

**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*

D.E.L.T.A., 25:2, 2009 (285-317)

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; Deliyiski, 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, 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. 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 disфония 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, videolaringoscopia 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 freqüência 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 [αPαP], 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 suaprararíngeos, 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 seqüência, 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  $f_0$ , 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

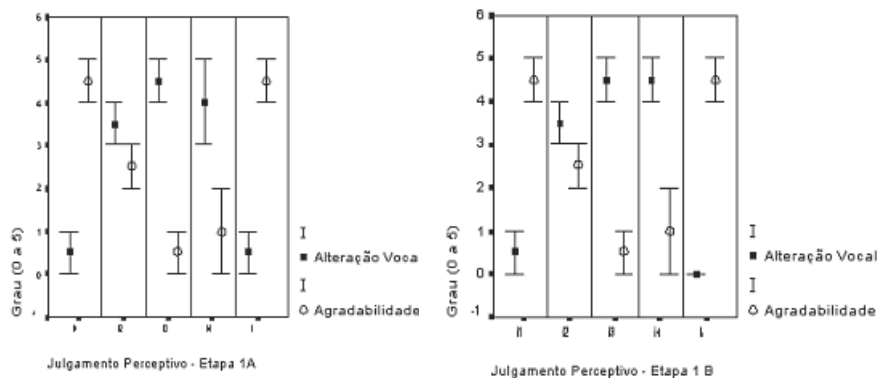


Paulo. O equipamento para avaliação incluiu endoscópio rígido 70° *LY-C530 Machida*, estroboscópio *Brüel & Kjaer type 4914* e videoquimógrafo 8900-*Kay Elemectrics Corp.* No caso de exacerbado reflexo, foi utilizado endoscópio flexível ENT 30 P3 *Machida* em associação com o referido videoestroboscópio. 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 de 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		• □		• □	
Áspera		• □	• □	• □	
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	<b>LARÍNGEOS</b>							
	Laringe alta				○			
	Laringe baixa			◇				
	<b>LABIAIS</b>							
	Labiodontalização			■				
	Protrusão							
SUPRALARÍNGEOS TRANSVERSAIS	<b>LABIAIS</b>							
	Arredondados							
	Estirados							
	<b>MANDIBULARES</b>							
	Fechada					▲		
	Aberta							
	<b>LINGUAIS - PONTA/ LÂMINA</b>							
	Avançada							
	Recuada							
	<b>LINGUAIS - CORPO</b>							
	Avançado							
	Retraído				▲			
	Elevado							
	Abaixado							
	<b>LINGUAIS - BASE</b>							
	Constricção faríngea						■	
Expansão faríngea								
VELOFARÍNGEOS	<b>ACOPLAGEM VELAR</b>							
	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.

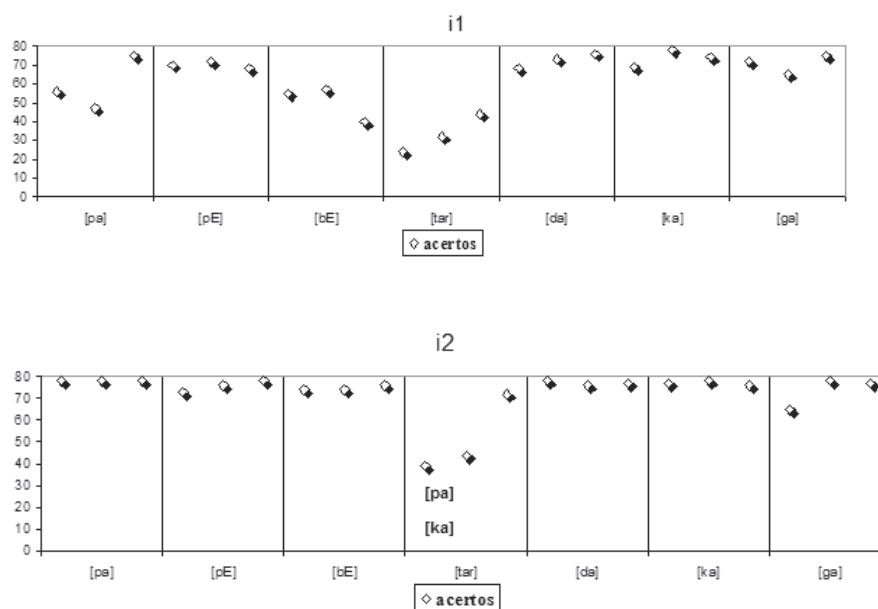
No segundo grupo, o das oposições, destacaram-se ações que apresentam maior dificuldade, ou mesmo impossibilidade de combinação:

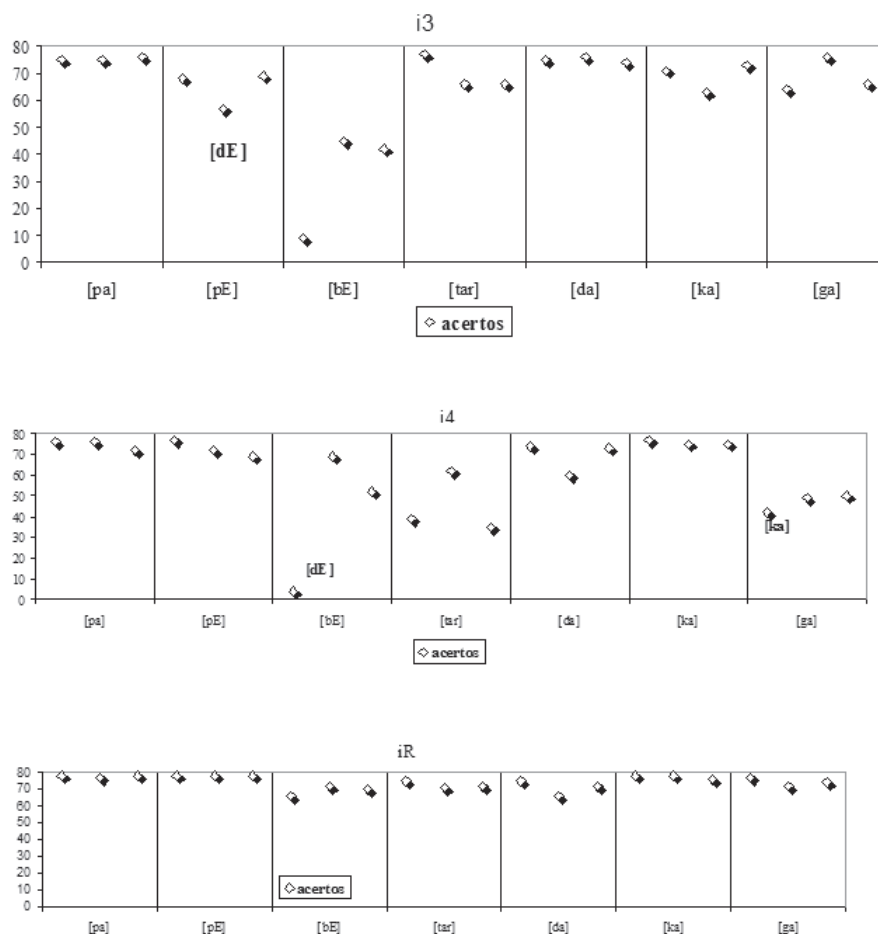
- ajuste de voz áspera ausente e ajuste modal presente *versus* ajuste de voz áspera presente e ajuste modal ausente.

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] [p"], [b"], [taɾ], [da], [ka] e [ga] dos indivíduos do grupo estudado (i1 a i4) e indivíduo referência (iR).

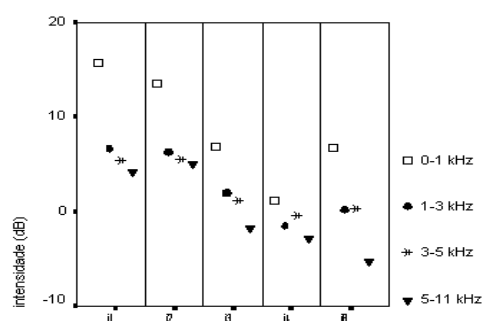
Diferenças estatisticamente significantes ( $p < 0,05$ ) foram identificadas para as sílabas [pa] ( $p = 0,046$ ), [p"] ( $p = 0,002$ ), [b"] ( $p = 0,30$ ) e [taɾ] ( $p = 0,03$ ) de i1, [pa] ( $p = 0,08$ ), [p"] ( $p = 0,037$ ), [b"] ( $p = 0,040$ ), [b"] ( $p = 0,028$ ) de i3 e [p"] ( $p = 0,042$ ), [taɾ] ( $p = 0,040$ ) 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).

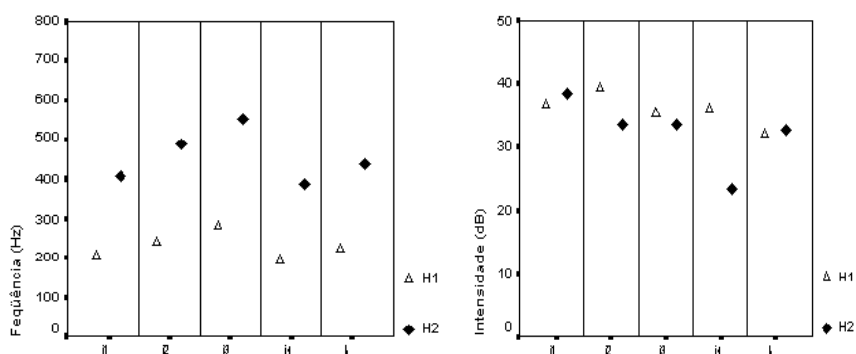
A análise estatística aplicada ao cômputo das médias anteriormente representadas revelou o agrupamento das três emissões distintas de i3, i4 e iR, enquanto i1 e i2 mesclaram suas emissões por dois grupos. O fator

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).



