

Análise acústica da voz captada na faringe próximo à fonte glótica através de microfone acoplado ao fibrolaringoscópio

Acoustic analysis of voice captured in the pharynx above the glottic source through a microphone on a laryngo-fiberscope

Erica E. Fukuyama¹

Palavras-chave: distúrbios da voz, qualidade da voz, glote, laringoscopia, análise acústica.

Key words: voice disorders, voice quality, laryngoscopy, acoustic analysis, glottis.

Resumo / Summary

Objetivo: O objetivo deste trabalho foi estudar a voz próximo à sua fonte produtora, as pregas vocais, através de um microfone miniaturizado de aparelho auditivo que foi adaptado para ser acoplado à extremidade de um fibrolaringoscópio, permitindo a captação da voz durante a laringoscopia direta. **Forma de estudo:** Experimental. **Material e Método:** A voz foi estudada em um grupo de 50 indivíduos, 25 homens e 25 mulheres sem doenças, através de um programa de análise acústica MDVP (*Multi-Dimensional Voice Program*) do laboratório de voz *Computerized Speech Lab*, Model 4300B, da *Kay Elemetrics*. Amostras de vogais sustentadas /a/, /i/ e /u/ foram captadas de três formas diferentes, primeiramente com um microfone comum externo a 15 cm da boca, em segundo lugar com o microfone especial na faringe a 1,5 cm acima das pregas vocais e por último com o microfone especial externamente a 2 cm da boca. Doze parâmetros acústicos relacionados a frequência fundamental, amplitude e ruído de cada uma das vogais foram comparadas estatisticamente conforme à sua forma de captação. **Resultados:** Os resultados mostraram diferenças estatisticamente significativas entre a voz captada pelo microfone comum externo e o microfone especial, em relação à frequência fundamental, aos parâmetros de variação de periodicidade de frequência, amplitude e ruído. **Conclusão:** A diferença do som da fonte glótica do som da voz externa pode mostrar as modificações sofridas pela voz no decorrer da passagem pelo trato vocal.

Aim: The aim of the present study is to examine the voice to its acoustic source – the vocal folds – with a miniature hearing-aid microphone coupled to the extremity of a laryngo-fiberscope allowing the voice to be captured during direct laryngoscopy. **Study design:** Experimental. **Material and Method:** The voice of 50 individuals – 25 males and 25 females bearing no pathologies – was collected by the Multi-Dimensional Voice Program (MDVP) by Kay Elemetrics' Computerized Speech Lab 4300B Model. Samples of the sustained vowels /a/, /i/ and /u/ were picked up in three distinct ways. Firstly, by a common external microphone placed at 15 cm from the mouth. Secondly, a special microphone was placed on the pharynx 1.5 cm above the vocal folds. Lastly, the same special microphone was placed externally at 2 cm from the mouth. Twelve acoustic parameters regarding fundamental frequency, amplitude and noise of each and every vowel were compared statistically as to the way the voice was picked up. **Results:** Results show statistically significant differences between the voice picked up by the common external microphone and by the special one as regards to the fundamental frequency, frequency and amplitude variability and noise. **Conclusion:** The difference between the sound coming from the glottic source and the sound from the external voice shows alterations experienced by the voice during its passage through the vocal tract.

¹ Especialista em Cirurgia de Cabeça e Pescoço – Instituto do Câncer “Arnaldo Vieira de Carvalho”, SP
Faculdade de Medicina da Universidade de São Paulo, SP
Rua Martinico Prado 26, cj 151 – 15º andar – Higienópolis, São Paulo – 01224-010
Tel/fax: (11) 3337-7813 celular: 9702-2645 – E-mail: tamaris@ig.com.br
Artigo recebido em 25 de abril de 2001. Artigo aceito em 29 de junho de 2001.

INTRODUÇÃO

O ser humano é o único ser capaz de produzir voz. Através deste som articulado expressamos nossos pensamentos, sentimentos e transmitimos nossas vontades. Por ser a fala o meio de expressão e comunicação mais importante, qualquer distúrbio da voz pode ter profundas implicações na vida social e profissional de uma pessoa.

O aparelho fonador é constituído pelo aparelho respiratório, pela laringe (a fonte de vibração) e o trato vocal (o sistema ressonador composto pela faringe, boca e nariz). O fluxo aéreo respiratório, ao passar pelos ciclos de abertura e fechamento das pregas vocais, constituirá uma vibração que irá ressonar pelo trato vocal.

O diagnóstico dos distúrbios fonatórios depende de uma abordagem multidisciplinar, e para um diagnóstico preciso deve-se realizar a anamnese minuciosa, o exame loco-regional, a avaliação do comportamento vocal e a análise psicodinâmica.

A laringe e o trato vocal podem ser sede de lesões benignas e malignas, além de distúrbios neurológicos que alteram a motilidade e a sensibilidade destes órgãos.

A avaliação da voz é um referencial para a terapia fonoaudiológica, para averiguar o progresso do tratamento instituído, além de ser muito importante para o entendimento da fisiologia dos órgãos da voz. Esta avaliação era inicialmente realizada de forma subjetiva através de análise perceptiva auditiva, mas a falta de consenso entre examinadores experientes, mesmo com o uso de diferentes escalas de alterações vocais, tornou necessária a pesquisa de um tipo de avaliação objetiva, na qual a voz fosse analisada através de aparelhos que são capazes de medir vários parâmetros acústicos da voz.

A teoria de produção da fala baseia-se na teoria linear fonte-filtro de Gunnar Fant (1970)⁸, na qual a fonte é a vibração laríngea e o filtro (sistema de transmissão seletiva ou sistema ressonador) é o trato vocal. A laringe é um transdutor de energia aerodinâmica (fluxo de ar) em acústica, através dos ciclos de abertura e fechamento das pregas vocais.

A ressonância do tubo ressonador depende de seu comprimento e de seu diâmetro transversal (Kent, 1992)¹⁴. Os formantes são o padrão de frequências das ressonâncias do trato vocal, onde se concentram as maiores energias acústicas. Através dos formantes podemos reconhecer as características segmentais da fala. Os três primeiros formantes são os principais determinantes da qualidade fonética de uma vogal (Camargo, 2000³; Kent, 1992¹⁴).

A fonte de energia e o ressonador são independentes, a frequência de vibração das pregas vocais não afeta as propriedades do ressonador.

Na avaliação objetiva da voz é necessário o estudo de vários seguimentos envolvidos no processo de fonação como o fluxo aéreo pulmonar, a atividade laríngea e a ressonância.

A história da análise acústica da fala começou neste século com métodos analógicos até chegar aos modernos processos de digitalização. Em 1920 surgiu o oscilograma que se constitui num gráfico da amplitude em relação ao tempo num sinal contínuo. Em 1940, surgiu o analisador de Henrici com análise de Fourier e em 1950 a espectrografia que é a análise da onda complexa em seus componentes de frequência pelo uso de filtros, que mostra a concentração de energia da fala ao longo do tempo. Em 1970, uma nova era se iniciou com a introdução da tecnologia digital (Kent, 1992)¹⁴.

Desde a década de 50, houve um aumento progressivo de estudos voltados à análise objetiva da voz, paralelamente ao aparecimento de teorias que explicam a produção da voz e desenvolvimento de inúmeros laboratórios de voz.

Existem vários instrumentos de análise, chamados dedicados ou baseados quando contém a própria placa de som, e não-dedicados caso os programas sejam acoplados à placa de som de computadores (Camargo, 2000)³. Dentre os diversos laboratórios de voz disponíveis comercialmente nos Estados Unidos, temos o CSpeech desenvolvido por Paul Milenkovich, *Computadorized Speech Laboratory da Kay Elemetrics* e o *SoundScope* (GW Instruments), (Bielamowicz, 1996)².

As amostras de voz necessárias dependem do objetivo da análise. As vogais sustentadas podem ser usadas para casos de patologias, e para padronização de novas medidas e parâmetros (Horii, 1979)¹². A fala encadeada verifica o impacto da alteração vocal nas situações de comunicação oral. A vogal em posição de acentuação frasal, por exigir maior atividade laríngea, detecta dificuldades de controle laríngeo (Camargo, 2000)³. As porções iniciais e finais de emissão sustentada podem revelar maiores graus de perturbação e instabilidade (De Krom, 1995)⁶.

Outros cuidados técnicos são necessários, como a gravação digital, a repetição das análises e o tipo do microfone a ser utilizado. A distância do microfone para a captação da voz, uma vez padronizada para o programa a ser usado, deve ser sempre a mesma, e o trabalho de Titze & Winholtz (1993)²⁶ demonstrou a superioridade do microfone tipo condensador em relação ao dinâmico.

A onda sonora, neste caso a voz, possui três qualidades físicas, a altura vocal ou frequência fundamental (hertz-Hz) relacionada à frequência de vibração das pregas vocais, a intensidade vocal ou amplitude (decibel-dB) que depende da pressão sub-glótica e da energia transportada pela onda sonora e o timbre, que é a qualidade fornecida pela combinação harmônica do som, decorrente das características da fonte sonora que o produziu.

Os principais parâmetros acústicos utilizados atualmente são os seguintes:

- A. Medidas de Freqüência fundamental (Fo – Hz) – corresponde ao número de ciclos glóticos por segundo. Segundo Behlau e Pontes (1995)¹, os valores normais são de 80 a 150Hz para os homens, 150 a 250Hz para as mulheres e acima de 250Hz para as crianças;
- B. Medidas de perturbação – referem-se a quanto um determinado período de vibração diferencia-se do outro que o sucede, quanto à freqüência e amplitude, representam o nível de estabilidade vibratória e dependem do controle sobre o sistema fonatório. O *jitter* representa a variação de periodicidade de freqüência e o *shimmer* a variação de periodicidade de amplitude;
- C. Medidas de Ruído – avaliam o ruído em diferentes faixas de freqüência do espectro. O ruído é um sinal acústico sem periodicidade, originado da superposição de vários movimentos de vibração com diferentes freqüências, as quais não apresentam relação entre si.

Os primeiros trabalhos envolvendo a análise acústica da voz foram os de Saito, Kato & Teranishi (1958)²⁵ que analisaram freqüência fundamental na fala japonesa, de Von Leden e colaboradores (1958)²⁷ que realizaram o primeiro trabalho importante à respeito da extração da freqüência fundamental e parâmetros de variação de periodicidade (*Jitter* e *Shimmer*), de Lieberman (1961)¹⁷ que estudou a perturbação de *pitch*, uma rápida variação da periodicidade fundamental de uma fala encadeada, de Risberg (1962)²¹ que avaliou a variação da freqüência fundamental na fala encadeada em língua sueca e inglesa, de Lieberman (1963)¹⁸ que estudou 32 indivíduos entre 24 e 71 anos, e avaliou a periodicidade da freqüência fundamental em um computador IBM; a perturbação de *pitch* em uma fala encadeada reflete alterações no formato da glote e na periodicidade da glote e Wendahl (1966)²⁸ que correlacionou *jitter* e aspereza vocal.

Rontal (1975)²³ defendia a qualidade visual dos espectrogramas, de melhor compreensão do que a quantificação exata dos parâmetros acústicos, porém os mesmos quantificados têm sido amplamente estudados ao longo da história da análise acústica, visando a elucidação dos correlatos acústicos das alterações vocais (Wendahl, 1966²⁸; Koike, 1977¹⁵; Hammarberg e colaboradores, 1980¹⁰; Yumoto & Gould, 1982³⁰; Eskenazi e colaboradores, 1990⁷; De Krom, 1995⁶; Hillenbrand, 1996¹¹; Omori e colaboradores 1997²⁰; Michaelis e colaboradores, 1998¹⁹). Existe uma interdependência entre os parâmetros, os quais são sensíveis às várias propriedades acústicas, o que torna difícil sua interpretação, levando a resultados contraditórios na literatura.

Sumariamente, os parâmetros de variação de periodicidade estão mais relacionados à rouquidão, e os parâmetros de ruído estão relacionados com a soproidade e fenda glótica. No estudo da voz em pacientes com patologia vocal, não há como comparar simplesmente os

dados acústicos com critérios de normalidade, é necessário o estabelecimento de novas correlações de natureza fisiológica, auditiva e acústica (Camargo, 2000)³.

Vários trabalhos foram realizados para padronização em falantes do português brasileiro. Behlau e colaboradores (1995)¹ obtiveram valores de freqüência fundamental, *Jitter* e *Shimmer* para cada um grupos de 30 homens, 30 mulheres e 30 crianças falantes do português brasileiro. Castro & Pegoraro-Krook (1993)⁴ estudaram a freqüência fundamental (Fo) de 150 indivíduos masculinos, falantes do português brasileiro, e constataram que a Fo das vogais altas /i/ e /u/ foram maiores que a da vogal baixa /a/. Castro, Pegoraro-Krook (1994)⁵ estudaram 140 indivíduos do sexo feminino falantes do português brasileiro, e os valores médios de freqüência fundamental da fala para a leitura oral e contagem de números foram respectivamente 213,9Hz e 200,1Hz. Rodrigues e colaboradores (1994)²² avaliaram a proporção harmônico-ruído (PHR) da vogal sustentada /a/ para 40 indivíduos (20 homens e 20 mulheres), e encontraram valores de 8,63 para o sexo masculino e 10,17 para o sexo feminino.

Em sua tese de doutorado, Dr. José Francisco de Góis Filho⁹ demonstrou a viabilidade de um sistema de captação vocal próximo à sua fonte produtora, as pregas vocais, estudando 10 pacientes masculinos submetidos à laringectomia supracricóideia e 10 homens com voz normal. Um microfone miniaturizado de aparelho auditivo foi adaptado para que o mesmo fosse colocado internamente ao canal de biópsia de um laringoscópio flexível; desta forma este microfone poderia captar a voz da fonte glótica, sem as alterações sofridas pela voz através do trato vocal quando é utilizado o microfone comum externo.

O objetivo deste presente estudo foi o de analisar o funcionamento e o tipo de voz obtido por este método de captação da voz com microfone em posição faríngea, comparando-o com a captação externa, através da análise acústica da voz de 50 indivíduos normais (25 homens e 25 mulheres).

Na literatura não havia ainda nenhum estudo similar, que capta e analisa o som da voz da fonte glótica, podendo trazer novos dados na compreensão da fisiologia e patologia da voz.

CASUÍSTICA E MÉTODO

No período compreendido entre agosto de 1999 a maio de 2000, no Serviço de Cirurgia de Cabeça e Pescoço do Instituto do Câncer “Arnaldo Vieira de Carvalho” (ICAVC - SP), foram examinadas 50 pessoas, 25 homens com idade variando de 20 a 58 anos (média de 32,96) e 25 mulheres com idade variando entre 25 e 40 anos (média de 36,68), sem antecedentes de alteração da voz ou problemas auditivos. A pesquisa foi aprovada pela Comissão de Ética para Análise de Projetos de Pesquisa (CAPPesq) do Hospital

das Clínicas e da Faculdade de Medicina da Universidade de São Paulo.

Todos os indivíduos foram submetidos aos exames de oroscopia, laringoscopia com telescópio laríngeo de 70°, rinofaringolaringoscopia com fibrolaringoscópio flexível, palpação do pescoço e avaliação da voz (perceptiva auditiva e acústica).

A avaliação perceptiva auditiva da voz foi realizada por duas fonoaudiólogas experientes que coletaram os dados de qualidade vocal, de ressonância, de *pitch* e de *loudness*, os quais foram normais.

A avaliação acústica foi realizada através do laboratório de voz *Computerized Speech Lab, Model 4305B* da *Kay Elemetrics*, utilizando-se microfone padrão externo do tipo dinâmico e microfone interno especialmente projetado para este estudo, utilizado em prótese auditiva.

Especificações dos microfones:

A - Microfone externo: Shure Modelo SM48S, tipo dinâmico, resposta de freqüências de 55 a 14.000 Hz, impedância de 270 ohms, nível de saída: (a 1kHz) de 77,5 dB(0,13mV), ruído de 32 dB equivalente SLP por milioersted(60Hz) e peso de 672 g;

B - Microfone especial: Knowles Modelo EM 3046, tipo condensador electreto, resposta de freqüências de 100 a 10.000 Hz, impedância de 4400 ohms, nível de saída (a 1 kHz) de 97,0 dB(0,4mV), ruído de 31 dB equivalente SLP(a 1KHz) e peso de 0,08 g.

O microfone especial foi encapado com um envoltório de silicone para evitar a entrada de saliva no sistema.

A voz foi captada através de três formas, utilizando-se as vogais sustentadas /a, i, u/:

1. Microfone externo padrão do *Computerized Speech Lab* (CSL) – os pacientes foram colocados em pé, com os braços estendidos ao longo do corpo, e microfone à 15 cm da boca,
2. Microfone especial em posição externa (MIE) – os pacientes foram colocados em pé, com os braços estendidos ao longo do corpo, e microfone à 2 cm da boca,
3. Microfone especial em posição interna (MII) – este microfone foi colocado internamente ao canal de biópsia do fibrolaringoscópio, e captou a voz a 1,5 cm acima da glote.

Os pacientes foram instruídos a realizar respiração profunda antes de cada emissão e a sustentar as vogais /a/, /i/ e /u/ de forma mais natural e no maior tempo possível. Foram escolhidas as emissões mais representativas quanto à freqüência, intensidade e qualidade, portanto as porções mais estáveis, eliminando-se o início e o final de cada emissão. A duração foi de 3 segundos para a extração dos parâmetros acústicos.

Os dados vocais foram armazenados no computador, gravados em CD e analisados através do programa

Multi-Dimensional Voice Program (MDVP) – *Model 4305* da *Kay Elemetrics*. Este programa calcula 32 parâmetros acústicos, e ainda compara estes parâmetros a um extenso banco de dados de vozes, dispondo-os em forma de gráfico que possibilita rápida identificação de parâmetros normais e alterados.

Dentre os 32 parâmetros foram escolhidos 12 que representam acusticamente a voz, de forma satisfatória, para fins de comparação entre os 3 métodos de captação.

Os parâmetros selecionados foram os seguintes:

- A. Fo – Freqüência fundamental (Hz);
- B. Medida de perturbação de freqüência fundamental - PPQ(%) *Pitch Period Perturbation Quotient*;
- C. Medida de perturbação de amplitude - APQ(%) *Amplitude Perturbation Quotient*;
- D. Medidas de ruído – NHR (*Noise to Harmonic Ratio*), VTI(*Voice Turbulence Index*) e SPI (*Soft Phonation Index*).

A comparação estatística foi realizada entre os mesmos parâmetros captados das três formas diferentes para uma mesma vogal. A análise estatística utilizada foi a Análise de Variância (ANOVA), e o nível de significância (F) de cada comparação foi de 5%. Quando F foi menor que 0,05, houve diferença significativa. O coeficiente de variação até 30%, mostrou a homogeneidade da amostra.

RESULTADOS

As Tabelas a seguir apresentam os resultados comparativos entre os parâmetros acústicos das 3 vogais sustentadas captadas de três formas diferentes:

- CSL – microfone do **Computerized Speech Lab** em posição externa padronizada para o Laboratório de voz da Kay Elemetrics;
- MII – microfone especial em posição interna;
- MIE – microfone especial em posição externa.

É apresentada a média, o desvio padrão, o coeficiente de variação, o nível de significância F e a comparação final.

DISCUSSÃO

Que transformações o som da fonte glótica sofre ao passar pelo trato vocal? A comparação entre o som captado próximo às pregas vocais e a voz final captada no meio externo pode responder a esta pergunta. Partindo deste princípio necessitávamos de um dispositivo que medisse separadamente a energia da fonte glótica. Na literatura não encontramos nenhum dispositivo semelhante ao anteriormente citado. Existem métodos indiretos de avaliação da vibração glótica como a glotografia com filtragem inversa e a eletroglotografia (Camargo, 2000)³.

Prof. Dr. Pedro Luiz Mangabeira Albernaz pensou em utilizar um microfone de prótese auditiva, e o Dr. José

Tabela 1. Média da frequência fundamental (fo) das vogais /a/ /i/ /u/ em homens.

Vogal	Estatística	CSL	MIE	MII	Significância	Resultados
A	Média	130,19	131,70	129,33	0,6285	CSL = MIE = MII
	Desvio-padrão	19,17	19,24	17,92		
	Coef.de variação	14,72	14,61	13,86		
I	Média	151,11	149,71	143,73	0,1313	CSL = MIE = MII
	Desvio-padrão	24,96	25,98	20,95		
	Coef.de variação	16,52	17,35	14,58		
U	Média	156,12	153,22	147,03	0,0304 *	CSL = MIE > MII
	Desvio-padrão	27,80	27,59	24,36		
	Coef.de variação	17,81	18,01	16,56		

Tabela 2. Média de PITCH PERIOD PERTURBATION QUOTIENT (PPQ) das vogais /a/ /i/ /u/ em homens.

Vogal	Estatística	CSL	MIE	MII	Significância	Resultados
A	Média	0,29	0,35	0,40	0,0183 *	CSL < MII
	Desvio-padrão	0,14	0,15	0,14		
	Coef.de variação	48,28	42,86	35,00		
I	Média	0,30	0,56	0,49	0,0105 *	CSL < MIE = MII
	Desvio-padrão	0,24	0,43	0,19		
	Coef.de variação	80,00	76,79	38,78		
U	Média	0,51	0,77	0,50	0,0015 *	MIE > CSL = MII
	Desvio-padrão	0,27	0,37	0,18		
	Coef.de variação	52,94	48,05	36,00		

Tabela 3. Média de AMPLITUDE PERTURBATION QUOTIENT (APQ) das vogais /a/ /i/ /u/ em homens.

Vogal	Estatística	CSL	MIE	MII	Significância	Resultados
A	Média	2,67	2,30	0,87	< 0,0001 *	CSL > MIE > MII
	Desvio-padrão	0,79	0,66	0,91		
	Coef.de variação	29,59	28,70	104,6		
I	Média	2,10	2,61	0,36	< 0,0001 *	MIE > CSL > MII
	Desvio-padrão	0,75	0,90	0,25		
	Coef.de variação	35,71	34,48	69,44		
U	Média	1,36	2,16	0,38	< 0,0001 *	MIE > CSL > MII
	Desvio-padrão	0,36	0,93	0,25		
	Coef.de variação	26,47	43,06	65,79		

Tabela 4. Média de NOISE TO HARMONIC RATIO (NHR) das vogais /a/ /i/ /u/ em homens.

Vogal	Estatística	CSL	MIE	MII	Significância	Resultados
A	Média	0,14	0,15	0,14	0,0059 *	MII = CSL < MIE
	Desvio-padrão	0,02	0,02	0,02		
	Coef.de variação	14,29	13,33	14,29		
I	Média	0,13	0,16	0,12	< 0,0001 *	MII = CSL < MIE
	Desvio-padrão	0,02	0,03	0,02		
	Coef.de variação	15,38	18,75	16,67		
U	Média	0,12	0,14	0,12	0,0115 *	MII = CSL < MIE
	Desvio-padrão	0,02	0,03	0,02		
	Coef.de variação	16,67	21,43	16,67		

Tabela 5. Média de VOICE TURBULENCE INDEX (VTI) das vogais /a/ /i/ /u/ em homens.

Vogal	Estatística	CSL	MIE	MII	Significância	Resultados
A	Média	0,05	0,05	0,04	0,0089 *	CSL = MIE > MII
	Desvio-padrão	0,01	0,02	0,02		
	Coef.de variação	20,00	40,00	50,00		
I	Média	0,11	0,08	0,03	0,0460 *	CSL = MIE > MII
	Desvio-padrão	0,18	0,03	0,01		
	Coef.de variação	163,64	37,50	33,33		
U	Média	0,04	0,04	0,06	0,7092	CSL = MIE = MII
	Desvio-padrão	0,06	0,01	0,13		
	Coef.de variação	150,00	25,00	216,67		

Tabela 6. Média de SOFT PHONATION INDEX (SPI) das vogais /a/ /i/ /u/ em homens.

Vogal	Estatística	CSL	MIE	MII	Significância	Resultados
A	Média	8,81	10,72	17,95	0,0176 *	CSL = MIE < MII
	Desvio-padrão	4,75	4,31	19,51		
	Coef.de variação	53,92	40,21	108,69		
I	Média	4,52	4,55	19,08	0,0001 *	CSL = MIE < MII
	Desvio-padrão	3,74	3,61	21,09		
	Coef.de variação	82,74	79,34	110,53		
U	Média	36,60	28,34	21,68	0,0412 *	CSL > MII
	Desvio-padrão	22,57	16,96	27,53		
	Coef.de variação	61,67	59,84	126,98		

Tabela 7. Média da frequência fundamental (fo) das vogais /a/ /i/ /u/ em mulheres.

Vogal	Estatística	CSL	MIE	MII	Significância	Resultados
A	Média	222,96	220,22	206,22	0,0013 *	CSL = MIE > MII
	Desvio-padrão	43,64	36,06	33,45		
	Coef.de variação	19,57	16,37	16,22		
I	Média	254,69	247,62	227,35	< 0,0001 *	CSL = MIE > MII
	Desvio-padrão	50,31	45,41	35,77		
	Coef.de variação	19,75	18,34	15,73		
U	Média	257,58	252,28	233,40	< 0,0001 *	CSL = MIE > MII
	Desvio-padrão	53,24	45,02	41,83		
	Coef.de variação	20,67	17,85	17,92		

Tabela 8. Média de PITCH PERIOD PERTURBATION QUOTIENT (PPQ) das vogais /a/ /i/ /u/ em mulheres.

Vogal	Estatística	CSL	MIE	MII	Significância	Resultados
A	Média	0,30	0,41	0,51	< 0,0001 *	CSL < MIE < MII
	Desvio-padrão	0,14	0,12	0,11		
	Coef.de variação	46,67	29,27	21,57		
I	Média	0,21	0,51	0,90	< 0,0001 *	CSL < MIE < MII
	Desvio-padrão	0,14	0,34	0,53		
	Coef.de variação	66,67	66,67	58,89		
U	Média	0,45	0,86	0,71	< 0,0001 *	CSL < MII < MIE
	Desvio-padrão	0,25	0,39	0,17		
	Coef.de variação	55,56	45,35	23,94		

Tabela 9. Média de AMPLITUDE PERTURBATION QUOTIENT (APQ) das vogais /a/ /i/ /u/ em mulheres.

Vogal	Estatística	CSL	MIE	MII	Significância	Resultados
A	Média	2,10	2,03	0,61	< 0,0001 *	CSL = MIE > MII
	Desvio-padrão	0,58	0,58	0,55		
	Coef.de variação	27,62	28,57	90,16		
I	Média	1,93	2,35	0,40	< 0,0001 *	MIE > CSL > MII
	Desvio-padrão	0,43	0,83	0,47		
	Coef.de variação	22,28	35,32	117,50		
U	Média	1,30	2,07	0,32	< 0,0001 *	MIE > CSL > MII
	Desvio-padrão	0,33	0,85	0,12		
	Coef.de variação	25,38	41,06	37,50		

Tabela 10. Média de NOISE TO HARMONIC RATIO (NHR) das vogais /a/ /i/ /u/ em mulheres.

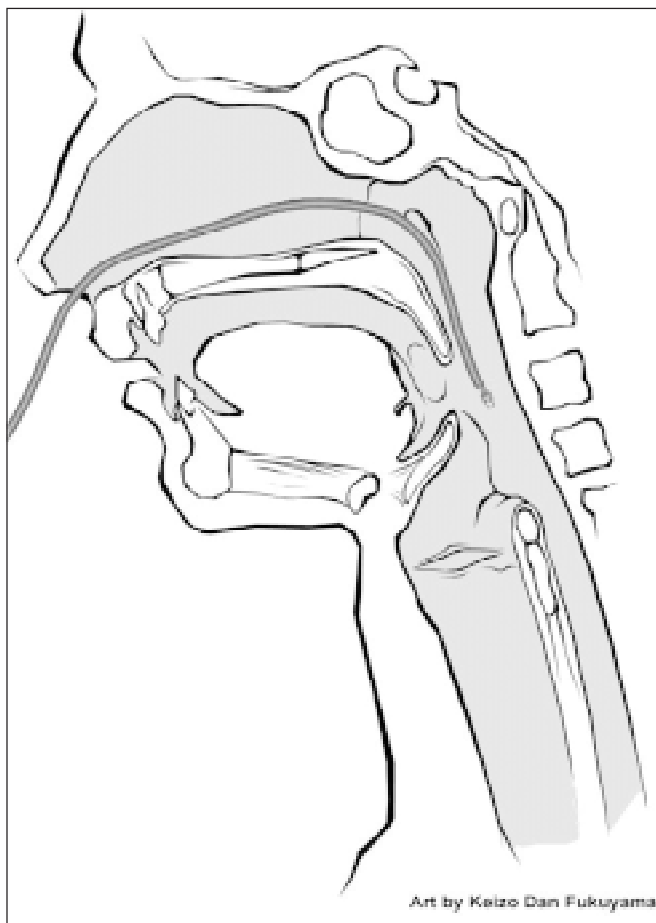
Vogal	Estatística	CSL	MIE	MII	Significância	Resultados
a	Média	0,12	0,16	0,12	<0,0001 *	MIE > CSL = MII
	Desvio-padrão	0,02	0,03	0,02		
	Coef.de variação	16,67	18,75	16,67		
i	Média	0,12	0,17	0,12	<0,0001 *	MIE > CSL = MII
	Desvio-padrão	0,02	0,04	0,03		
	Coef.de variação	16,67	23,53	25,00		
u	Média	0,11	0,15	0,11	<0,0001 *	MIE > CSL = MII
	Desvio-padrão	0,02	0,03	0,02		
	Coef.de variação	18,18	20,00	18,18		

Tabela 11. Média de VOICE TURBULENCE INDEX (VTI) das vogais /a/ /i/ /u/ em mulheres.

Vogal	Estatística	CSL	MIE	MII	Significância	Resultados
a	Média	0,06	0,06	0,03	< 0,0001 *	CSL = MIE > MII
	Desvio-padrão	0,02	0,02	0,02		
	Coef.de variação	33,33	33,33	66,67		
i	Média	0,08	0,07	0,03	< 0,0001 *	CSL = MIE > MII
	Desvio-padrão	0,03	0,02	0,01		
	Coef.de variação	37,50	28,57	33,33		
u	Média	0,03	0,04	0,04	0,5536	CSL = MIE = MII
	Desvio-padrão	0,03	0,01	0,06		
	Coef.de variação	100,00	25,00	150,00		

Tabela 12. Média de SOFT PHONATION INDEX (SPI) das vogais /a/ /i/ /u/ em mulheres.

Vogal	Estatística	CSL	MIE	MII	Significância	Resultados
a	Média	4,16	5,57	14,17	0,0014 *	CSL = MIE < MII
	Desvio-padrão	2,01	2,66	17,46		
	Coef.de variação	48,32	47,76	123,22		
i	Média	3,53	3,85	9,36	< 0,0001 *	CSL = MIE < MII
	Desvio-padrão	2,84	2,16	5,35		
	Coef.de variação	80,45	56,10	57,16		
u	Média	28,54	27,05	10,22	< 0,0001 *	CSL = MIE > MII
	Desvio-padrão	21,70	14,41	7,00		
	Coef.de variação	76,03	53,27	68,49		



Posicionamento do Microfone Especial dentro da Faringe.

Francisco de Góis Filho obteve na Audibel da Philips um protótipo que permitiu colocar um microfone especial auditivo na ponta do fibrolaringoscópio, com seus fios passando internamente ao canal de biópsia, este protótipo foi testado inicialmente por Góis Filho (2000)⁹ em 10 homens normais e 10 homens submetidos a laringectomia parcial supracricóideia no Serviço de Cirurgia de Cabeça e Pescoço do Instituto do Câncer “Arnaldo Vieira de Carvalho”, demonstrando a viabilidade do microfone.

Neste estudo o funcionamento do microfone especial foi analisado comparando o som captado em duas situações diferentes (internamente a 1,5 cm acima da glote e externamente a 2 cm da boca) com o som captado externamente a 15cm da boca através de um microfone padrão utilizado no Laboratório de voz Computerized Speech Lab da Kay Elemetrics.

A análise dos resultados obtidos requer uma visão multidimensional, desde a fisiologia dos órgãos da voz, a acústica da fala, até o funcionamento do Laboratório de voz e do microfone especial.

Em nosso estudo o microfone especial usado foi

do tipo condensador e o microfone externo padrão do *Computerized Speech Lab* foi do tipo dinâmico. O fato de termos realizado a captação do som com o microfone especial externamente e internamente teve o intuito de compararmos a performance do microfone especial com a do microfone do *Computerized Speech Lab*. Mais uma particularidade foi a necessidade de encapar o microfone especial com um envoltório de silicone para evitar a entrada de saliva no sistema.

O Laboratório de voz realiza a análise acústica que consiste num processo de extração e quantificação dos dados objetivos do sinal vocal, utilizando a representação digital do mesmo. A energia da fala se estende por mais de 10 kHz e por 60 dB e variações importantes podem ocorrer num tempo de 10ms.

Através do programa de análise acústica *Multi Dimensional Voice Program*, analisando os parâmetros de frequência, amplitude, variação de periodicidade e ruído tentamos verificar o tipo de som captado pelo microfone especial.

Todas as comparações de resultados foram estatisticamente significativas (nível de significância de 5%), através da Análise de Variância (ANOVA). Para facilitar o discurso utilizaremos as abreviaturas para indicar os microfones e suas posições, como segue:

- CSL – microfone comum externo padrão do *Computerized Speech Lab*;
- MIE – microfone especial em posição externa;
- MII – microfone especial em posição interna.

A – Medidas de Frequência

A frequência fundamental (F_0) das vogais sustentadas em homens captadas pelo CSL para as vogais /a/, /i/ e /u/, respectivamente 130,19Hz, 151,11Hz, 156,12Hz estão próximas às faixas de normalidade de 80 a 150Hz para homens segundo Behlau & Pontes (1995)¹. A F_0 das vogais sustentadas em mulheres captadas pelo CSL para as vogais /a/, /i/ e /u/, respectivamente 222,96Hz, 254,69Hz, 257,58Hz também estão próximas às faixas de normalidade de 150 a 250Hz para mulheres segundo Behlau & Pontes (1995)¹. O estudo de Castro & Pegoraro-Krook (1993)⁴ demonstrou que a frequência fundamental (F_0) das vogais /i/ e /u/ (136,2 e 140Hz respectivamente) foram maiores que a da vogal /a/ (128,3Hz) em 150 indivíduos do sexo masculino falantes do português brasileiro, o mesmo ocorreu em nosso estudo, F_0 de 151,11Hz para a vogal /i/ e F_0 de 156,12Hz para a vogal /u/ e F_0 de 130,19Hz para a vogal /a/, porém com F_0 mais altas.

O CSL e o MIE apresentaram valores de F_0 comparáveis em todas as vogais em homens e mulheres, podendo concluir-se que em termos de frequência fundamental os microfones captaram sons comparáveis externamente.

O MII só apresentou Fo comparável ao CSL e MIE nas vogais /a/ e /i/ em homens, em todas as outras vogais (/u/ masculino e todas do feminino), as Fo captadas pelo MII foram estatisticamente menores. Se o CSL e o MIE tiveram Fo comparáveis, o MII também deveria captar sons com Fo comparáveis, pois teoricamente a frequência fundamental depende exclusivamente das vibrações das pregas vocais.

A fonação em vigência do exame com o fibrolaringoscópio flexível pode estar alterada por ser incômodo ao paciente. O reflexo do paciente ao corpo estranho introduzido na faringe pode ter produzido o aumento da massa e do comprimento com respectiva diminuição da tensão das pregas vocais. Este fato explicaria as Fo mais graves encontradas no MII em relação ao CSL, para as vogais mais agudas.

B – Parâmetro de variação de periodicidade de frequência (PPQ)

Todos os valores de PPQ apresentaram-se dentro da faixa de normalidade que é abaixo de 0,84%, exceto a vogal /i/ do MII e a vogal /u/ do MIE.

A grande variabilidade do MIE dificulta sua comparação, tanto com o MII, como com o CSL. Como veremos posteriormente nos parâmetros de ruído, devido à grande sensibilidade do microfone especial, este quando colocado externamente captou muito ruído, que pode ter influenciado na variabilidade do PPQ.

Em todas as vogais o PPQ do som extraído pelo CSL foi menor que o do MII, exceto na vogal /u/ masculina. O PPQ maior no MII também pode ser explicado pela dificuldade maior para sustentar uma vibração periódica pela presença do fibrolaringoscópio na faringe.

C – Parâmetro de variação de periodicidade de amplitude (APQ)

Todas as vogais, sob todos os tipos de captação apresentaram resultados de APQ dentro da normalidade que é abaixo de 3,07.

Também houve grande variabilidade dos resultados de MIE. Mas os resultados de APQ captados pelo CSL foram todos maiores que o MII. Para que o muco não obstruísse o canal de entrada do som, foi colocada uma capa de silicone no microfone com um orifício lateral, o que pode ter amortecido os efeitos de amplitude e conseqüentemente de sua variabilidade. Outra hipótese é de que o trato vocal cause uma certa turbulência na passagem do som, o que implicaria no maior APQ para o CSL e menor APQ para o MII que capta a voz logo ao ser produzido pela fonte glótica.

D – Medidas de Ruído (NHR, VTI, SPI)

O NHR (*Noise to Harmonic Ratio*) é a proporção média entre o espectro inarmônico de energia (1500 a

4500Hz) e o espectro harmônico de energia (70 a 4500Hz), este parâmetro avalia o ruído em geral, que pode estar relacionado a variações de frequência e amplitude, a ruído turbulento, a componentes sub-harmônicos e quebras de voz. Todos os resultados de NHR se mostraram dentro da normalidade que é abaixo de 0,19. Em todos os casos os resultados de NHR do som captado pelo CSL e MII foram comparáveis, mas os resultados de MIE foram maiores que as outras duas captações. Isto demonstra um incremento na captação de ruído pelo MIE, devido à sensibilidade do microfone auditivo quando é usado em ambiente aberto, o que pode prejudicar a averiguação de outros parâmetros.

O VTI (*Voice Turbulence Index*) é a proporção média entre o espectro inarmônico de energia de alta frequência (2800 a 5800Hz) e o espectro de energia harmônica (70-4500Hz), em áreas em que a influência de variação de frequência e amplitude, quebras de voz e componentes sub-harmônicos são mínimas. É um parâmetro novo relacionado a sopro, ainda pouco estudado na literatura.

Os resultados de VTI estão dentro da normalidade (<0,061), exceto a vogal /i/ masculina captada por CSL e MIE, e a vogal /i/ feminina captada por CSL e MIE. Na vogal /a/ e /i/ masculinos e femininos, o CSL e o MIE foram semelhantes, mas ambos foram maiores que o MII. Na vogal /u/ o VTI foram semelhante nas três captações. A sopro geralmente está relacionada à perda de adução ou adução incompleta das pregas vocais. Se isto fosse correto os VTI que se apresentaram alterados na vogal /i/ feminina e masculina captadas pelo CSL e pelo MIE, também deveriam estar alterados nos VTI da vogal /i/ captadas pelo MII. Portanto o componente inarmônico da vogal /i/ pode estar relacionado ao trato vocal e não à prega vocal.

Conceitualmente o VTI mede o ruído de frequência alta e está relacionado à turbulência causada por adução incompleta ou incorreta das pregas vocais. Neste caso, pode estar havendo uma amplificação dos ruídos de alta frequência pelo trato vocal.

O SPI (*Soft Phonation Index*) é sensível à estrutura dos formantes, pois é a proporção média entre a energia harmônica de baixa frequência (70 a 1600Hz) e a energia harmônica de alta frequência (1600 a 4500Hz). Os formantes são as ressonâncias do trato vocal, isto é, as faixas de frequências (dentro de uma constituição harmônica) que concentram maior energia acústica. A mudança da forma e da elasticidade do trato vocal proporciona combinações acústicas variadas, o que resulta na grande variedade de sons da fala (Russo, 1999)²⁴.

A padronização no programa é para a vogal /a/, portanto devemos analisar apenas esta vogal. No CSL e no MIE foram comparáveis, mas no MII o SPI foi maior. O CSL e o MIE captou uma estrutura de formantes semelhante, enquanto o MII captou um som com uma

estrutura diferente, com uma proporção de energia harmônica de baixa frequência maior.

Apesar das diferenças que podem ter decorrido durante a fonação com o fibrolaringoscópio, modificando os parâmetros de frequência fundamental e amplitude, os resultados de SPI mostram que a estrutura de formantes do som captado internamente é diferente daquele captado externamente. Em estudos futuros será necessário definirmos a estrutura de formantes do som captado internamente, através de métodos espectrográficos e seus derivados, para que possamos minimizar ao máximo a interferência do trato vocal neste som. Isto pode ser conseguido através de vários métodos, relacionados à forma de captação, ao microfone e ao Laboratório de voz.

Uma máscara para o nariz e a boca pode ser utilizada no intuito de delimitar o tubo ressonador. Se a característica do tubo for conhecida, será possível a criação de filtros que eliminará qualquer influência do trato vocal no som captado próximo à prega vocal.

Todas as inferências acima descritas são preliminares, e necessitam de um profundo estudo da acústica da voz, e testes para cada uma das hipóteses levantadas.

Se tudo nasce de uma idéia, e esta idéia necessita de um princípio para se materializar, este estudo é mais um passo rumo à compreensão da fisiologia da voz.

CONCLUSÕES

Pelas comparações realizadas entre a captação externa da voz pelo microfone padrão do *Computerized Speech Lab* (CSL), a captação externa com o microfone especial (MIE) e a captação interna com o microfone especial (MII), pudemos concluir que:

1. microfone especial testado em posição externa (MIE) demonstrou, devido à sua sensibilidade, um incremento na captação de ruído, o que alterou o resultado dos parâmetros de variação de periodicidade de frequência e amplitude. Quanto à frequência fundamental, captou um som semelhante ao captado pelo CSL.
2. microfone especial em posição interna (MII) apresentou diferenças com o MIE e o CSL. As diferenças com o CSL podem ter decorrido da fonação com o fibrolaringoscópio, modificando os parâmetros de frequência fundamental e amplitude, mas os resultados de *Soft Phonation Index* (SPI) mostraram que a estrutura de formantes do som captado internamente é diferente daquele captado externamente. Isto é, a influência do trato vocal foi menor na captação MII, mas ainda houve uma certa interferência, que necessita ser elucidada em estudos futuros.

REFERÊNCIAS BIBLIOGRÁFICAS

1. BEHLAU, M.; PONTES, P. *Avaliação e tratamento das disfonias*. São Paulo, Lovise. 312p, 1995.
2. BIELAMOWICZ, S.; KREIMAN, J.; GERRATT, B.R.; DAUER, M.S.; BERKE, G.S. Comparison of voice analysis systems for perturbation measurement. *J. Speech. Hear. Res.*, 39:126-34, 1996.
3. CAMARGO, Z. Avaliação objetiva da voz. In: A atuação da fonoaudiologia no câncer de cabeça e pescoço. São Paulo, Lovise. p.175-94, 2000.
4. CASTRO, V.C.; PEGORARO-KROOK, M.I. Medida da Frequência Fundamental das Vogais em Indivíduos do Sexo Masculino Falantes do Português Brasileiro. *Pró-Fono Revista de Atualização Científica*, 15:19-22, 1993.
5. CASTRO, V.C.; PEGORARO-KROOK, M.I. Características da Frequência Fundamental da Fala de Indivíduos do Sexo Feminino Falantes do Português Brasileiro. *Pró-Fono Revista de Atualização Científica*, 6:5-7, 1994.
6. DE KROM, G. Some spectral correlates of pathological breathy and rough voice quality for different types of vowel fragments. *J. Speech. Hear. Res.*, 38:794-811, 1995.
7. ESKENAZI, L.; CHILDERS, A.G.; HICKS, D.M. Acoustic correlates of vocal quality. *J. Speech. Hear. Res.*, 33:298-306, 1990.
8. FANT, G. Acoustic theory of speech production. Paris, Mouton. 1970 apud CAMARGO, Z. *Avaliação objetiva da voz*. In: A atuação da fonoaudiologia no câncer de cabeça e pescoço. São Paulo, Lovise, 2000. p.175-94.
9. GÓIS FILHO, J.F. *Análise acústica da voz com microfones externo e na faringe*. São Paulo, 2000. 107p. Tese (Doutorado) – Escola Paulista de Medicina, Universidade Federal de São Paulo.
10. HAMMARBERG, B.; FRITZELL, B.; GAUFFIN, J.; SUNDBERG, J.; WEDIN, L. Perceptual and acoustic correlates of abnormal voice qualities. *Acta. Otolaryngol.*, 90:441-51, 1980.
11. HILLENBRAND, J.; HOUDE, R.A. Acoustic correlates of breathy vocal quality: dysphonic voices and continuous speech. *J. Speech. Hear. Res.*, 39:311-21, 1996.
12. HORII, Y. Fundamental frequency perturbation observed in sustained phonation. *J. Speech. Hear. Res.*, 22:5-19, 1979.
13. KAY ELEMETRICS CORP. *Operations Manual: Multi-Dimensional Voice Program* (MDVP) Model 4305, New Jersey. 239p, 1993.
14. KENT, R.D.; READ, C. *The Acoustic Analysis of Speech*. San Diego, Singular Publishing Group. 238p, 1992.
15. KOIKE, Y.; TAKAHASHI, H.; CALCATERRA, T.C. Acoustic Measures for detecting laryngeal pathology. *Acta. Otolaryngol.*, 84:105-17, 1977.
16. KREIMAN, J.; GERRATT, B.R.; KEMPSTER, G.B.; ERMAN, A.; BERKE, C.S. Perceptual evaluation of voice quality: review, tutorial, and a framework for future research. *J. Speech. Hear. Res.*, 36:21-40. 1993.
17. LIEBERMAN, P. Perturbations in Vocal Pitch. *J. Acoust. Soc. Am.*, 33:597-603, 1961.
18. LIEBERMAN, P. Some acoustic measures of the fundamental periodicity of normal and pathologic larynges. *J. Acoust. Soc. Am.*, 35:344-53, 1963.
19. MICHAELIS, D.; FROHLICH, M.; STRUBE, H.W. Selection and combination of acoustic features for the description of pathologic voices. *J. Acoust. Soc. Am.*, 103:1628-39, 1998.
20. OMORI, K.; KOJIMA, H.; KAKANI, R.; SLAVIT, D.H.; BLAUGRUND, S.M. Acoustic characteristics of rough voice: subharmonics. *J. Voice*, 11:40-47, 1997.
21. RISBERG, A. Statistical studies of fundamental frequency range and rate of change. In: Speech Transmission Lab. Quart. Prog. And Status Rept., Royal Institute of Technology. Stockholm, 1962. P.7-8 apud LIEBERMAN P. Some acoustic measures of the fundamental periodicity of normal and pathologic larynges. *J. Acoust. Soc. Am.*, 35:344-53, 1963.
22. RODRIGUES, S.; BEHLAU, M.S.; PONTES, P.A.L. Proporção harmônico-ruído: valores para indivíduos adultos brasileiros. *Acta AWHO*, 13:112-6, 1994.

-
23. RONTAL, E.; RONTAL, M.; ROLNICK, M.I. Objective evaluation of vocal pathology using voice spectrography. *Ann. Otol. Rhinol. Laryngol.*, 84:662-71, 1975.
 24. RUSSO, I.C.P. *Acústica e Psicoacústica aplicadas à fonoaudiologia*. 2. ed. São Paulo, Lovise. 263p, 1999.
 25. SAITO, S.; KATO, K.; TERANISHI, N. Statistical Propertires of the Fundamental Frequencies of Japanese Speech Voices. *J. Acoust. Soc. Japan*, 14:111, 1958 apud LIEBERMAN P. Some acoustic measures of the fundamental periodicity of normal and pathologic larynges. *J. Acoust. Soc. Am.*, 35:344-53, 1963.
 26. TITZE, I.R.; WINHOLTZ, W.S. Effect of microphone type and placement on voice perturbation measurements. *J. Speech. Hear. Res.*, 36:1177-90, 1993.
 27. VON LEDEN, H.; TIMKE, A.; MOORE, P. Laryngeal vibrations: measurements of the glottic wave. Part I. The normal vibration cicle. *Arch. Otolaryngol.*, 68:1-19, 1958 apud LIEBERMAN P. Some acoustic measures of the fundamental periodicity of normal and pathologic larynges. *J. Acoust. Soc. Am.*, 35:344-53, 1963.
 28. WENDAHL, R.W. Laryngeal analog synthesis of jitter and shimmer auditory parameters of harshness. *Folia Phoniat.*, 18:98, 1966.
 29. WOLFE, V.; CORNELL, R.; PALMER, C. Acoustic correlates of pathologic voice types. *J. Speech. Hear. Res.*, 34:509-16, 1991.
 30. YUMOTO, E.; GOULD, W.J. Harmonics-to-noise as an index of the degree of hoarseness. *J. Acoust. Soc. Am.*, 71:1544-50, 1982.