

8.00.00.00-2 LINGUÍSTICA, LETRAS E ARTES
8.01.00.00-7 LINGUÍSTICA

PROSÓDIA AUDIOVISUAL E QUALIDADE VOCAL: UMA PROPOSTA DE
AVALIAÇÃO FONÉTICA

THAMIRES RODRIGUES DE FREITAS

Curso de Fonoaudiologia – Faculdade de Ciências Humanas e da Saúde

ZULEICA CAMARGO - ORIENTADORA

Programa de Estudos Pós-Graduados em Linguística Aplicada e Estudos da
Linguagem

Departamento de Linguística - Faculdade de Filosofia, Comunicação, Letras e
Artes (Faficla).

RESUMO: ESTUDOS RECENTES APONTAM A INFORMAÇÃO VISUAL DO FALANTE COMO FATOR IMPORTANTE NA PERCEPÇÃO DA FALA. NOSSO OBJETIVO FOI ANALISAR GRAVAÇÕES EM VÍDEO REPRESENTATIVAS DE AJUSTES DE QUALIDADE VOCAL DOS PLANOS SUPRALARÍNGEOS E DE TENSÃO, COM BASE EM MODELO FONÉTICO. O CONJUNTO DE 38 VÍDEOS FOI ANALISADO POR MEIO DE DECUPAGEM. OS AJUSTES SUSCEPTÍVEIS AOS EFEITOS DE PERCEPÇÃO VISUAL FORAM AQUELES DE LÁBIOS, MANDÍBULA, LÍNGUA, FARINGÉ E POSIÇÃO VERTICAL DE LARINGE NO PLANO SUPRALARÍNGEO E HIPERFUNÇÃO DE LARINGE E DE TRATO VOCAL NO PLANO DOS AJUSTES DE TENSÃO. OS ACHADOS REFORÇAM OS EFEITOS DA PERCEPÇÃO VISUAL NOS PLANOS DOS AJUSTES SUPRALARÍNGEOS E DE TENSÃO.

Palavras-Chave: Voz, Percepção visual, Fonética.

Introdução

A linha de estudos em prosódia audiovisual é recente para o campo da qualidade vocal, porém estudos atuais apontam a informação visual do falante como fator importante na percepção da voz e fala (ABELIN, 2008). No entanto, os argumentos a favor da importância e contribuição das informações visuais na percepção de fala por ouvintes remontam à década de 80. Tais estudos intensificaram-se na década de 90, com contribuições marcantes de pesquisadores das áreas de Psicologia, Linguística, Engenharia e Ciências da Fala. Tal panorama reforça a indicação de que estes estudos voltaram-se à leitura labial, ou mais precisamente à leitura orofacial, em ouvintes e não em portadores de deficiência auditiva, como na literatura pertinente à reabilitação auditiva (WHALEN, 2002).

Estudos atuais também ressaltam a importância de julgamentos visuais da qualidade vocal, inclusive quanto à exploração do efeito de McGurk para voz (ABELIN, 2008). Tal experimento, explica a tentativa de nosso cérebro em integrar informações conflitantes, de naturezas auditiva e visual, quando apresentadas simultaneamente e, dessa forma, reforça a importância das informações visuais na percepção dos sons. O efeito é melhor exemplificado quando se apresenta um vídeo de uma pessoa pronunciando a sílaba [ba] enquanto estímulo sonoro, porém com a imagem da articulação da sílaba [ga]. Quando ouve o vídeo, sem a pista visual, a maior parte das pessoas escuta a sílaba indicada pela informação auditiva, isto é, o [BA]. No entanto, ao observar o vídeo, o som passa a soar como [da]. O trabalho, assim, demonstra a influência da percepção visual no universo da percepção sonora e revela a integração no sentido de resolução de informações perceptuais conflitantes (audição e visão), em que também consta o desempenho do cérebro. O experimento também pode ser experienciado com outras sílabas, além das já citadas. O efeito foi especialmente descrito para o campo segmental, entretanto há alguns indícios de que se estenda pelos elementos prosódicos (ABELIN, 2008).

No campo da prosódia, a abordagem fonética da qualidade vocal propõe uma análise integrada do sinal sonoro que emana da boca do falante, fruto do desempenho nos níveis glótico, supraglótico e de tensão do aparelho fonador (LAVER, 1980). De acordo com o mesmo autor, a unidade de análise proposta

para avaliação perceptiva da qualidade vocal refere-se ao ajuste muscular de longo termo do aparelho fonador intrínseco, recorrente no ato de fala, denominado *setting* (termo traduzido como ajuste).

Do ponto de vista fonético, os segmentos (vogais e consoantes) são tidos como elementos mais ou menos susceptíveis aos efeitos auditivos dos ajustes da qualidade vocal. Neste sentido é destacado como um dos princípios do modelo fonético de descrição da qualidade vocal a susceptibilidade dos segmentos aos efeitos dos ajustes de qualidade vocal. De acordo com tal conceito, torna-se importante dividir os segmentos entre aqueles mais susceptíveis à influência de um determinado ajuste e aqueles não susceptíveis. (LAVÉ, 1980).

Nesse contexto, esta pesquisa teve como objetivo avaliar a aplicabilidade do princípio da susceptibilidade dos segmentos aos efeitos de ajustes de qualidade vocal (supralaríngeo e de tensão) na percepção visual da qualidade vocal, segundo modelo fonético de descrição da qualidade vocal. Os resultados são apresentados com base no detalhamento das imagens e nas correlações à avaliação perceptiva da qualidade vocal e discutidos quanto às particularidades da percepção visual da qualidade vocal, no campo da Prosódia Audiovisual.

A partir das reflexões anteriormente esboçadas, são ressaltadas as características de abordagem interpretativa e integrada da qualidade vocal adotadas no Grupo de Estudos sobre a Fala (CNPq) do Laboratório Integrado de Análise Acústica e Cognição (LIAAC) da Pontifícia Universidade Católica de São Paulo (PUCSP), no Projeto Abordagem Fonética da Qualidade Vocal, na linha de pesquisa Linguagem e Patologias da Linguagem do Departamento de Linguística da Faculdade de Comunicação e Filosofia.

1. Método

O *corpus* de pesquisa foi constituído por 38 estímulos em formato de vídeos referentes ao VPAS-PB (ajustes de qualidade vocal dos planos supralaríngeos e de tensão – Figura 1) constantes do banco de dados do Laboratório Integrado de Análise Acústica e Cognição (LIAAC), tendo sido autorizado o uso das imagens para a presente pesquisa (comitê de ética da instituição sob o número 101/11). Nas gravações, cada sujeito procedeu a

leitura de 10 sentenças (Figura 2) elaboradas especificamente para análise de qualidade vocal, levando-se em conta a susceptibilidade dos segmentos aos efeitos de ajustes (CAMARGO, MADUREIRA, 2008), segundo modelo fonético. Para cada falante, também constou uma amostra de vídeo de um trecho de fala semi-espontânea, na qual cada falante foi solicitado a falar sobre a cidade em que nasceu.

QUALIDADE VOCAL	PRIMEIRA PASSADA		SEGUNDA PASSADA						
	Neutro	Não neutro	AJUSTE	Moderado			Extremo		
				1	2	3	4	5	6
A. ELEMENTOS DO TRATO VOCAL									
1. Lábios			<i>Arredondados/protraídos</i>						
			Estirados						
			Labiodentalização						
			<i>Extensão diminuída</i>						
			<i>Extensão aumentada</i>						
2. Mandíbula			<i>Fechada</i>						
			Aberta						
			Protraída						
			<i>Extensão diminuída</i>						
			<i>Extensão aumentada</i>						
3. Língua ponta/lâmina			<i>Avançada</i>						
			<i>Recuada</i>						
4. Corpo de língua			<i>Avançado</i>						
			<i>Recuado</i>						
			Elevado						
			Abaixado						
			<i>Extensão diminuída</i>						
			<i>Extensão aumentada</i>						
5. Faringe			Constricção						
			Expansão						
6. Velofaringe			Escape nasal audível						
			Nasal						
			Denasal						
7. Altura de laringe			Elevada						
			Abaixada						
B. TENSÃO MUSCULAR GERAL									
8. Tensão do trato vocal			Hiperfunção						
			Hipofunção						
9. Tensão laríngea			Hiperfunção						
			Hipofunção						

Figura 1. Roteiro *Voice Profile Analysis Scheme* para o Português Brasileiro (VPAS-PB) referente aos grupos de ajustes supralaríngeos e de tensão.

Sentenças Chave	
1	O objeto de estudo da fonética é essa complexa, variável e poderosa face sonora da linguagem: a fala.
2	Na cidade de São Paulo a contribuição que cada grupo étnico ou regional deu à cidade é vista em cada esquina. Italianos, japoneses, árabes, judeus, portugueses, coreanos e pessoas de todo o país ajudaram e muito a construir esta metrópole. São Paulo é hoje uma metrópole cosmopolita um lugar onde todos se sentem em casa.
3	A Roberta gosta muito de comprar livros de pássaros. Ela também costuma ir ao jardim zoológico para ver suas aves preferidas: a arara, a garça, o sabiá, o periquito, o tico-tico, a coruja e o tucano.
4	A Lara guarda figuras de pássaros em uma caixa e suas preferidas são a da arara, da patativa, da garça, do canário e do sabiá amarelo.
5	Liliane diverte-se imitando os trinos do periquito, do bicudo, do bem-te-vi e do tico-tico.
6	O garoto tirou muitas fotografias do tucano, da coruja, do pombo e do jaburu.
7	Soube que a casa dos bispos é visitada por turistas todos os dias e que o roteiro de visita dura cerca de duas horas para ser percorrido.
8	Detesto ir à casa dele, pois fica do outro lado da cidade e o acesso é difícil.
9	Não mencionei anteriormente, mas minha mãe morou muitos anos em Santos, numa mansão à beira-mar.
10	Em uma palestra para professores de filosofia, um famoso filósofo da linguagem atacou, de maneira implacável, o que considerava os equívocos e confusões da linguagem. Ao longo da palestra o filósofo foi trazendo à tona suas reflexões de natureza pessoal, bem como suas memórias e angústias pessoais.

Figura 2. Sentenças chave para avaliação de qualidade vocal por meio do roteiro VPAS-PB

A partir dos vídeos representativos dos ajustes supralaríngeos e de tensão, selecionamos as sentenças-chave formadas por segmentos-chaves para detecção dos ajustes específicos. Tais estímulos foram editados e redimensionados por meio dos *software* de livre acesso GIMP 2.6 (<http://www.gimp.org/>). Em seguida, os vídeos foram decupados por meio dos *softwares* de livre acesso PITIVI versão 0.13.4 (<http://www.pitivi.org/>) e KINO versão 1.1.1 (<http://www.kinodv.org/>) compatíveis com sistema operacional *Linux*, o que permitiu a geração de cerca de 30 quadros por segundo de imagens.

As sequências de imagens geradas pelo processo de decupagem foram analisadas segundo as mobilizações prevalentes em quadros sucessivos e

segundo aquelas específicas em segmentos-chave, por meio de um roteiro de análise visual de ajustes de qualidade vocal (Figura 3), elaborado com base no protocolo MBGR (GENARO et al, 2009).

<i>Ajustes de qualidade vocal</i>		Segmento (s) chave	<i>Aspectos de motricidade orofacial (atividade de grupos musculares)</i>										<i>Qualidade da gravação em vídeo</i>	
Do trato vocal	De tensão		Orbicular	Músculo Mental	Músculo Abaix Ang.boca	Músculo nasal	Muscu. cervical	. Lab. Infer..evers	Músculo Levant. Lab.	Mus. Da mast. e fac.	Mus. Extrin. laringe	foco	Enquadramento	luminosidade

Figura 3. Roteiro de análise visual de ajustes de qualidade vocal (aspectos de motricidade orofacial e da qualidade da gravação).

Legenda: *Músculo Abaix.Ang. boca*= *Músculo Abaixador do ângulo da boca*. *Mus. Cervical*= *musculatura cervical*. *Lab. Infer. evers.*= *Lábio inferior com eversão*. *Músculo Levant. Lab. Sup.*= *Músculo levantador do lábio superior*. *Mus. da mast. e fac.*= *músculos da mastigação e faciais*. *Mus. Extrin. Laringe*= *Musculatura extrínseca da laringe*.

Os resultados são apresentados a partir do detalhamento das imagens e das correlações à avaliação perceptivo-auditiva com base nos diferentes segmentos-chave para detecção de ajustes específicos de qualidade vocal.

1.1 Resultados e discussão

Os ajustes de qualidade vocal susceptíveis aos efeitos de percepção visual foram aqueles de lábios (estirados, arredondados, labiodentalização e extensão diminuída), mandíbula (aberta, fechada, protraída, extensão aumentada e diminuída), ponta de língua (avançada), faringe (constricção) e posição vertical de laringe (laringe elevada) no plano supralaríngeo e de hiperfunção de laringe e de trato vocal no plano dos ajustes de tensão.

A síntese dos dados referentes à inspeção visual das amostras em vídeo é apresentada na Figura 4 com relação aos ajustes de qualidade vocal detectados na esfera perceptivo-auditiva (informações do material instrutivo do VPAS-PB) e segmentos chave para detecção dos ajustes. Na Figura 5 são apresentados alguns exemplos com imagens.

Roteiro Visual			
Grupo de ajustes	Segmento(s) Chave	Ajuste de qualidade vocal	Aspectos observados nas imagens de falantes
Do trato vocal	[p]; [b]; [m]	Labiodentalização	Tensão dos músculos: mental, levantador do lábio superior e da asa do nariz
	[i]; [u]	Lábios extensão diminuída	Músculos faciais tendendo à hipotonia funcional
	[o]; [u]	Lábios estirados	Tensão dos músculos orbicular da boca e do abaixador do ângulo da boca
	[t]; [d]; [s]; [z]	Ponta de língua avançada	Lábio inferior com eversão
	[a]; [i]; [u]	Mandíbula extensão diminuída	Músculos faciais tendendo à hipotonia funcional
	[i]; [o]; [u]	Mandíbula aberta	Tensão do músculo orbicular da boca
	[a]; [i]; [u]	Mandíbula extensão aumentada	Tensão dos músculos faciais
	[f]; [s]; [o]; [u]	Mandíbula protraída	Lábio inferior com eversão
	[i]; [u]	Constricção faríngea	Tensão muscular supraglótica Mandíbula fechada
	[p]; [t]; [k]; [f]; [s]	Escape nasal audível	Tensão do músculo orbicular da boca Movimento do músculo nasal
	Vogais	Laringe elevada	Tensão dos músculos extrínsecos da laringe
De Tensão	Vogais	Hiperfunção laríngea	Tensão da musculatura cervical Tensão dos músculos extrínsecos da laringe
	segmentos em geral	Hiperfunção do trato vocal	Tensão de musculatura cervical e facial

--	--	--	--

Figura 4. Resultados da análise visual com relação à correspondência entre ajustes de qualidade vocal e aspectos de motricidade orofacial

Ajuste: Lábios estirados
[u] estudo



Quadro 1

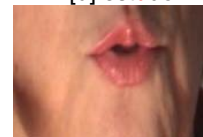


Quadro 2



Quadro 3

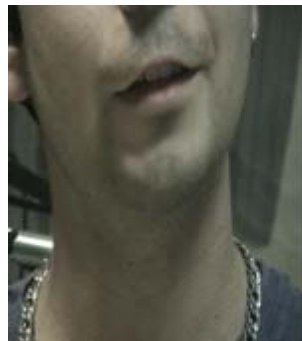
Ajuste de referência
[u] estudo



Ajuste: Mandíbula fechada
[a] face

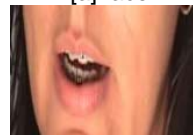


Quadro 1

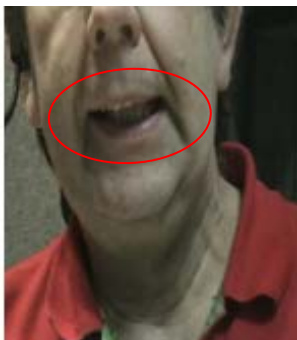


Quadro 2

Ajuste de Referência
[a] face



Ajuste: Hiperfunção de trato vocal
[e] detesto



Quadro 3

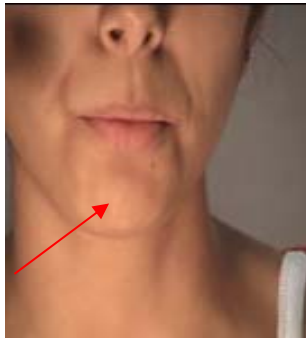


Quadro 4

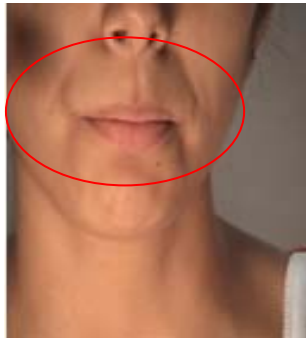
Ajuste de Referência
[e] beira



**Ajuste: Labiodentalização
[m] anteriormente**



Quadro 1



Quadro 2



Quadro 3

Ajuste de
Referência
[m] imitando



Figura 5. Resultados da análise visual com relação à correspondência entre ajustes de qualidade vocal.

2. Conclusão

Os achados reforçam o princípio da susceptibilidade dos segmentos aos efeitos da percepção visual de ajustes de qualidade vocal nos planos supralaríngeo e de tensão e apontam demandas específicas no registro de imagens em vídeo para a finalidade de avaliação de ajustes de qualidade vocal.

REFERÊNCIAS

ABELIN, Å. Seeing glee but hearing fear? Emotional McGurk effect in Swedish. Conference on Speech Prosody, 2008, Campinas. 2008 p713-716.

CAMARGO, Z. MADUREIRA, S. Voice quality analysis from a phonetic perspective: Voice Profile Analysis Scheme Profile for Brazilian Portuguesees (BP- VPAS) In: **Fourth Conference on Speech Prosody**, 2008, Campinas: Capes, FAPESP, CNPq, V.1. p 57-60. 2008.

GENARO, K.F. et al. Avaliação Miofuncional Orofacial- Protocolo MBGR. Revista CEFAC, São Paulo, p. 237-255, Abr-Jun. 2009.

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

WHALEN, D.H. **Of hearing by Eye II: Advances in the Psychology os speechreading and audition-Visual Speech**. Am J Psychol ,p 631-640. 2002.