

8.00.00.00-2 LINGÜÍSTICA, LETRAS E ARTES

| 8.01.00.00 – 7 - LINGÜÍSTICA

CORRELATOS ACÚSTICOS DOS AJUSTES SUPRAGLÓTICOS DE
| QUALIDADE VOCAL

PERPÉTUA COUTINHO GOMES

Curso de Fonoaudiologia- Faculdade de Ciências Humanas e da Saúde
| (FACHS)

ZULEICA CAMARGO

Departamento de Lingüística - Faculdade de Filosofia, Comunicação, Letras e
Artes (FAFICLA)

RESUMO: HÁ POUCAS DESCRIÇÕES REFERENTES AOS CORRELATOS ACÚSTICOS DOS AJUSTES SUPRAGLÓTICOS DE QUALIDADE VOCAL. O OBJETIVO DESTA PESQUISA FOI INVESTIGAR DADOS ACÚSTICOS DE FREQUÊNCIA E INTENSIDADE DE FORMANTES DA VOGAL ORAL [a] EM POSIÇÃO TÔNICA E PÓS-TÔNICA DE 40 AMOSTRAS DE FALA REPRESENTATIVAS DE AJUSTES SUPRAGLÓTICOS. TAIS AMOSTRAS FORAM ANALISADAS POR MEIO DO *SOFTWARE PRAAT*, A PARTIR DA EXTRAÇÃO DOS VALORES DE FREQUÊNCIA (F1, F2 E F3) E INTENSIDADE (I1, I2 E I3) DOS TRÊS PRIMEIROS FORMANTES. OS ACHADOS FORAM SUBMETIDOS A ANÁLISE ESTATÍSTICA DE NATUREZA MULTIVARIADA. A ANÁLISE DE *CLUSTER* REVELOU AGRUPAMENTO DIFERENCIADO DOS AJUSTES SUPRAGLÓTICOS PARA FALANTES DOS GRUPOS MASCULINO E FEMININO. A ANÁLISE DISCRIMINANTE REVELOU A POSSIBILIDADE DE PREDIÇÃO DO GÊNERO DO FALANTE E DA POSIÇÃO ACENTUAL DA VOGAL POR MEIO DAS MEDIDAS DE FREQUÊNCIAS FORMÂNTICAS (F1 E F2) EM AMBOS OS GÊNEROS. NO CASO DAS MULHERES, HOUE A SEGREGAÇÃO DO AJUSTE DE HIPERFUNÇÃO LARÍNGEA E, NOS HOMENS, DE CORPO DE LÍNGUA AVANÇADA. OS VALORES PARA AJUSTES NEUTROS PODEM REPRESENTAR UMA ESTIMATIVA DE REFERÊNCIA PARA A VOGAL [a]: HOMENS F1 (Hz): 647, F2 (Hz): 1358 E F3 (Hz): 2657 HZ E MULHERES F1 (Hz): 791, F2 (Hz) :1667 E F3 (Hz): 2855 HZ; HOMENS I1 (dB): 56 , I2(dB):33 E I3 (dB): 17 dB E MULHERES: I1(dB): 47 I2 (dB): 36 E I3(dB): 21 dB.

PALAVRAS-CHAVE: Percepção da Fala, Acústica da Fala, Voz.

Introdução

No universo das Ciências da Fala, a interdisciplinaridade possibilita a integração e a colaboração de profissionais de diversas áreas (medicina, engenharia, fonoaudiologia, linguística, entre outras) na investigação inesgotável das particularidades relacionadas à faculdade da fala, bem como na tentativa de desenvolvimento de recursos tecnológicos (Tecnologias de Fala). Dentre suas múltiplas vantagens, tal integração permite o avanço de tecnologias para telecomunicações, de desenvolvimento de recursos mais avançados para a formação de profissionais da fala e auxiliam na reabilitação de um grande número de distúrbios da comunicação.

Nesse contexto, a qualidade vocal tem sido alvo de muitas investigações, tanto em termos de detalhamento de seus mecanismos de produção como de percepção. Os estudos na área abordam desde as propriedades estéticas da sonoridade, no campo da expressividade, até aquelas impostas por distúrbios nos sistemas de produção e de percepção de sons, no campo das disfonias. Entretanto, as descrições de correlatos acústicos de ajustes supraglóticos de qualidade vocal, os quais respondem pela maior parte das informações relacionadas ao padrão de formantes, são ainda escassas.

Diante disso, o objetivo desta pesquisa foi investigar dados acústicos de frequência (F) e intensidade (I) de formantes da vogal oral [a] em posição tônica e pós-tônica de amostras de fala representativas de ajustes supraglóticos de qualidade vocal.

1. Método

O *corpus* da pesquisa foi constituído por 40 estímulos (15 estímulos de falantes do sexo masculino e 25 estímulos de falantes do sexo feminino com idades entre 17 e 58 anos) em formato de áudio, constantes no banco de dados do LIAAC e referentes ao roteiro *Vocal Profile Analysis Scheme* para o Português Brasileiro (VPAS-PB): ajustes de qualidade vocal do trato vocal supralaríngeo.

As gravações são constituídas por um trecho de fala semi-espontânea, no qual cada sujeito foi solicitado a falar sobre a cidade em que nasceu e a realizar a leitura de 10 sentenças-veículos elaboradas para análise de qualidade vocal e, portanto, levando em consideração o princípio de

susceptibilidade do segmento aos efeitos dos ajustes de qualidade vocal (CAMARGO, MADUREIRA, 2008).

Após estudos e análises, foi decidido que das 10 sentenças acima referidas, utilizaríamos apenas as três destacadas abaixo, uma vez que são apenas estas que revelam a ocorrência da vogal [a] em posição tônica (grifada) e pós-tônica (em negrito):

- O objeto de estudo da Fonética é essa complexa, variável e poderosa face sonora da linguagem: a **fala**;
- A **Lara** guarda figuras de pássaros em uma caixa e suas preferidas são a da **ara**, da patativa, da garça, do canário e do sabiá amarelo;
- Detesto ir à **casa** dele, pois fica do outro lado da cidade e o acesso é difícil.

Das três sentenças acima, selecionamos apenas os enunciados em que constava o segmento-chave [a] em posição tônica e pós-tônica para extrair as medidas, isto é, os enunciados referentes aos segmentos-chave **ara**, **casa**, **fala** e **Lara**.

As sentenças foram segmentadas com o auxílio do *software* de livre acesso *Praat* (<http://www.fon.hum.uva.nl/praat/>) por meio de extração manual dos valores de frequência (F1, F2 e F3) e intensidade (I1, I2 e I3) dos três primeiros formantes, a partir dos algoritmos FFT (*Fast Fourier Transform*) e LPC (*Linear Predictive Code*). As etapas para realizar tal procedimento encontram-se minuciosamente descritas abaixo e foram conduzidas para cada vogal analisada:

- *Etapa 1*: Geração do espectrograma de banda estreita por meio dos parâmetros *spectrum*, *spectrum settings*, *Window length: 0.05*.
- *Etapa 2* : Seleção do ponto mais estável no período estacionário da vogal e geração do espectro FFT por meio dos parâmetros *spectrum*, *view spectrum slice*.
- *Etapa 3*: Para a obtenção dos valores de frequência (em Hz) e intensidade (em dB) dos três primeiros formantes, geração do espectro LPC a partir do arquivo *spectrum* constante na janela *Praat objects*. Para isso, foi utilizado o parâmetro *LPC Smoothing/ Number of peaks*, no qual considerando a frequência de amostragem (o limite de frequência possível de análise), foi inserido o número de formantes. Dessa forma, para frequência de

amostragem de 22050 Hz, por exemplo, o limite de análise é de 11025 Hz e o número de formantes será de aproximadamente 11-12 (esse número pode variar se a amostra for de uma voz mais grave ou aguda, portanto homem, mulher ou criança). Por meio do pico da onda foram extraídos manualmente os valores de frequência e de intensidade.

- *Etapa 4:* Traçado do espectro FFT por meio dos parâmetros *draw frequency*, *spectrum range* 0 a 5000Hz e, em seguida, traçado e sobreposição do espectro de LPC utilizando os mesmos parâmetros..
- *Etapa 5:* Retorno ao espectrograma e, por meio dos parâmetros *spectrum*, *spectrum settings*, *Window length* : 0.005, modificação para um espectrograma de banda larga:
- *Etapa 6:* Adição à opção de marcação de formantes utilizando os parâmetros *Formant*, *Show formant*.
- *Etapa 7:* Extração das medidas de frequência dos formantes por meio dos parâmetros *Formant*, *Formant listing*. As medidas de frequência dos formantes aparecem na janela *Praat Info*. Tais medidas foram comparadas com os valores extraídos manualmente.

Após finalizar a extração, os dados acústicos foram submetidos à análise estatística de natureza multivariada: Análise Aglomerativa Hierárquica de *Cluster*, Análise Discriminante e Análise de Correlação Canônica. Todas as etapas de testes estatísticos foram realizadas com apoio do *software XLSTAT* da *Addinsoft*.

1.1 Resultados

A análise aglomerativa hierárquica de *cluster* revelou que as medidas de frequência e intensidade de formantes agruparam-se de forma mais uniforme no grupo feminino, em que as medidas de frequência (F1, F2 e F3) estiveram mais próximas, assim como aquelas de intensidade (I1, I2 e I3). No caso dos homens, somente F1 e F2 agruparam-se entre si, enquanto F3 agrupou-se às medidas de intensidade conforme figura 1.

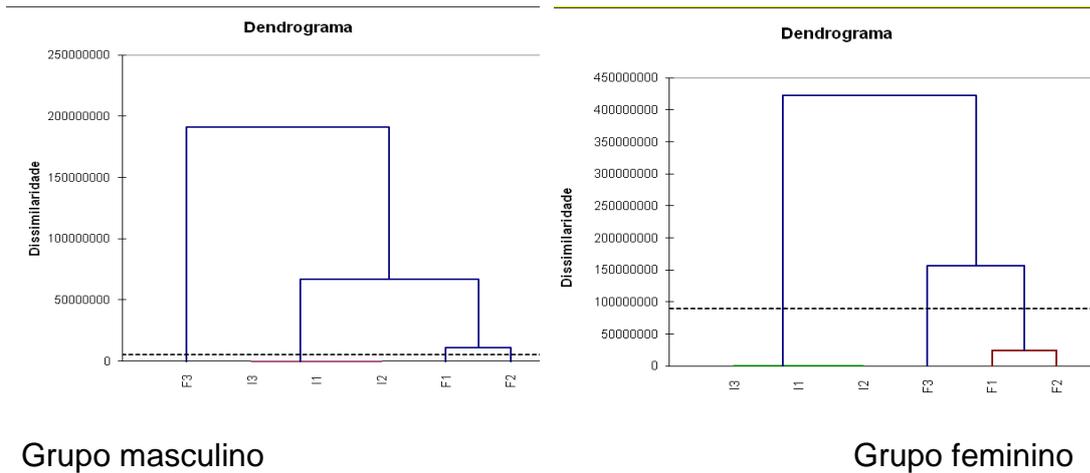


Figura 1. Dendrogramas da análise aglomerativa hierárquica de *cluster* para as medidas de frequência (F1, F2 e F3) e intensidade (I1, I2 e I3) de formantes dos grupos masculino e feminino.

A análise aglomerativa hierárquica de *cluster* revelou agrupamento dos ajustes de qualidade vocal no plano supralaríngeo tanto para falantes do gênero masculino, quanto feminino, conforme figuras 2 e 3.

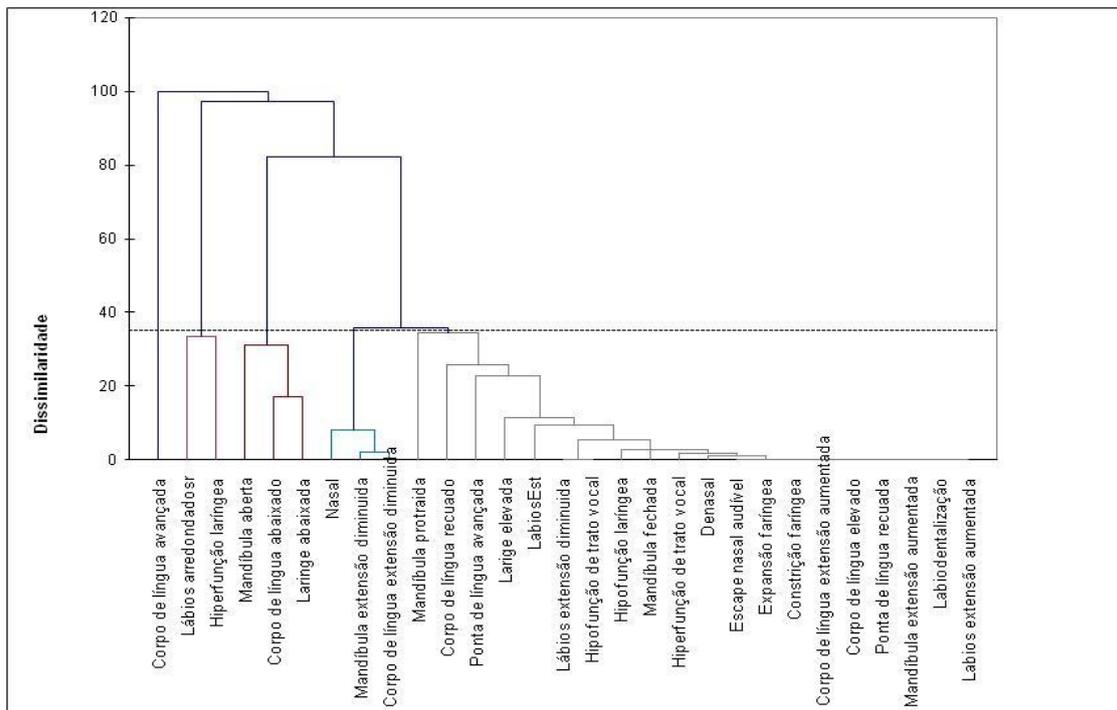


Figura 2: Dendrograma da análise aglomerativa hierárquica de *cluster* para os ajustes supralaríngeos e de tensão de qualidade vocal do grupo feminino.

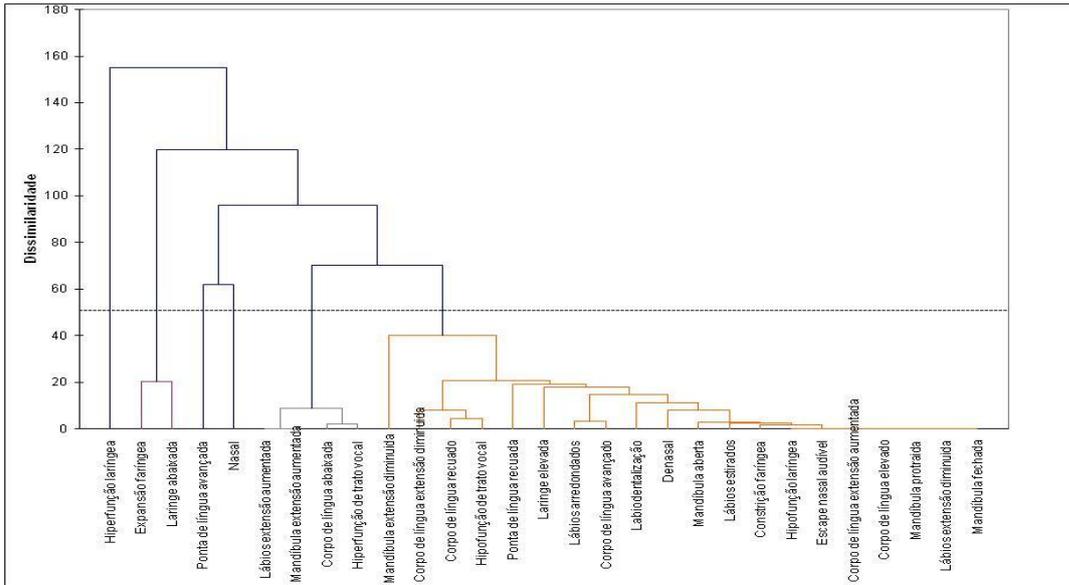


Figura 3: Dendrograma da análise aglomerativa hierárquica de *cluster* para os ajustes supralaríngeos e de tensão de qualidade vocal do grupo feminino.

Os valores mínimo, máximo, média e desvio padrão de frequências e intensidades formânticas encontrados nos falantes com ajustes neutros são apresentados nas tabelas 1 e 2, respectivamente para homens e mulheres.

Variável	Mínimo	Máximo	Média	Desvio Padrão (DP)
F1	455	926	647	92,479
F2	1033	1688	1358	135,203
F3	1957	3119	2657	229,724
I1	9	343	56	29,712
I2	6	58	35	10,379
I3	-4	42	17	9,829

Tabela 1: Valores mínimo, máximo, médio e DP de frequências (F1, F2 e F3) e intensidades (I1, I2 e I3) formântica da vogal [a] em posição tônica dos falantes do grupo masculino (azul) com ajustes neutros da esfera supralaríngea.

Variável	Mínimo	Máximo	Média	Desvio Padrão (DP)
F1	494	1232	785	144,1
F2	993	2052	1652	196,673
F3	2310	3894	2815	243,659
I1	16	70	47	8,181
I2	7	54	36	8,002
I3	-9	43	21	9,773

Tabela 2: Valores mínimo, máximo, médio e DP de frequências (F1, F2 e F3) e intensidades (I1, I2 e I3) formântica da vogal [a] em posição tônica dos falantes do grupo feminino (vermelho) com ajustes neutros da esfera supralaríngea.

A análise de correlação canônica para o conjunto de informações acústicas e perceptivas (supraglóticos) é apresentada nas figuras 4 e 5.

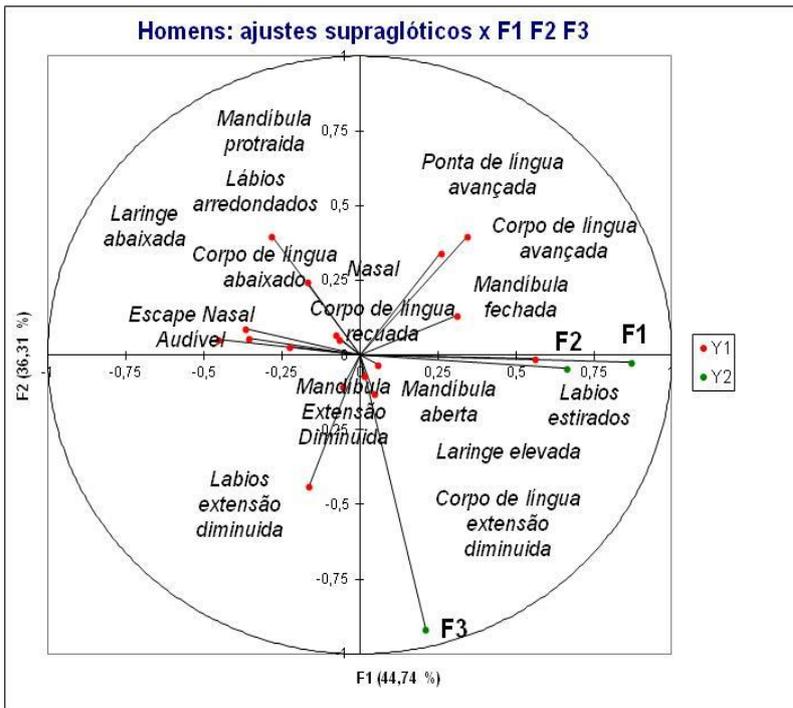


Figura 4: Diagrama circular resultante da análise de correlação canônica de dados perceptivos (supraglóticos) e acústicos do grupo masculino.

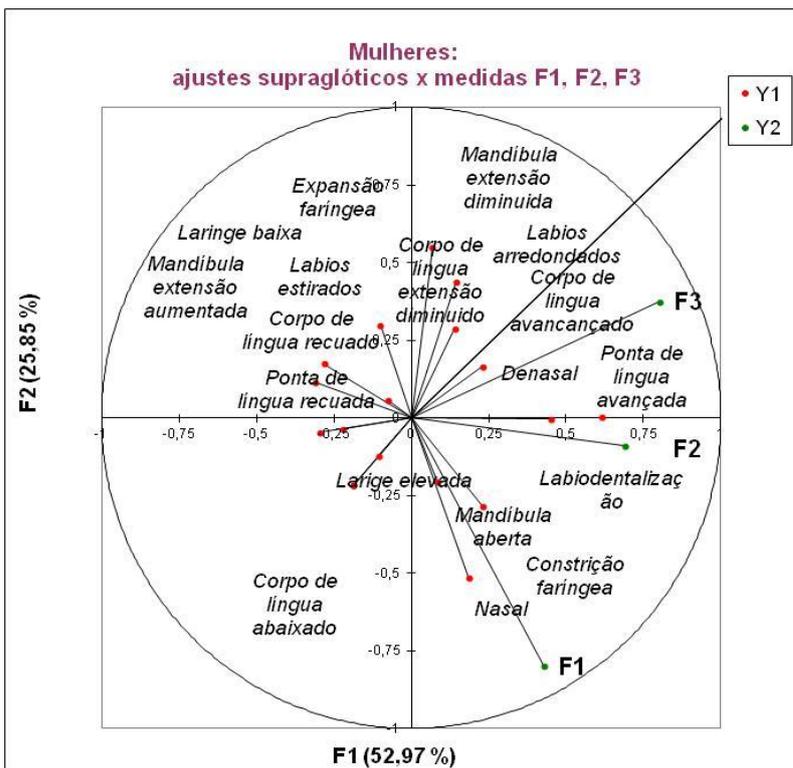


Figura 5: Diagrama circular resultante da análise de correlação canônica de dados perceptivos (supraglóticos) e acústicos do grupo masculino.

A análise de correlação canônica para o conjunto de informações acústicas e perceptivas (ajustes supralaríngicos e de tensão) é apresentada nas figuras 6 e 7.

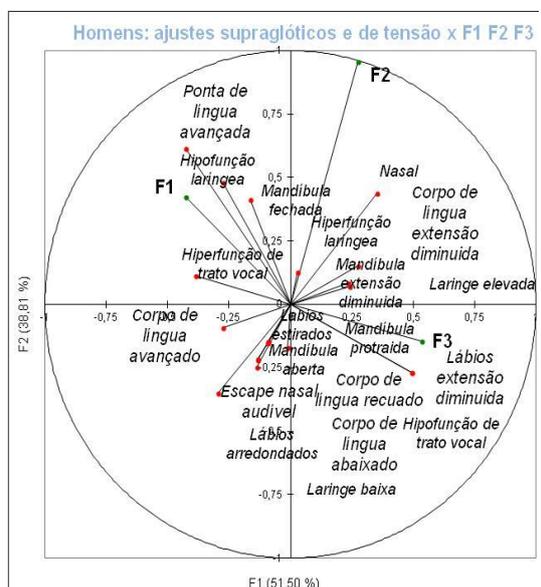


Figura 6: Diagrama circular resultante da análise de correlação canônica de dados perceptivos (supraglóticos e de tensão) e acústicos do grupo masculino.

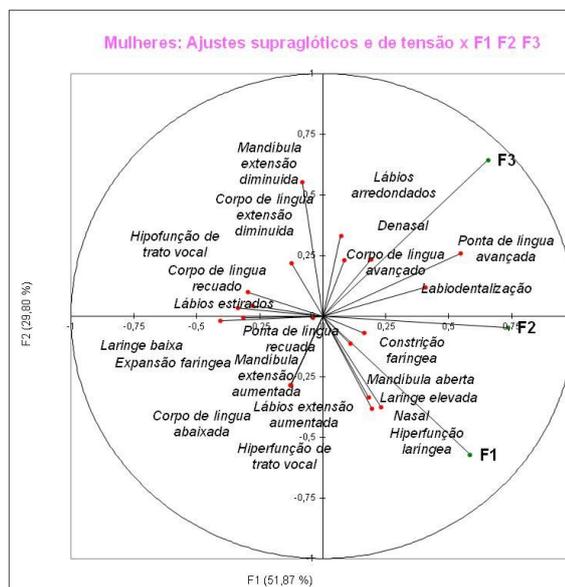


Figura 7: Diagrama circular resultante da análise de correlação canônica de dados perceptivos (supraglóticos e de tensão) e acústicos do grupo feminino.

Na análise de correlação canônica, as correlações entre dados perceptivos e acústicos diferenciaram-se entre os gêneros. Para falantes do gênero feminino, foram detectadas correlações (em ordem decrescente de influência) entre: ajuste de ponta de língua avançada e F3 (27,2%), F1 (21%) e F2 (15,6%), laringe elevada e F1 (17%), labiodentalização e F2 (16,4%), lábios arredondados e F2 (15,2%), nasal e F1 (14,8%) e hiperfunção laríngea e F2 (12,5%). Para falantes do gênero masculino, foram detectadas correlações (em ordem decrescente de influência): entre lábios estirados e F1 (47,9%) (além de F2: 18,3% e F3: 16,1%), hipofunção laríngea e F2 (29,1%) (além de F1:21,9%), hipofunção de trato vocal e F3 (26,2%), corpo de língua extensão diminuída e F2 (24,6%), corpo de língua avançado e F1 (24,4%), ponta de língua avançada e F1 (24,1%), nasal e F2 (20,1%) e mandíbula fechada e F2 (20%) (além de F1:19,6%).

A análise discriminante, por sua vez, revelou a possibilidade de predição do gênero e da posição acentual (tônica e pós-tônica) por meio das medidas de frequências formânticas (F1 e F2), tanto em homens quanto em mulheres. O

poder discriminatório deste grupo de medidas não aumentou quando consideradas as medidas de F3, em conjunto com F1 e F2. No caso da ocorrência de ajustes neutros e não neutros supraglóticos e de tensão, a combinação de medidas de F1 e F2 revelou poder segregatório de 72% para os ajustes neutros e 74,12% para os não neutros no grupo masculino. Para o grupo feminino, os valores foram inferiores, com 54,55% de poder discriminatório para ajuste neutro e 51,75% para os não neutros. Os valores foram similares à combinação F1, F2 e F3. Os valores de lambda e p-valor são apresentados nas tabelas a seguir:

Variável	Lambda	p-valor
F1	0,998	0,824
F2	0,988	0,321

Tabela 3: valores de lambda e p-valor de F1 e F2 referentes à ocorrência de ajustes neutros e não-neutros no grupo feminino

Variável	Lambda	p-valor
F1	0,985	0,203
F2	0,984	0,181

Tabela 4: valores de lambda e p-valor de F1 e F2 referentes à ocorrência de ajustes neutros e não-neutros no grupo masculino.

No caso da posição acentual, as medidas de F1 e F2 revelaram poder discriminatório da posição de acento pós-tônica de 83,64% e 81,82% de posição tônica do grupo masculino. No grupo feminino, as medidas de F1 e F2 alcançaram poder discriminatório da posição de acento pós-tônica de 91,21% e 84,38% na posição tônica. Os valores de lambda e p-valor são apresentados na tabela a seguir:

Variável	Lambda	p-valor
F1	0,508	<0,0001
F2	0,928	0

Tabela 5: valores de lambda e p-valor de F1 e F2 referentes à posição tônica e pós-tônica do grupo feminino.

Variável	Lambda	p-valor
F1	0,626	<0,0001
F2	0,862	<0,0001

Tabela 6: valores de lambda e p-valor de F1 e F2 referentes à posição tônica e pós-tônica do grupo masculino.

Os valores de intensidades formânticas não revelaram poder segregatório em termos de gêneros e de posição acentual.

1.2 Discussão

Quanto à distribuição de ajustes de qualidade vocal, a análise de *cluster* propiciou observar diferenciações em termos dos grupos masculino e feminino,

em que se destacou, no caso das mulheres a diferenciação do ajustes de hiperfunção laríngea e, nos homens, o ajuste de corpo de língua avançado. Tais dados remontam às descrições de particularidades das vozes femininas e masculinas, especialmente de aspectos da configuração laríngea e proporção glótica (BEHLAU, 2001), as quais podem interferir no grau de tensão laríngea e interferir nas respostas formânticas.

Para ambos os gêneros, as medidas de F1 e F2 formaram uma classe distinta na análise de *cluster*. Tais dados da congruência de F1 e F2 para descrição de ajustes de qualidade vocal de natureza supraglótica foi revelada pela análise de correlação canônica, em que as medidas de F1 e F2 agruparam-se a ajustes de ponta e corpo de língua, mandíbula, tensão laríngea e de trato vocal para falantes do gênero masculino. Para o grupo feminino, as distâncias entre F1, F2 e F3 foram similares, sendo que os ajustes de altura e tensão de laringea, faringe, mandíbula e nasal estiveram situados entre F1 e F2 e ajustes de lábios, ponta e corpo de língua situaram-se no espaço entre as medidas de F2 e F3. Tais dados são congruentes a achados de segregação entre grupos masculino e feminino em função de medidas acústicas de f_0 , intensidade e declínio espectral (CAMARGO et al., 2012)

As medidas de F1 e F2 foram as mais relevantes na identificação de ajustes de natureza neutra e não neutra da esfera supralaríngea e de tensão para a vogal [a], de maneira que os valores para ajustes neutros podem representar uma estimativa de valores de referência para homens e mulheres adultos falantes do PB: homens F1 (Hz): 647, F2 (Hz):1358 e F3(Hz): 2657 e mulheres F1:(Hz)791 , F2 (Hz): 1667 e F3 (Hz): 2855 . Com relação às medidas de intensidade foram encontradas: Homens I1 (dB): 56 , I2(dB): 33 e I3 (dB): 17 dB e mulheres; I1(dB): 47 I2(dB): 36 e I3(dB):21.

Provavelmente, a associação a outras medidas previamente utilizadas (CAMARGO, MADUREIRA, 2010; CAMARGO et al., 2012) aumente o poder discriminatório do grupo de medidas acústicas em predizer mobilizações específicas do trato vocal supraglótico. Este pode ser um caminho natural de continuidade desta linha de pesquisas, em que, na busca por caracterizações dos aspectos supraglóticos de qualidade vocal, encontramos esferas de interrelação de mecanismos supralaríngeos, de tensão e, inclusive, fonatórios. Tais achados respaldam o princípio da combinabilidade de ajustes de

qualidade vocal no plano acústico e fisiológico, segundo modelo fonético de descrição da qualidade vocal (LAYER, 1980)

As medidas de F1 e F2 foram as mais relevantes na identificação de ajustes de natureza neutra e não neutra da esfera supralaríngea e de tensão para a vogal [a].

1.3 Conclusão

As medidas de frequências formânticas F1 e F2, permitem diferenciar ajustes supraglóticos de qualidade vocal.

REFERÊNCIAS

BEHLAU, M; AZEVEDO, R; PONTES, P. **Conceito de voz normal e classificação das disfonias**. In: Behlau M (org). Voz: o livro do especialista. Rio de Janeiro: Revinter, 2001. p. 53-84.

CAMARGO, Z. MADUREIRA, S. **Voice quality analysis from a phonetic perspective: Voice Profile Analysis Scheme Profile for Brazilian Portuguese (BP- VPAS)**. In: SPEECH PROSODY, 4, 2008, Campinas. Fourth Conference on Speech Prosody - Abstract Book and CD-Rom Proceedings, 2008. v. 1. p.57–60.

CAMARGO, Z.A.; MADUREIRA, S. **The acoustic analysis of speech samples designed for the Voice Profile Analysis Scheme for Brazilian Portuguese (BP-VPAS): long-term f0 and intensity measures**. In: ISCA, 2010, Athens. Edited by Antonis Botinis, v. 1. p. 33-36.

CAMARGO, Z. ET AL. **Voice quality and gender: some insights on correlation between perceptual and acoustic dimensions** In: 6th International Conference on Speech Prosody, 2012, Shanghai. Abstract Book Speech Prosody 2012. Shangai: Tongji University Press, 2012. 1. P.115 – 118.

LAYER, J. **The phonetic description of voice quality**. Cambridge: Cambridge University Press, 1980. p186.