2	i3	i4	iR
Normal/adaptada					● □
Modal/neutra					□
Rouca		● □		● □	
Aspera		● □	● □	● □	
Soprosa	● (-) □ (-)	● □	● □	● □	
Diplofônica		□	● □	● □	
Hiponasal	□				□
Hiperfunção	●	● □	●		●
Ressonância Laringo-faríngea	□ (-)				● (-)
Instabilidades	●	● □			

Legenda: ● Etapa 1A □ Etapa 1B (-) referida em menor grau ou por metade dos juízes

Figura 2: Julgamentos de qualidade vocal atribuídos às amostras de fala dos indivíduos do grupo estudado (i1 a i4) e indivíduo referência (iR) no julgamento da qualidade vocal para emissão vocálica (etapa 1A) e fala encadeada (etapa 1B).

A Figura 3 apresenta os resultados referentes aos julgamentos da qualidade vocal a partir do roteiro *VPAS*. Apesar da escassa indicação de ajustes supraglóticos na etapa anterior, observa-se, para os mesmos estímulos (fala encadeada) e os mesmos juízes, a referência a vários ajustes supralaríngeos (articulatórios) longitudinais e transversais, não indicados previamente. Do ponto de vista perceptivo-auditivo, a validade de um roteiro de avaliação da qualidade vocal foneticamente orientado (Laver, 2000) foi confirmada. Quando os juízes foram solicitados a indicar a combinação de ajustes supralaríngeos e laríngeos na caracterização da qualidade vocal das amostras representativas dos indivíduos avaliados, praticamente um terço de suas referências referiu-se à região supraglótica do aparelho fonador.

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

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

Legenda: i1= ▲ i2= ○ i3= ■ i4= ◇ iR= ■ i= ajuste intermitente

Figura 3: Perfil da qualidade vocal segundo julgamentos atribuídos às amostras de fala encadeada dos indivíduos do grupo estudado (i1 a i4) e indivíduo referência (iR) por meio do roteiro *VPAS*.

A análise estatística revelou a separação de três grupos para os julgamentos da qualidade vocal com motivação fonética, nos quais i1, i2 e iR formaram um grupo e i3 e i4 formaram, da um, um grupo distinto. Os parâmetros responsáveis por esta distribuição responderam por praticamente 70% da distribuição (mais precisamente 69,20%). O fator 1 representou 39,69 % de influência na composição, com destaque para os ajustes modal (especialmente presença), escape de ar (especialmente presença) e voz áspera (especialmente ausência). O fator 2 representou 29,51% de

influência na distribuição para os ajustes de hiperfunção (especialmente a presença), constrição faríngea (especialmente presença), falsete (especialmente a presença) e diplofonia (especialmente a presença).

As correlações entre ajustes que influenciaram a composição dos grupos também foram investigadas, de forma que vários ajustes supralaríngeos influenciaram a composição das três classes e, inclusive, para cada um dos indivíduos (grupo estudado e referência); i1 teve sua composição definida basicamente pela presença de ajustes supralaríngeos de mandíbula fechada e dorso de língua recuado; i2 definiu-se pela presença de hiperfunção, laringe alta, aspereza e escape de ar; i3 diferenciou-se pela presença de diplofonia e constrição faríngea; i4 destacou-se dos demais pela presença de laringe baixa e *vocal fry*. iR destacou-se por ausência de aspereza e escape de ar, presença de ajustes modal e labiodentalização. Vale recordar que i1, i2 e iR formaram uma classe, mais distante de cada uma das classes geradas pelos julgamentos das emissões de i3 e i4. Agrupamento semelhante fez-se presente quando da exploração dos dados acústicos, a ser exposta mais adiante. Neste sentido, i1 e i2 mostraram maiores semelhanças entre si e com iR, ao passo que i3 e i4 apresentaram-se mais distantes das demais emissões estudadas e, inclusive, entre si.

Outra informação importante da análise de componentes principais e composição de *clusters* referiu-se à correlação entre presença e ausência de ajustes, os quais indicaram dimensões perceptivas da qualidade vocal do grupo estudado. Foi possível indicar algumas tendências de combinações e outras de oposições de ajustes. No primeiro grupo, destacaram-se:

- ajuste de laringe baixa e ajuste de *vocal fry*/crepitação ausentes *versus* ajuste de laringe baixa e ajuste de *vocal fry*/crepitação presentes;
- ajuste de dorso de língua recuado e ajuste mandíbula fechada presentes *versus* ajuste de dorso de língua recuado e ajuste de mandíbula fechada ausentes;
- ocorrência de curto termo de diplofonia, ajustes de falsete e de constrição faríngea ausentes *versus* ocorrência de curto termo de diplofonia, ajustes de falsete e de constrição faríngea presentes.

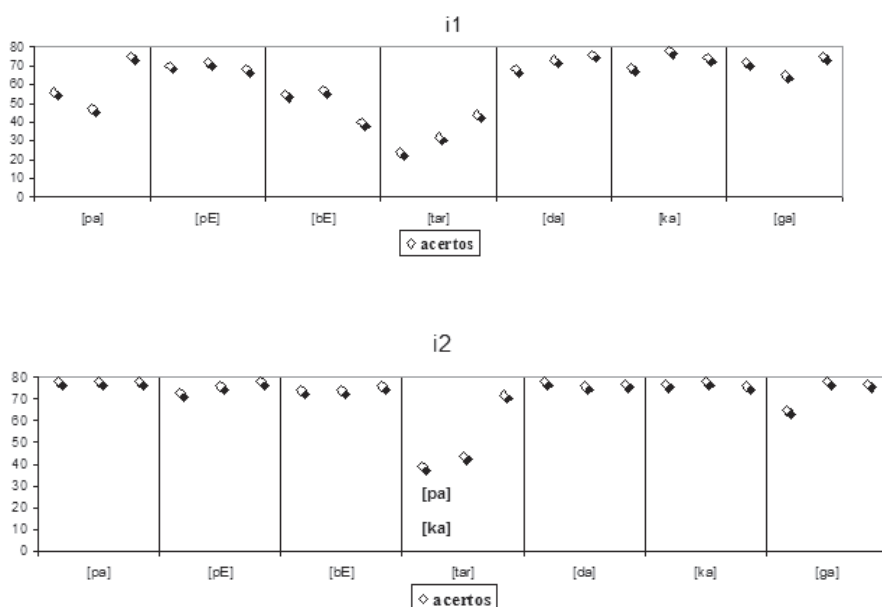
No segundo grupo, o das oposições, destacaram-se ações que apresentaram maior dificuldade, ou mesmo impossibilidade de combinação:

- ajuste de voz áspera ausente e ajuste modal presente *versus* ajuste de voz áspera presente e ajuste modal ausente.

Tais descrições permitiram a discussão mais detalhada a respeito da plasticidade do trato vocal, bem como da interação entre elementos glóticos e supraglóticos na composição da qualidade vocal. Vários dos achados coletados nesta etapa permitem a apreciação de combinação de elementos que possivelmente não se fazem presentes apenas nas vozes com alterações.

3. Reconhecimento de fala

Os dados referentes ao reconhecimento de sílabas são apresentados na Figura 4 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.



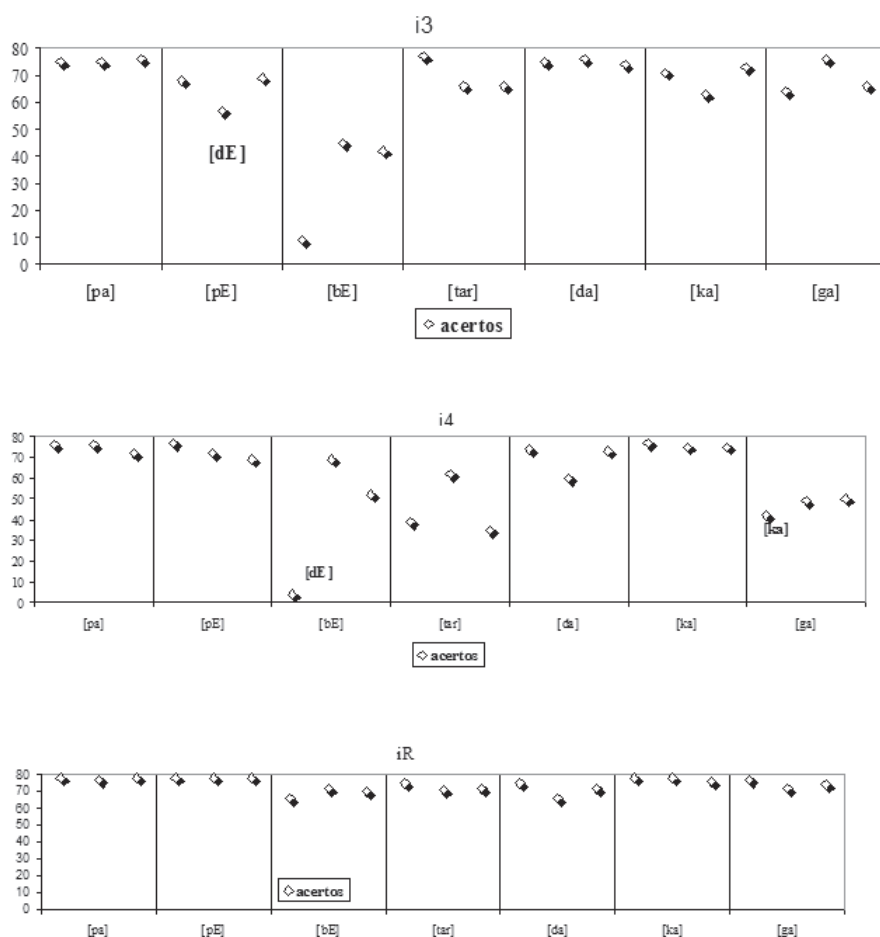


Figura 4: Total de acertos em julgamentos de reconhecimento para as três emissões das sílabas [pa] [pʰ], [bʰ], [taʃ], [da], [ka] e [ga] dos indivíduos do grupo estudado (i1 a i4) e indivíduo referência (iR).

Diferenças estatisticamente significantes ($p < 0,05$) foram identificadas para as sílabas [pa] ($p = 0,046$), [pʰ] ($p = 0,002$), [bʰ] ($p = 0,30$) e [taʃ] ($p = 0,03$) de i1, [pa] ($p = 0,08$), [pʰ] ($p = 0,037$), [bʰ] ($p = 0,040$), [bʰ] ($p = 0,028$) de i3 e [pʰ] ($p = 0,042$), [taʃ] ($p = 0,040$) e [ga] ($p = 0,001$) de i4, todos comparativamente aos valores de iR, adotado como parâmetro de referência para o PB.

Os julgamentos perceptivo-auditivos referentes ao reconhecimento de sons da fala, os quais foram baseados nas amostras de sílabas editadas (dos vocábulos selecionados), revelaram dificuldades relacionadas à identificação de ponto de articulação e à distinção do contraste de vozeamento por parte dos juízes.

4. Análise acústica

Do ponto de vista acústico, os resultados indicaram a relevância dos eventos de longo termo, especialmente aqueles referentes aos traçados dos espectros de longo termo (ELT – Figura 5) e à sua correspondência àqueles da dimensão de curto termo (estrutura harmônica, de formantes e medidas de *VOT*).

A Figura 5 apresenta as médias de intensidade por faixas de frequências dos espectros de longo termo de três emissões de cada falante do grupo estudado e de iR.

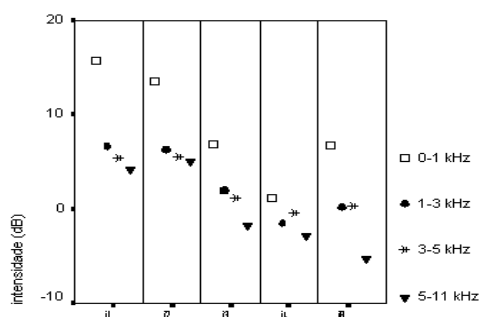


Figura 5: Distribuição das médias de concentrações de amplitude (em dB) por intervalos de frequência (0-1 kHz, 1-3 kHz, 3-5 kHz e 5-11 kHz) dos ELTs de três leituras realizadas pelos falantes do grupo estudado (i1 e i4) e indivíduo referência (iR).

A análise estatística 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 de frequência analisadas para i1 e i2 em relação ao restante dos falantes.

Análise de componentes principais e composição de *clusters* revelou que medidas de longo termo diferenciaram as emissões dos falantes avaliados, uma vez que cada grupo de três emissões de um mesmo falante compôs um *cluster* diferente, reforçando o aspecto de individualidade da qualidade vocal, inclusive na situação de disfonia. A divisão de classes foi semelhante para os aspectos perceptivo-auditivos e acústicos de longo termo, revelando dimensões de correlação possíveis entre percepção e acústica. Tais achados podem sinalizar direções mais confiáveis na busca pelo estabelecimento de correlatos perceptivos e acústicos da qualidade vocal.

A continuidade da exploração dos aspectos acústicos de curto termo procurou detalhar as ocorrências anteriormente esboçadas, as quais permitissem entender como as ações investigadas se organizam na dimensão temporal, de forma a revelar a geração de gestos motores no aparelho fonador de indivíduos com algum grau de limitação da produção vocal.

A Figura 6 apresenta as médias de frequência e intensidade do primeiro (H1) e segundo (H2) harmônicos das emissões estudadas, como forma de detalhar a descrição de eventos (ajustes glóticos/fonatórios) relacionados à qualidade vocal.

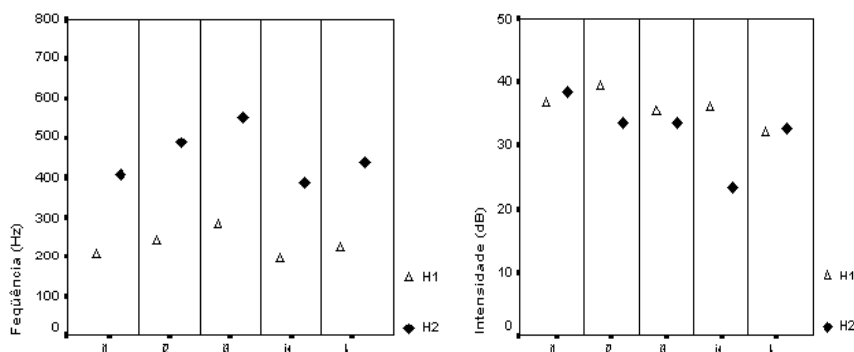


Figura 6: Médias de frequência (Hz) e amplitude (dB) do primeiro (H1) e segundo (H2) harmônicos de três emissões da vogal [a] em posição tônica da emissão de “a arara” em três leituras dos indivíduos do grupo estudado (i1 a i4) e referência (iR).

Os valores de H1(frequência fundamental- f_0) foram mais estáveis (intra-falantes) entre as várias emissões analisadas para i1, i2, i3 i4 e iR no que tange à emissão encadeada, compatíveis com dados de EGG a serem apresentados na seqüência dos resultados e de 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 aspectos de declínio espectral guardam relação com os padrões detectados nos traçados ELT anteriormente abordados.

A Figura 7 apresenta as médias de frequência e intensidade dos quatro primeiros formantes (F1, F2, F3 e F4) das emissões estudadas, como forma de detalhar a descrição de ajustes supraglóticos/articulatórios relacionados à qualidade vocal e ao reconhecimento de fala.

Os falantes avaliados apresentaram valores de frequência de F1 e F2 próximos entre si, dentro dos limites previstos para a vogal [a] no português e mantiveram suas diferenças estáveis entre suas próprias repetições

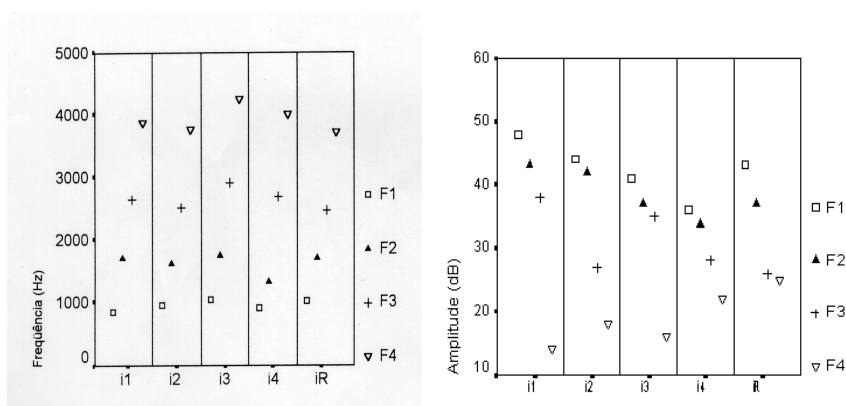


Figura 7: Médias de frequência (Hz) e amplitude (dB) dos quatro primeiros formantes (F1, F2, F3 e F4) da emissão da vogal [a] em posição tônica em “a arara” para três leituras pelos indivíduos do grupo estudado (i1 a i4) e do indivíduo referência (iR).

de emissões, garantindo a manutenção de aspectos de informações relativas ao plano do segmental (qualidade da vogal no PB). Vale recordar que a identidade fonética da vogal encontra correspondência com os valores de F1 e F2 (Fant, 1970). A maior variabilidade interfalantes ocorreu para freqüências de F3 e, especialmente F4, a qual sinaliza a interferência de alguns ajustes individuais, compatíveis com ajustes da qualidade vocal.

A análise estatística das medidas formânticas 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 na 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. Tais resultados indicam que elementos de influência supraglótica também permitem diferenciar os vários falantes do grupo estudado.

A distribuição das medidas de *VOT* é apresentada na Figuras 8 a 10 de forma diferenciada para o grupo de consoantes plosivas desvozeadas e vozeadas do PB, de acordo com as médias calculadas para a emissão de cada

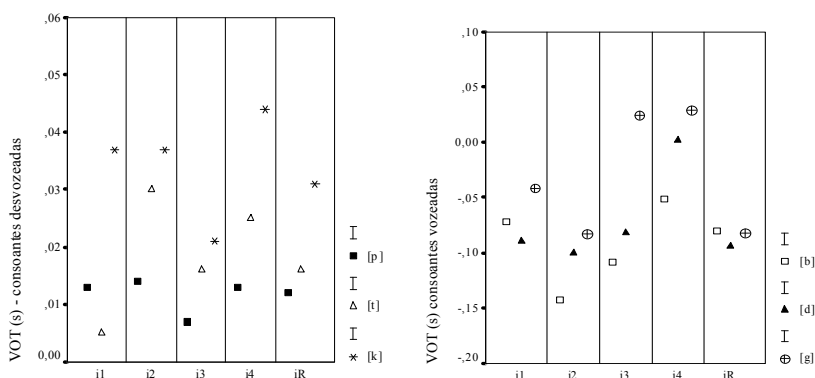


Figura 8: Médias de medidas de *VOT* em segundos de três emissões para as consoantes plosivas desvozeadas [p], [t] e [k] e vozeadas [b] [d] e [g] produzidas pelo grupo estudado (i1 a i4) e indivíduo referência (iR).

uma das sílabas analisadas nas três leituras dos indivíduos do grupo estudado (i1 a i4) e do indivíduo referência (iR).

Os dados referentes à análise das sílabas com sons consonantais plosivos são apresentados separadamente com relação às médias para os sons consonantais desvozeados (Figura 9) e vozeados (Figura 10) do grupo estudado (i1 a i4) e do indivíduo referência (iR).

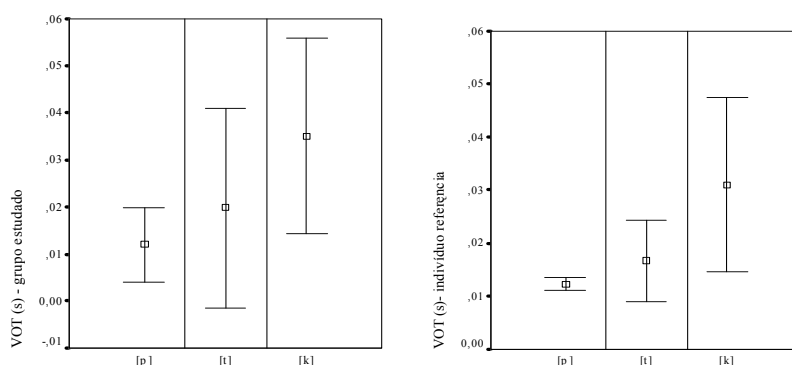


Figura 9: Distribuição das médias dos valores de *VOT* em segundos de consoantes plosivas desvozeadas [p], [t] e [k] nos respectivos vocábulos para, estar e aplicada emitidos pelo grupo estudado (i1 a i4) e indivíduo referência (iR).

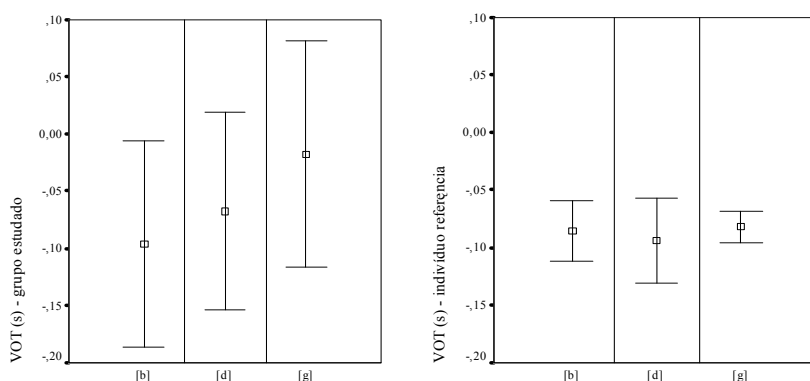


Figura 10: Distribuição das médias dos valores de *VOT* em segundos de consoantes plosivas vozeadas [b], [d] e [g] nos respectivos vocábulos bela, humildade e lugares emitidos pelo grupo estudado (i1 a i4) e indivíduo referência (iR).

A análise estatística revelou diferenças significativas ($p < 0,05$) para as sílabas [ta] ($p = 0,005$) e [ga] ($p = 0,0005$) de i1, [b"] ($0,001$) e [ta] ($p = 0,003$) de i2, [pa] ($p = 0,004$) [ka] ($p = 0,021$) [ga] ($p = 0$) de i3 e [ta] ($0,012$) [da] ($p = 0,0055$) [ka] ($p = 0,041$) e [ga] ($p = 0$) de i4, todos comparativamente aos valores revelados por iR.

5. Achados fisiológicos

Na abordagem dos aspectos fisiológicos, dados da EGG revelaram correspondência com os do exame otorrinolaringológico, especialmente em termos de informações de periodicidade, simetria de movimentos vibratórios, fechamento glótico e a presença de onda mucosa (Figuras 11 a 13).

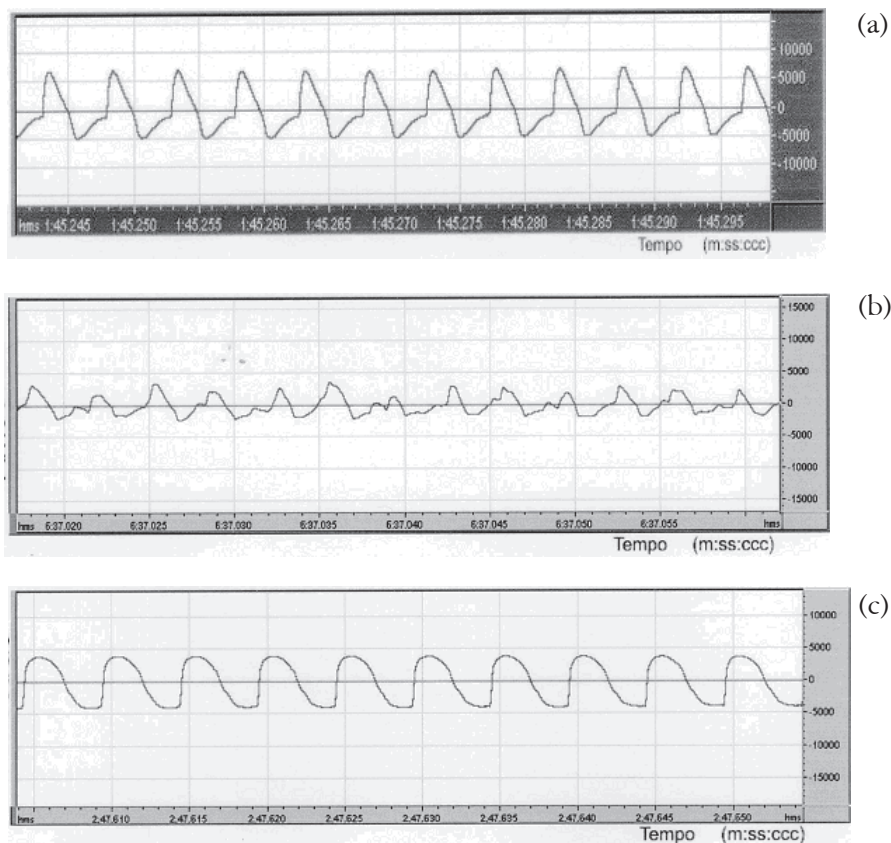


Figura 11: Ondas Lx da vogal [a] de i2 (a) (paralisia de prega vocal esquerda), i3 (b) (fibrose e atrofia de pregas vocais) e iR (pregas vocais sem alterações).

Os espectros de longo termo – ELT também revelaram aspectos relacionados a influências subglóticas do trato vocal que puderam ser observadas na inspeção acústica e indicados pela medida Lx de índice de velocidade (indicativa da natural assimetria de movimentos de lábios superior e inferior da prega vocal no tempo).

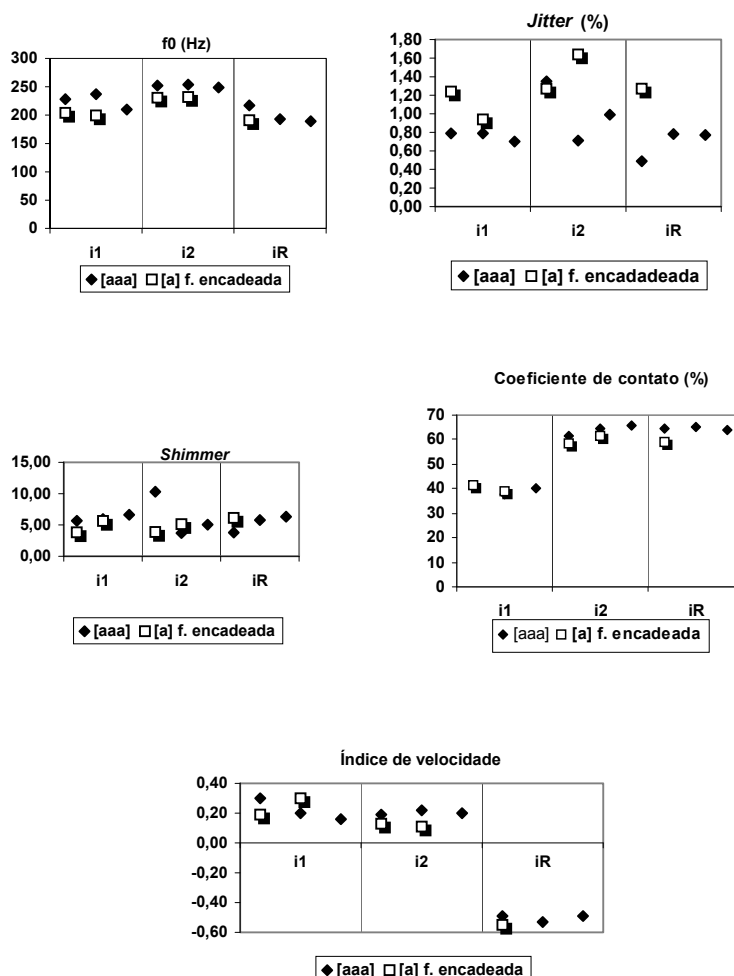


Figura 12: Distribuição de medidas de f0 (Hz), Jitter (%), Shimmer (%), coeficiente de contato (%) e índice de velocidade para três emissões da vogal central das seqüências de emissões [aaa] e duas emissões da vogal tônica das emissões “a arara” para dois indivíduos do grupo estudado (i1 e i2) e indivíduo referência (iR).

Neste campo, a validade da análise da onda Lx (Figura 11) foi ressaltada, graças à possibilidade de uso de instrumentos desenvolvidos na tentativa de superar as flutuações de linha de base da onda do EGG, que poderiam mascarar a real atividade de pregas vocais (Vieira et al., 1996; Vieira, 1997). Foram selecionados traçados de indivíduos representativos dos agrupamentos sinalizados na análise fatorial das etapas anteriores.

Conforme apontado na metodologia, para os casos em que a onda Lx apresentou características de regularidade (i1, i2 e iR), foram extraídas medidas EGG, cujos valores são ilustrados na Figura 12 para medidas referentes a f0, perturbação (*jitter* e *shimmer*), coeficiente de contato e índice de velocidade.

As informações referentes ao exame otorrinolaringológico são apresentadas na Figura 13, com relação aos aspectos gerais estruturais e funcionais, complementados pelos dados dos procedimentos de videolaringo-estroboscopia e de videoquimografia.

Correlações entre dados acústicos, fisiológicos e perceptivo-auditivos na análise das alterações de qualidade vocal.

Os achados foram 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 opuseram simplesmente em termos dos extremos de presença ou ausência de parâmetros vocais, porém partilharam algumas características entre si, diferenciam-se em outras e assim sucessivamente.

A análise estatística por agrupamentos expôs tal particularidade, em que os aspectos perceptivo-auditivos e acústicos 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). 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 (Fant, 1970, Laver, 1980, Mackenzie-Beck, 2005).

Parâmetros do exame	i1	i2	i3	i4
Aspecto de mobilidade	paralisia de prega vocal direita	paralisia de prega vocal esquerda	estenose posterior	paralisia de prega vocal direita
Aspecto estrutural	pregas vocais levemente arqueadas	aparente desnível de pregas vocais	fibrose de pregas vocais	
Videostroboscopia				
- fechamento glótico	Fechamento incompleto (fenda triangular posterior)	fechamento completo	fechamento incompleto	fechamento incompleto
- periodicidade	ausente	presente	ausente	presente
- simetria de fase e amplitude	assimetria de fase discreta	assimetria de fase discreta	Assimetria de fase, inclusive na mesma prega vocal (esquerda)	assimetria de amplitude (maior à esquerda)
- onda mucosa	diminuída em amplitude do lado paralisado	presente	diminuída em amplitude bilateral	praticamente ausente
- indícios de atividade supraglótica	mobilização de prega vestibular esquerda em grau leve	assimetria de movimentação de aritenóides, constrição supraglótica (ântero-posterior e medial) em grau moderado	constrição supraglótica (ântero-posterior) em grau leve	constrição supraglótica (medial), com padrão variável de movimentação de pregas vestibulares e vibração esporádica
Videoquimografia		não foi possível a realização		
- periodicidade	presente		ausente	presente
- simetria	assimetria de fase discreta		assimetria de fase	assimetria de fase
- onda mucosa	diminuída do lado paralisado		diminuída	diminuída
- fases do ciclo vibratório	abertura lenta e fechamento atrasado do lado paralisado		fase fechada de curta duração	amplitude maior à esquerda

Figura 13: Resultados de exame otorrinolaringológico do grupo estudado.

Na tentativa de desvendar a complexidade da relação da produção vocal e da alteração de estrutura do aparelho fonador, o indivíduo sem queixas de disfonia, nomeado como referência, foi 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 ado-

ção 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 focar a produção vocal. Por várias ocasiões, os dados do indivíduo referência partilharam algumas semelhanças com aqueles de i1, i2, i3 e i4.

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. Tal concepção pode ser ilustrada pelos dados da análise acústica de longo termo, em que 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. 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.

O fator subjacente de forte influência (88,9%) na composição dos agrupamentos da análise acústica de longo termo referiu-se à amplitude registrada nas faixas de frequências 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.

Ao se estabelecer correspondências com os elementos acústicos de curto termo, gradativamente, de i1 a i4, detectou-se 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 rouquidão (Kitajima, 1981; Yumoto et al, 1982; Hammarberg, Gaffin, 1995; Holmberg et al, 1995).

Dessa maneira, i3 e i4 parecem ter se diferenciado na análise de composição de *clusters* perante o grupo, e inclusive entre si, pela falta de definição de estrutura harmônica. As emissões da primeira falante revelaram mais harmônicos, porém mesclados às irregularidades vibratórias (bifurcações), e a segunda praticamente não registrou harmônicos.

Diante de tais observações das características acústicas, torna-se importante focar o julgamento perceptivo-auditivo atribuído às amostras

de fala. O perfil da qualidade vocal dos falantes i2, i3 e i4 revelou, juntamente a vários graus de mobilizações laríngicas, ajustes supralaríngicos que ajudaram inclusive a diferenciá-los entre si. Outra informação importante refere-se à possibilidade de levantamento de correlações entre possibilidades de combinação oposição de ocorrência de ajustes, as quais remetem aos componentes de compatibilidade entre os ajustes no modelo de descrição fonética da qualidade vocal (Laver, 1980, Laver et al, 1981; Laver, 2000; Mackenzie-Beck, 2005), os quais podem colaborar na discussão a respeito da multidimensionalidade da qualidade vocal, ou mais precisamente a dificuldade em estabelecer aspectos da percepção com base numa única propriedade do sinal vocal.

Das tendências de correlações destacadas, cabe observar que para aqueles que expressaram alguma tendência a combinação de ações, os ajustes de dorso de língua recuado e mandíbula fechada apontaram que aspectos de hiperfunção das estruturas supralaríngicas facilitaria a presença de ambas mobilizações de longo termo. Para os ajustes de laringe baixa e *vocal fry*/crepitação pode-se destacar que a posição vertical da laringe tende a interferir em alguns ajustes glóticos (fonatórios) (Pinho, 2003). Como último grupo da tendência de combinação, ajustes de falsete e de constrição faríngea, com ocorrências em curto termo de diplofonia, revelaram tendências de mobilizações que acentuam a percepção de irregularidades na emissão e de agudização do *pitch* vocal, tanto por efeitos de fonte glótica (estiramento de pregas vocais no ajuste de falsete), como de filtro supraglótico (constrição faríngea).

Na direção oposta, da tendência de oposição de ajustes, em que novamente se destaca o princípio da combinabilidade no modelo teórico, os ajustes de voz áspera e modal revelam que o fator aspereza (irregularidade vibratória da mucosa provavelmente devida a rigidez de mucosa) não se apresenta diante da vibração regular, periódica de pregas vocais, indicada pela ocorrência de ajuste de voz modal sem modificação pelo fator aspereza). Tal ocorrência ilustra a dimensão de perturbação da atividade vibratória imposta pelo fator aspereza à vibração natural das pregas vocais durante a produção vocal.

Quanto aos julgamentos perceptivo-auditivos de grau geral de alteração e de agradabilidade, a menção aos ajustes de voz áspera e soprosa ou escape de ar e ainda combinados ao *vocal fry* (crepitação) 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 (Holmberg et al, 1995; Vieira et al, 1996; Vieira, 1997; Blaj et al, 2007).

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 EGG e do exame otorrinolaringológico. De forma sintética, a regularidade vibratória manifesta por (iR), no caso da eletroglotografia, evoluiu 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 de atividade de pregas vocais, altamente refinado no tempo, que representa o substrato fisiológico das distribuições de energia anteriormente referidas e abordadas por meio de análise acústica espectral de curto e de longo termo.

Os achados de harmônicos e suas concentrações na emissão por faixas de frequências encontraram 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, denotou aumento dos componentes de rigidez e fechamento incompleto que colaborou para a diminuição de amplitude espectral e prejuízo do padrão de harmônicos. Há dificuldades na detecção de onda mucosa, enquanto i4 representou a situação em que o escape aéreo superou a atividade vibratória, com escasso registro de onda mucosa, levando-a a apresentar diminuição de energia exatamente nas faixas de frequência em que os demais falantes apresentavam, coincidente com a escassa definição de sonoridade em suas emissões.

Com o avanço gradativo da análise integrada dos dados perceptivo-auditivos, acústicos e fisiológicos, procurou-se por elemento que pudesse

signalizar a complexidade de combinação de eventos subjacentes à geração da qualidade vocal, de forma a revelar a relação e a implicação entre esferas de atividade relativas à fonte e ao filtro, além de permitir uma incursão ao campo segmental, de forma a refletir inclusive a influência dos ajustes de qualidade vocal na produção de segmentos da fala. Para o referido propósito, as medidas de *VOT* permitiram relacionar dificuldades em sincronizar gestos orais e glóticos na caracterização dos ajustes da qualidade vocal e no grau de reconhecimento de fala. Quando o sinal EGG não permitiu a identificação do padrão de atividade laríngea, mais precisamente glótica, as medidas de *VOT* puderam qualificar alteração vibratória presente e seu impacto no reconhecimento do sinal de fala, a qual pode interferir na inteligibilidade do sinal de fala.

Os falantes que apresentaram as maiores limitações de atividade glótica revelaram maiores alterações nos valores de *VOT* para consoantes vozeadas, especialmente para aquela de ponto de articulação velar (Figuras 9 e 10). Diferenças foram encontradas entre o grupo estudado e o indivíduo referência. Os falantes cujas vozes foram caracterizadas também por ajustes supralaríngeos na qualidade vocal tenderam a revelar valores fora da estimativa para as medidas de *VOT* referentes às consoantes plosivas desvozeadas bilabiais e dentais/alveolares no PB. 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.

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 (Camargo, 2002; Andrade, 2004; Gregio et al, 2006), talvez por não contemplar somente aspectos da dinâmica laríngea e por ser pouco usual a investigação do impacto das alterações de qualidade vocal nos campos do reconhecimento e da inteligibilidade de fala.

Para o grupo de sons consonantais plosivos vozeados, 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 suprasegmentais,

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.

Os resultados das medidas de *VOT* foram comparados aos resultados do teste de reconhecimento de fala, revelando correspondências entre medidas de *VOT*, identificação de sílabas e aspectos fisiológicos glóticos (configuração da onda *Lx*, medida *Lx* de índice de velocidade e presença de onda mucosa no exame laringológico). Observou-se que, embora em várias situações de análise os dados de *i1*, *i2* e *iR* estivessem próximos, o traçado da onda *Lx*, bem como a medida de índice de velocidade (Figura 12) permitiram diferenciar *iR* em relação a *i1* e *i2*. Tais modalidades de análise permitiram dimensionar o refinamento de atividade na seqüência da fala, de forma a oferecer maiores informações a respeito dos mecanismos compensatórios desenvolvidos por cada um dos falantes estudados.

Outro achado de interesse para a análise da qualidade vocal disfônica foi o de valores de *VOT* zero para algumas consoantes vozeadas produzidas por *i4*, cujos mecanismos compensatórios para atividade glótica não se mostraram eficientes. Cabe ressaltar que os julgamentos perceptivos indicaram que as sílabas foram todas julgadas como vozeadas, apesar de valores de *VOT* zero não serem achados comuns no PB. Os dados apresentados revelaram a relevância de se aprofundar as descrições das relações entre eventos glóticos e supraglóticos. Nesse sentido, as modalidades de análise acústica de *VOT* e de longo termo revelaram-se particularmente importantes.

O exercício teórico a que se propôs este estudo, a partir de amostras de fala de um grupo particular de indivíduos disfônicos, ressalta a necessidade de embasamento teórico a permear os processos de avaliação, reabilitação e de 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.

Diante da dimensão dos achados discutidos, salienta-se a possibilidade de caracterização dos mecanismos compensatórios presentes nos por-

tadores de alterações vocais, como tentativa de resgate do processo de sonorização. Ao se considerar a plasticidade do aparelho fonador, devemos 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, avançar 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.

6. Conclusões

Ao contrário da maioria dos estudos de qualidade vocal, os quais adotam a abordagem tradicional de descrição das qualidades vocais a partir de um padrão ideal de normalidade, o presente estudo introduziu uma abordagem rara em nosso meio ao enfatizar a sincronização de gestos glóticos (fonatórios) e supraglóticos (articulatórios), como resultado do impacto da alteração da fonte sonora vibratória sobre os movimentos articulatórios, desvendados por meio de índices perceptivo-auditivos (descrição da qualidade vocal a partir de modelo fonético e reconhecimento de fala), acústicos (espectro de longo termo e medidas de VOT) e fisiológicos (onda Lx, medidas de coeficiente de contato e de índice de velocidade). Os resultados focam a multiplicidade de manifestações relacionadas à disfonia de forma a inspirar futuros instrumentos de avaliação e de reabilitação no campo clínico, com embasamento teórico das Ciências Fonéticas.

Recebido em novembro de 2007

Aprovado em julho de 2008

E-mails: zcamargo@pucsp.br
madusali@pucsp.br

REFERÊNCIAS BIBLIOGRÁFICAS

- ANDRADE, Luciana GC. 2004. *Estudo da correlação entre qualidade vocal e disfagia pós-acidente vascular cerebral: aspectos acústicos, fisiológicos e perceptivos* [dissertação]. Pontifícia Universidade Católica de São Paulo.
- ASKENFELT ANDERS G & Hammarberg BRITTA. 1986. Speech waveform perturbation analysis: a perceptual-acoustical comparison of seven measures. *J Speech Hear Res* 29: 50-64.
- AWAN, Shaheen N & Frenkel MICHAEL L. 1994. Improvements in estimating the harmonics-to-noise ratio of the voice. *J Voice* 8: 255-262.
- BIELAMOWICZ, Steven; KREIMAN, Jody; GERRATT, Bruce R; DAUER, Marc S; BERKE, Gerald S. 1996. Comparison of voice analysis systems for perturbation measurement. *J Speech Hear Res*; 39:126-134.
- BLAJ, Sabrina C; MADUREIRA, Sandra; CAMARGO, Zuleica. 2007. Vocal quality in asmathic individuals with and without paradoxical vocal fold dysfunction: perceptual, acoustic and physiologic correlates. In: 2nd International Composium of IALP (International Association of Logopedics and Phoniatrics, São Paulo. *Abstract Book 2nd International Composium of IALP*. São Paulo: IALP, 2007. v. I. : 126-127.
- BONFIM, Maria F; CAMARGO, Zuleica; FERREIRA, Leslie P; MADUREIRA, Sandra. Qualidade vocal e formantes das vogais de falantes adultos da cidade de João Pessoa. *Revista CEFAC* 2007; 9(1): 99-109.
- CAMARGO, Zuleica A. 1996. *Parâmetros vocais e configurações laríngeas na fonação de indivíduos submetidos às laringectomias parciais verticais* [dissertação]. São Paulo: Pontifícia Universidade Católica de São Paulo.
- _____. 2002. *Análise da qualidade vocal de um grupo de indivíduos disfônicos: uma abordagem interpretativa e integrada de dados de natureza acústica, perceptiva e eletroglotográfica* [tese]. São Paulo: Pontifícia Universidade Católica de São Paulo.
- CAMARGO, Zuleica; VILARIM, Geisa S; CUKIER, Sabrina. 2004. Parâmetros perceptivo-auditivos e acústicos de longo termo da qualidade vocal de indivíduos disfônicos. *Revista CEFAC*, 6(2):189-196.
- COX, Neil B; ITO, Mabo R; MORRISON, Murray D. 1989. Data labeling and sampling effects in harmonics-to-noise ratios. *J Acoust Soc Am* 85:2165-2176.
- DEJONCKERE, Philippe H; REMACLE, Marc; FRESNEL-ELBAZ, Elizabeth; WOISNARD, Virginie; CREVIER-BUCHMAN, Lise; MILLET, Benoit. 1996. Differentiated perceptual evaluation of pathological voice quality:

- reliability and correlations with acoustic measurements. *Rev Laryngol Otol Rhinol (Bourd)* 117: 219-224.
- DELIYSKI, Dimitar D. 1993. Acoustic model and evaluation of pathological voice production. In: 3rd Conference on Speech Communication and Technology; 1993; Berlin. *Proceedings. Berlin: Eurospeech*: 183-6.
- ESKENAZI, Laurent; CHILDERS, Donald G; HICKS, Douglas M. 1990. Acoustic correlates of vocal quality. *J Speech Hear Res* 33: 298-306.
- FANT, Gunnar. 1970. *Acoustic theory of speech production*. 2nd ed. Paris: Mouton.
- FEIJOO, Sergio; HERNÁNDEZ, Carmen. 1990. Short-term stability measures for the evaluation of vocal quality. *J Speech Hear Res*; 33: 324-334.
- GREGIO, Fabiana N; GAMA-ROSSI, Aglael J; MADUREIRA, Sandra; CAMARGO, Zuleica. 2006. Modelos teóricos de produção e percepção da fala como um sistema dinâmico. *Revista CEFAC* 8(2): 244-247.
- HAMMARBERG, Britta. & GAUFFIN, Jan. 1995. Perceptual and acoustics characteristics of quality differences in pathological voices as related to physiological aspects. In: Osamu FUJIMURA & Minoru HIRANO. *Vocal fold physiology*. San Diego: Singular Publishing Group Inc. 283-303.
- HAMMARBERG, Britta. 2000. Voice research and clinical needs. *Folia Phoniatrica et Logopaedica* 52: 93-102.
- HILLENBRAND, James M. 1987. A methodological study of perturbation and additive noise in synthetically generated voice signals. *J Speech Hear Res* 30: 448-461.
- HIRAOKA, Nobuaki; KITAZOE, Yasuhiro; UETA, Hisashi; TANAKA, Shinzo; TANABE, Masahiro. 1984. Harmonic-intensity analysis of normal and hoarse voices. *J Acoust Soc Am* 76: 1648-1651.
- HOLLIEN, Harry. 2000. The concept of ideal voice quality. In: Ray D KENT & Martin J BALL. *Voice quality measurement*. San Diego: Singular Publishing Group Inc. 13-24.
- HOLMBERG, Eva B; HILLMAN, Robert E; PERKELL, Joseph S; GUIOD, Peter C; GOLDMAN, Susan L. 1995. Comparisons among aerodynamic, electroglottographic and acoustic spectral measures of female voice. *J Speech Hear Res* 38:1212-1223.
- KASUYA, Hideki; OGAWA, Shigeki; MASIMA, Kasuhiko; EBIHARA, Satoshi. 1986. Normalized noise energy as an acoustic measure to evaluate pathologic voice. *J Acoust Soc Am* 80: 1329-1344.
- KENT, Ray D; VORPERIAN, Houri K; DUFFY, Joseph R. 1999. Reliability of the multi-dimensional voice program for the analysis of voice samples of subjects with dysarthria. *Am Journ Speech-Lang Pathol* 8: 129-136.

- KITAJIMA, Kazutomo. Quantitative evaluation of the noise in the pathologic voice. 1981. *Folia Phoniat* 33: 115-124.
- KLINGHOLZ, Fritz; MARTIN, Frank. 1985. Quantitative spectral evaluation of shimmer and jitter. *J Speech Hear Res* 28: 169-174.
- KOJIMA, Hiromi; GOULD, Wilbur J; LAMBIASE, Anthony; ISSHIKI, Nobuhiko. 1980. Computer analysis of hoarseness. *Acta Otolaryngol* 89: 547-554.
- LAVER, John. 1980. *The phonetic description of voice quality*. Cambridge: Cambridge University Press.
- LAVER, John; WIRZ, Sheila L.; MACKENZIE-BECK, Janet & HILLER, Steve M. 1981. A perceptual protocol for the analysis of vocal profiles. *Edinburgh University Department of Linguistics Work in Progress* 14: 139-155.
- LAVER, John. 2000. The phonetic evaluation of voice quality. In: Ray D KENT & Martin J BALL. *Voice quality measurement*. San Diego: Singular Publishing Group Inc. 37-48.
- MACKENZIE-BECK, Janet. 1988. *Organic variation and voice quality*. [PhD thesis]. Edinburgh: University of Edinburgh.
- MACKENZIE-BECK, Janet. 2005. Perceptual analysis of voice quality: the place of vocal profile analysis. In: William J HARDCASTLE & Janet MACKENZIE-BECK. *A figure of speech: a festschrift for John Laver*. Lawrence Erlbaum Associates, Mahwah, p. 285-322.
- MCALLISTER, Anita; SEDERHOLM, Elisabeth; TERNSTRÖM, Sten O; SUNDBERG, Johan. 1996. Perturbation and hoarseness: a pilot study of six children's voices. *J Voice* 10: 252-261.
- QI, Yingyong. 1992. Time normalization in voice analysis. *J Acoust Soc Am* 92: 2569-2576.
- RABINOVIC, C Rose; KREIMAN, Jody; GERRATT, Bruce R; BIELAMOWICZ, Steven. 1995. Comparing reliability of perceptual ratings and acoustic measures of jitter. *J Speech Hear Res* 38: 26-32.
- PERALTA, Juliana S. 2005. *A investigação da qualidade vocal de crianças deficientes auditivas: correlatos acústicos de longo termo e perceptivo-auditivos*. [Trabalho de conclusão de curso]. São Paulo: Pontifícia Universidade Católica de São Paulo.
- PINHO, Silvia M. 2003. Avaliação e tratamento da voz. In: Silvia M PINHO (org). *Fundamentos em fonoaudiologia*. Guanabara Koogan, Rio de Janeiro. p 3-40.
- SASAKI, Yumi; OKAMURA, Hiroshi; YUMOTO, Eiji. 1991. Quantitative analysis of hoarseness using a digital sound spectrograph. *J Voice* 5: 36-40.

- SCHERER, Ronald C; VAIL, Vernon J; GUO, Chewn G. 1995. Required number of tokens to determine representative voice perturbation values. *J Speech Hear Res* 38:1260-1269.
- SMITS, Ilse; CEUPPENS, Pieter; DE BODT, Marc S. 2005. A comparative study of acoustic voice measurements by means of Dr. Speech and Computerized Speech Lab. *J Voice* 19(2):187-196.
- TONER, Mary A; EMANUEL, Floyd W; PARKER, Donald. 1990. Relationship of spectral noise levels to psychophysical scaling of vowel roughness. *J Speech Hear Res* 33: 238-244.
- ULOZA, Virgilijus; SAFERIS, Viktoras; ULOZIENE, Ingrida. 2005. Perceptual and acoustic assessment of voice pathology and the efficacy of endolaryngeal phonomicrosurgery. *J Voice* 19(1): 138-145.
- VIEIRA, Maurílio N; MCINNES, Fergus G. & JACK, Mervyn A. 1996. Robust f0 and jitter estimation in pathological voices. In: *4th International Conference on Spoken Language Processing; Proceedings*. Philadelphia: ICSLP 745-748.
- VIEIRA, Maurílio N. 1997. *Automated measures of dysphonias and the phonatory effects of asymmetries in the posterior larynx*. [PhD thesis]. Edinburgh: University of Edinburgh.
- WUYTS, Floris L; DE BODT, Marc S; BRUCKERS, Liesbeth; MOLENBERGHS, Geert. 1996. Recent work of the Belgian study group of voice disorders. Results. *Acta Otorhinolaryngol Belg* 50: 331-341.
- YU, Ping; OUAKNINE, Maurice; REVIS, Joana; GIOVANNI, Antoine. 2001. Objective voice analysis for dysphonic patients: a multiparametric protocol including acoustic and aerodynamic measurements. *J Voice* 15(4): 529-542.
- YUMOTO, Eiji; GOULD, Wilbur J; BAUER, Thomas. 1982. Harmonics-to-noise ratio as an index of the degree of hoarseness. *J Acoust Soc Am* 71: 1544-1550.
- ZHANG, Yu; JIANG, Jack J.; BIAZZO, Laura; JORGENSEN, Malinda. 2005. Perturbation and nonlinear dynamic analyses of voices from patients with unilateral laryngeal paralysis. *J Voice* 19(4): 519-28.