

Marina Lang Fouquet¹
Mara Behlau²
Antônio José Gonçalves¹

Uma nova proposta de avaliação do segmento faringoesofágico e sua relação com a espectrografia acústica na voz traqueoesofágica

A new proposal for evaluation of the pharyngoesophageal segment and its relation with the narrow-band spectrogram in tracheoesophageal speakers

Descritores

Voz
Fístula traqueoesofágica
Espectrografia
Acústica
Laringectomia/reabilitação
Fluoroscopia/métodos

Keywords

Voice
Tracheoesophageal fistula
Spectrography
Acoustics
Laryngectomy/rehabilitation
Fluoroscopy/methods

RESUMO

Objetivo: Avaliar a relação entre a configuração do segmento faringoesofágico (SFE) e a espectrografia acústica nas vozes traqueoesofágicas (TE). **Métodos:** Participaram 30 sujeitos laringectomizados totais com fonação TE, que se submeteram à gravação da vogal /a/, para análise espectrográfica acústica, e ao exame videofluoroscópico (VF), durante a deglutição e a fonação. A análise espectrográfica foi feita pela avaliação visual do tipo do traçado e classificado como Tipo I, II, III ou IV. No exame, foi avaliado o grau de contato da proeminência do SFE com sua parede anterior, definido como ausente/leve (hipocontato), moderado (contato normal) e intenso (hipercontato); e mensurados: distância anteroposterior (DAP), comprimento do SFE (cSFE); área do reservatório de ar (área AR); área do SFE na deglutição (área DE) e na fonação (área FO). **Resultados:** O tipo de sinal espectrográfico I-II caracteriza-se por contato normal; Tipo III, hipercontato; e Tipo IV, hipocontato. O grupo com sinal Tipo I-II apresenta DAP maior e cSFE menor do que o Tipo IV. Há correlação entre DAP maior e cSFE menor. **Conclusão:** Os sinais espectrográficos Tipo I e II caracterizam o SFE com contato normal; Tipo III, com hipercontato e Tipo IV, com hipocontato. SFE com contato normal, curtos e com DAP maior caracterizam as melhores vozes TE. Somente o grau de contato da proeminência do SFE com sua parede anterior diferencia SFE com contato normal de segmentos com hipercontato.

ABSTRACT

Purpose: The primary purpose of this study was to assess the relationship between pharyngoesophageal segment (PES) configuration and narrow-band spectrogram of tracheoesophageal voices. **Methods:** This study included 30 total laryngectomees tracheoesophageal speakers. Patients were assessed by videofluoroscopy (VF), during deglutition and voicing, and the vowel /a/ was recorded for spectrographic analysis. The evaluation of VF recording consisted of visual perceptual rating of degree of contact between the prominence of PES and its anterior wall, defined as absent/mild (hypo contact), moderate (normo contact) and intense (hyper contact); and quantitative measures of PES: anteroposterior distance (APD) and length of the PES (lenPES); PES surface area in swallowing (areaSw), and phonation (areaPh), and the area of the air reservoir (areaAir). Visual inspection of a narrow-band spectrogram was made and four different acoustic signal typing were defined as Type I, II, III or IV. **Results:** Type I-II is correlated with moderate contact; Type III, with intense and Type IV, with absent/mild contact. Type I-II has bigger APD and PES with lower length than Type IV. There is a correlation between bigger APD and shorter PES. **Conclusion:** The group with I-II signal typing has PES with normo contact; Type III with hyper contact and Type IV has PES with hypo contact. The best tracheoesophageal voices are achieved by PES with moderate contact of the prominence and with shorter and larger anteroposterior PES distances. What differentiates the PES with hyper contact from PES with normal one is only the degree of contact between the prominence of the PES and its anterior wall.

Endereço para correspondência:
Marina Lang Fouquet
R. Ministro Alvaro de Souza Lima, 253,
apto 502/9, São Paulo (SP), Brasil,
CEP: 04664-020.
E-mail: marinafouquet@uol.com.br

Recebido em: 06/03/2012

Aceito em: 18/02/2013

Trabalho realizado na Faculdade de Ciências Médicas da Santa Casa de São Paulo – São Paulo (SP), Brasil.
(1) Disciplina de Cirurgia de Cabeça e Pescoço, Hospital Irmandade Santa Casa de Misericórdia de São Paulo – São Paulo (SP), Brasil.

(2) Centro de Estudos da Voz – CEV – São Paulo (SP), Brasil.

Conflito de interesses: nada a declarar.

INTRODUÇÃO

Na reabilitação da comunicação oral do paciente laringectomizado total, a voz traqueoesofágica (TE) é considerada uma opção superior à voz esofágica^(1,2), principalmente pela alta porcentagem de sucesso que é adquirida com essa forma de reabilitação, que chega a 88%^(3,4). Essa voz é possibilitada pela confecção de uma fístula TE, na qual é inserida uma prótese de silicone com válvula unidirecional⁽⁵⁾. O ar pulmonar é direcionado para o esôfago, através da prótese vocal, e vibra a fonte sonora, denominada segmento faringoesofágico (SFE). Por possibilitar o uso do fluxo de ar pulmonar, a voz TE apresenta fluência e velocidade de fala mais próximas da voz laríngea^(6,7). Entretanto, a vibração bastante irregular do SFE gera uma qualidade vocal predominantemente rugosa, com *pitch* grave^(8,9) e desviada em relação à voz laríngea^(10,11).

O SFE, por ser a fonte sonora da voz TE, foi o objeto de estudo de diversos trabalhos científicos, que utilizaram a videofluoroscopia para auxiliar na compreensão de sua fisiologia. Esse segmento localiza-se na região de transição entre a hipofaringe e o esôfago, no nível das vértebras cervicais C5 e C6 e é constituído por fibras do músculo constritor inferior da faringe, músculo cricofaríngeo e fibras estriadas do esôfago proximal⁽¹²⁾. Foi observado o crescimento de uma mucosa na parede posterior da hipofaringe, que se projeta ativamente em seu lúmen, entrando em contato com a parede anterior da mesma, vibra com a passagem do ar e, por consequência, participa da fonte sonora. Essa mucosa foi relacionada positivamente com o sucesso nas vozes esofágica e TE e foi denominada como “barra cricofaríngea”, “barra neoglótica”, “proeminência da neoglote” ou “proeminência do SFE”^(8,9,13-16).

Na tentativa de caracterizar de forma mais objetiva o SFE ideal para a voz esofágica e TE, os pesquisadores passaram a medi-lo em imagens geradas por videofluoroscopia, avaliando distância anteroposterior do SFE e do reservatório de ar, comprimento do SFE e área de sua superfície^(8,9,13,15-18). As melhores vozes TE apresentaram menor distância entre a proeminência do SFE e a parede anterior⁽¹⁵⁻¹⁷⁾; entretanto, essa medida não reflete a extensão do contato do tecido nessa área⁽¹⁸⁾, e essa extensão de contato, sendo maior ou menor, pode produzir uma voz com qualidade vocal mais tensa ou sopro, respectivamente.

Esse grau de contato entre a proeminência do SFE e sua parede anterior pode refletir o que a literatura mundial tem descrito como as diferenças de tônus do SFE, que variam desde hipotonia, na qual há pouca aproximação das paredes e a voz é fraca e sopro, passando por normotonia, hipertonia e culminando com espasmo, caracterizado por constrição intensa do SFE e ausência de emissão^(5,19,20). Vários estudos relacionaram o insucesso das vozes esofágica e TE com hipertonia e, principalmente, com o espasmo do SFE^(4,9,21,22).

Há uma limitação na avaliação acústica das vozes TE, pois estas são bastante aperiódicas e apresentam frequência muito grave, nem sempre permitindo extração confiável da frequência fundamental e demais parâmetros objetivos⁽²³⁻²⁵⁾, o que limita o uso dos parâmetros objetivos na avaliação acústica nessa população. Nesses casos, as espectrografias de faixa estreita

são muito interessantes, pois possibilitam a avaliação visual de toda a gama de vozes TE, independentemente da qualidade da emissão, permitindo visualizar a presença de componentes harmônicos e de ruídos. As espectrografias acústicas das vozes TE são classificadas em sinal Tipo I, II, III e IV, de acordo com a presença e continuidade, ou não, dos harmônicos durante toda a emissão, presença ou não de ruído espectral, estabilidade da frequência e amplitude do sinal^(23,26).

Com o intuito de compreender melhor a produção da voz TE, o presente trabalho teve como objetivo avaliar a configuração do SFE, por meio da avaliação do grau de contato da proeminência do SFE com sua parede anterior e medidas objetivas, e relacioná-la com o tipo de sinal da voz TE na espectrografia acústica.

MÉTODOS

Trata-se de um estudo transversal com 30 pacientes laringectomizados totais com prótese vocal TE, 23 homens e sete mulheres. A média de idade do grupo foi de 61,3 anos (37 a 79 anos) e o tempo de pós-operatório variou de 12 a 179 meses, com média de 59 meses. Foi realizada laringectomia total em 25 sujeitos, laringectomia total com faringectomia parcial em três e faringolaringectomia total com reconstrução com retalho em dois. Doze sujeitos foram submetidos à miotomia do SFE, dois à aplicação de botox no SFE e em dois não houve intervenção no tônus do SFE. Nos 14 pacientes restantes não havia dados no prontuário. O esvaziamento cervical foi seletivo em 17, radical em 11 e não havia dados no prontuário de dois pacientes. Foram submetidos à radioterapia 26 sujeitos e à quimioterapia, seis. A fístula traqueoesofágica foi primária em 13 pacientes e secundária em 17; o tempo de uso da prótese vocal variou entre dois e 153 meses (média de 44 meses). Todos estavam com a prótese vocal locada corretamente e utilizaram oclusão digital do traqueostoma para se comunicar com a voz TE. Foram incluídos no estudo pacientes laringectomizados totais com prótese vocal, independentemente da extensão cirúrgica ou reconstrução com retalho, manipulação do SFE (miotomia ou aplicação de toxina botulínica), sexo, idade, tratamento adjuvante, tempo de cirurgia, momento de colocação de prótese e qualidade vocal resultante.

O trabalho foi aprovado pelo Comitê de Ética em Pesquisa em Seres Humanos da instituição (Projeto 302/09). Todos os pacientes que aceitaram participar deste estudo assinaram o Termo de Consentimento Livre e Esclarecido.

Para cumprir o objetivo deste estudo, foi realizada a avaliação da qualidade da emissão da vogal /a/, por meio espectrográfica acústica, e a avaliação do SFE, por meio de análise visual e objetiva da videofluoroscopia, durante a deglutição e fonação.

Avaliação do segmento faringoesofágico

Trinta sujeitos realizaram o exame videofluoroscópico (VF), no aparelho *Phillips Optimus CDUO diagnostic-pele*, na visão lateral (perfil direito) da cabeça e pescoço, com uma moeda de 10 centavos colada na ponta de uma régua rígida de 30 cm,

presa a um pedestal, formando um ângulo de 90°. A moeda foi posicionada anteriormente ao pescoço do sujeito, em sua linha mediana, e serviu como parâmetro para as medições.

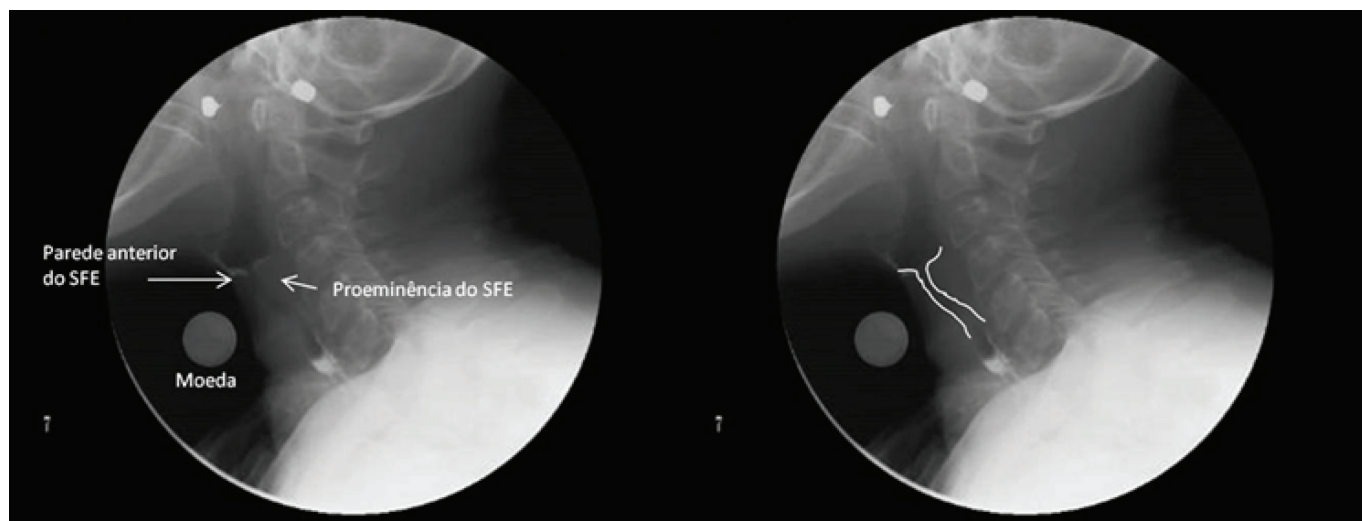
Primeiramente, os sujeitos foram instruídos a deglutir duas porções de alimento pastoso misturado ao contraste baritado, para melhor visualização das estruturas avaliadas. A seguir, os sujeitos foram solicitados a emitir cinco vogais sustentadas /a/. Utilizou-se radioescopia direta e imagem seriada na fonação. Na radioescopia direta, a imagem é constante, a dose é menor e a definição, maior. Na imagem seriada, com três imagens por segundo, o contraste é maior, o que permite observar com maiores detalhes onde termina uma estrutura e inicia a outra, facilitando a observação da movimentação do SFE durante a fonação. Os exames foram editados com o programa computadorizado *Windows Live Movie Maker*. As imagens para posterior medição foram selecionadas utilizando-se o aplicativo *Snapshot*. Para medir

comprimentos e áreas, foi utilizado o programa computadorizado *ImageJ 1.44p*⁽²⁷⁾.

Os exames de videofluoroscopia foram avaliados qualitativamente, por meio de análise visual, e objetivamente, por extração de diversas medidas.

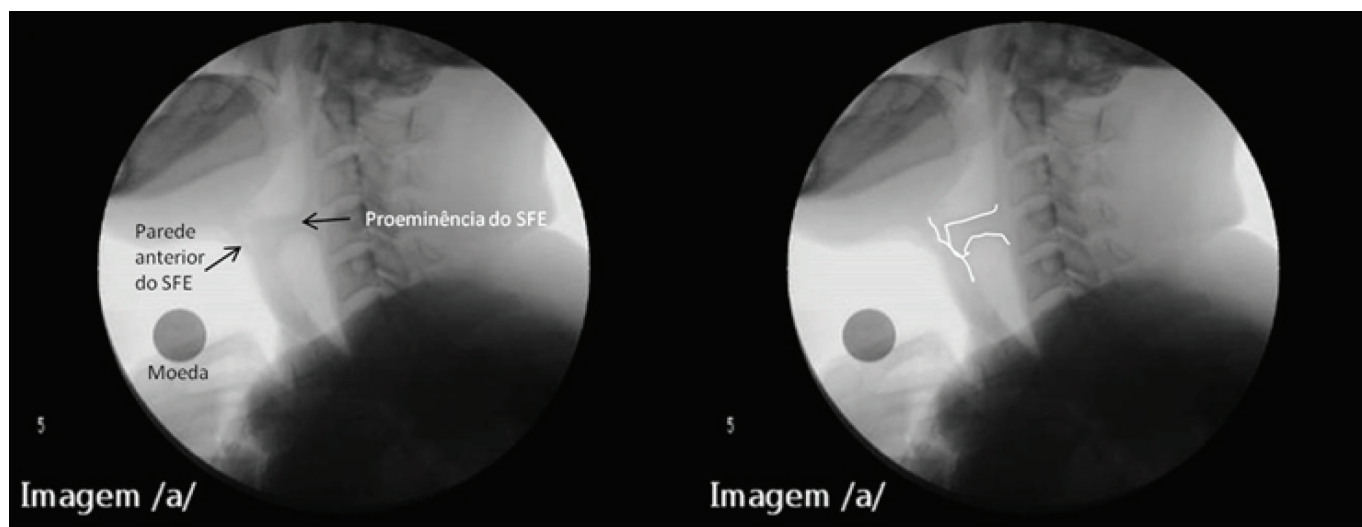
Avaliação qualitativa do segmento faringoesofágico

Na avaliação das imagens videofluoroscópicas da fonação, foi observado o grau de contato da proeminência do SFE com sua parede anterior, assinalando como contato ausente ou leve, moderado e intenso. Definiu-se como ausente ou leve, quando havia um contato de até aproximadamente $\frac{1}{4}$ da proeminência do SFE com sua parede anterior (Figura 1), moderado, quando havia contato de aproximadamente $\frac{1}{2}$ da proeminência (Figura 2), e contato intenso, quando havia contato maior que $\frac{3}{4}$ da proeminência com sua parede anterior (Figura 3).



Legenda: SFE = segmento faringoesofágico

Figura 1. Imagens videofluoroscópicas com exemplo de grau de contato ausente/leve da proeminência do segmento faringoesofágico com a parede anterior. Na imagem à direita, a região está demarcada com linhas



Legenda: SFE = segmento faringoesofágico

Figura 2. Imagens videofluoroscópicas com exemplo de grau de contato moderado da proeminência do segmento faringoesofágico com a parede anterior. Na imagem à direita, a região está demarcada com linhas

Avaliação objetiva do segmento faringoesofágico

Para a avaliação objetiva, selecionamos alguns parâmetros baseados em estudos anteriores^(9,15,17,18). Utilizamos como medida de referência a medida da moeda posicionada na linha média do paciente, com o intuito de reduzir o erro do fator de magnificação⁽¹⁷⁾. Foram medidas: distância anteroposterior do SFE durante fonação; comprimento do SFE durante fonação, distância mínima entre a proeminência do SFE e sua parede anterior durante a fonação; área da projeção do SFE durante a deglutição; área da projeção do SFE durante a fonação e acrescentamos a medida de área da projeção do reservatório de ar, durante fonação, do nível da prótese vocal até imediatamente abaixo do SFE.

A avaliação qualitativa e a objetiva do SFE foram realizadas pela primeira pesquisadora do presente estudo, sem que esta conhecesse a identidade dos pacientes, pois cada exame foi identificado por um número. Com o intuito de verificar a consistência das respostas, 20% das imagens foram reavaliadas. Na avaliação qualitativa, o grau de contato da proeminência do SFE com sua parede anterior apresentou concordância de 83,3% (valor Kappa 0,571). Para a verificação da confiabilidade intra-avaliador da avaliação objetiva do SFE, foi calculado o coeficiente de correlação intraclass. Todos os parâmetros foram avaliados com valores acima de 0,892.

Espectrografia acústica

Foram gravadas as vozes de 29 pacientes (1 paciente não conseguiu emitir voz TE) em sala acústica. Foi usado computador da *Helwett Packard* com o programa *Sound Forge 4.5c*, janela de 44.000 Hz e microfone *Headset Plantronics - Sound Innovation*, modelo *Audio 320*, posicionado lateralmente e a dois centímetros da boca do paciente. O volume foi ajustado para cada paciente utilizando-se o calibrador do *Sound Forge 4.5c*, para que a entrada do som estivesse entre -18 e -6 dB.

O material de fala era composto por três vogais /a/ sustentadas. Solicitou-se que os pacientes falassem de modo o mais natural possível.

Para a avaliação espectrográfica acústica, as vogais que haviam sido gravadas foram transferidas para o programa *Fono View*. Foram preparadas pranchas de análise com a seleção dos 2 s centrais e mais estáveis da melhor emissão da vogal /a/ de cada paciente⁽²³⁾. As pranchas apresentavam a espectrografia acústica de faixa estreita, com filtro de 40 Hz, janela de apresentação do traçado em 7,3 kHz e imagem em gradação de cinza. As imagens foram ordenadas aleatoriamente em apresentação *Power Point*, com informação apenas do gênero e idade do paciente, para avaliação de quatro fonoaudiólogas, especialistas em voz.

As fonoaudiólogas foram instruídas a observar o traçado espectroográfico e selecionar por tipo de traçado⁽²³⁾:

Tipo I – Estável e harmônico:

- Sinal estável por 2 s completos; e,
- Presença definida de harmônicos até pelo menos 1000 Hz.

Tipo II – Estável e com pelo menos um harmônico:

- Sinal estável por 2 s completos; e,
- Presença de pelo menos um harmônico na frequência fundamental por 2 s completos.

Tipo III – Instável e parcialmente harmônico:

- Sinal instável com harmônicos por 2 s completos; ou,
- Ausência de harmônicos por menos de 1 s.

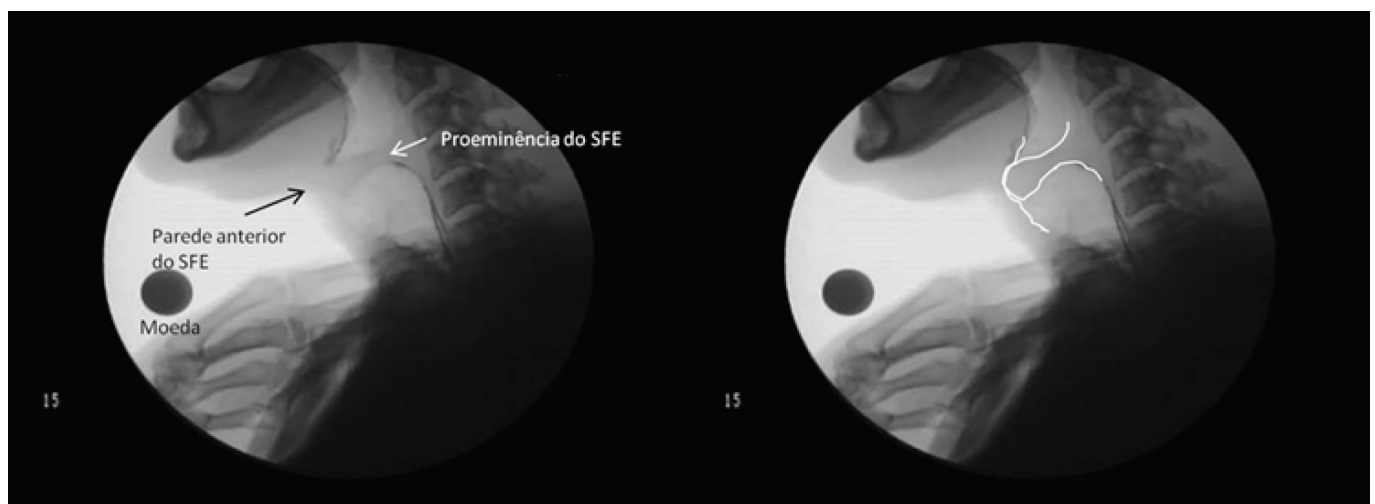
Tipo IV – Harmônicos escassos:

- Ausência completa de harmônicos; ou,
- Ausência de harmônicos por mais de 1 s.

Vinte por cento das imagens espectrográficas foram reavaliadas e apresentaram 100% de confiabilidade intra-avaliador e 93,1% de confiabilidade interavaliadores.

Análise estatística

Foi realizada a análise descritiva das variáveis avaliadas neste estudo e, para as variáveis qualitativas, foram calculadas as frequências absolutas e relativas (n e %). Já para as variáveis quantitativas, foram calculadas as medidas-resumo (média, mediana e DP) e preparados os diagramas de dispersão.



Legenda: SFE = segmento faringoesofágico

Figura 3. Imagens videofluoroscópicas com exemplo de grau de contato intenso da proeminência do segmento faringoesofágico (SFE) com a parede anterior. Na imagem à direita, a região está demarcada com linhas

As comparações das variáveis quantitativas, entre as qualitativas, foram realizadas por meio dos testes Kruskal-Wallis ou Mann-Whitney. As análises das variáveis quantitativas foram feitas por comparação umas com as outras por meio dos coeficientes de correlação (r) de Pearson ou Spearman, utilizados quando as variáveis não apresentavam distribuição normal. O nível de significância adotado foi de 5% (0,05). Foi utilizado o programa estatístico *Statistical Package for the Social Sciences* (SPSS) v. 13.

RESULTADOS

Para melhor compreensão, os resultados são apresentados na forma de análise descritiva e correlações.

Na análise descritiva, os dados de cada avaliação realizada são apresentados: para as variáveis qualitativas, mostram-se as frequências absolutas e relativas (n e %) e, para as variáveis quantitativas, as medidas-resumo de média e desvio-padrão. As correlações são mostradas como analogias das avaliações realizadas neste trabalho. Os resultados são apresentados em diagramas de dispersão, para a correlação de duas variáveis, e em tabelas, para correlações de mais de duas variáveis.

Análise descritiva

Na avaliação qualitativa do exame VF do SFE, os pacientes apresentaram contato da proeminência do SFE com a parede anterior em grau ausente ou leve em 11 (36,7%); moderado em nove (30%); e intenso em dez (33,3%).

Quanto à avaliação objetiva do exame VF do SFE de pacientes laringectomizados totais com prótese vocal, a distância anteroposterior do SFE apresentou medida média de 14,02 mm (DP=7,85 mm); o comprimento do SFE, medida média de 17,95 mm (DP=10,57 mm); e a distância mínima da proeminência do SFE com sua parede anterior, 0,65 mm (DP=1,38 mm). Nas medidas de área da projeção, o SFE na deglutição apresentou medida média de 229,78 mm² (DP=120,86 mm²); o SFE na fonação, 211,38 mm² (DP=85,58 mm²) e a área da projeção do reservatório de ar apresentou 430,79 mm² (DP=303,74 mm²). Observa-se grande variabilidade nas medidas objetivas, principalmente nas medidas de área das projeções.

Na avaliação do tipo de sinal na espectrografia acústica em pacientes laringectomizados totais com prótese vocal, o Tipo I foi observado em dois (6,9%) dos pacientes; Tipo II, em

seis (20,7%); o Tipo III, em sete (24,1%) e o Tipo IV, em 14 (48,3%). Quase a metade da amostra apresenta predominância de ruído na espectrografia acústica (Sinal Tipo IV).

Correlações

Para o estudo das correlações entre a avaliação qualitativa e objetiva do SFE e o tipo de sinal na espectrografia acústica, os pacientes com sinal Tipo I e Tipo II foram unidos em um só grupo, devido ao número pequeno de pacientes com sinal Tipo I, o que inviabiliza a análise estatística desse grupo isoladamente.

Os valores para o tipo de sinal na espectrografia acústica segundo o grau de contato da proeminência do SFE com a parede anterior (avaliação qualitativa do SFE) encontram-se na Tabela 1.

A maioria dos pacientes com o melhor sinal espectrográfico (Tipo I-II) apresenta contato moderado da proeminência do SFE com a parede anterior; os com sinal Tipo III apresentam contato intenso, e os com sinal Tipo IV apresentam contato ausente ou leve.

Os resultados para o tipo de sinal na espectrografia acústica, segundo os parâmetros da avaliação objetiva do SFE, podem ser vistos na Tabela 2.

Os pacientes com o sinal espectrográfico melhor (Tipo I-II) apresentam distância anteroposterior do SFE maior e comprimento do SFE menor.

A Figura 4 mostra a correlação entre a área da projeção do SFE na deglutição e a área da projeção do SFE na fonação, segundo o grau de contato da proeminência do SFE com sua parede anterior.

Quanto maior a área da projeção do SFE na deglutição, maior a área da projeção do SFE na fonação, no grupo com contato intenso da proeminência do SFE com a parede anterior.

Na Figura 5, o diagrama de dispersão A mostra a correlação moderada e negativa entre o comprimento do SFE e a distância anteroposterior ($r=-0,640$); o diagrama de dispersão B mostra a correlação moderada e positiva entre distância anteroposterior do SFE e a área da projeção do SFE na fonação ($r=0,514$); o diagrama C mostra a correlação moderada e negativa entre o comprimento do SFE e a área da projeção do reservatório de ar ($r=-0,624$); e o diagrama D mostra a correlação forte e positiva entre a distância anteroposterior do SFE e a área da projeção do reservatório de ar ($r=0,863$).

Tabela 1. Tipo de sinal na espectrografia acústica segundo o grau de contato da proeminência do segmento faringoesofágico com sua parede anterior

	Tipo de sinal na espectrografia acústica			Valor de p
	Tipo I-II n (%)	Tipo III n (%)	Tipo IV n (%)	
Grau de contato entre proeminência do SFE e sua parede anterior				
Ausente/leve	1(12,5)	1(14,3)	9 (64,3)	0,009*
Moderado	5 (62,5)	1(14,3)	3 (21,4)	
Intenso	2 (25,0)	5 (71,4)	2 (14,3)	

*Valor significativo (igual ou menor que 0,05)

Legenda: SFE = segmento faringoesofágico

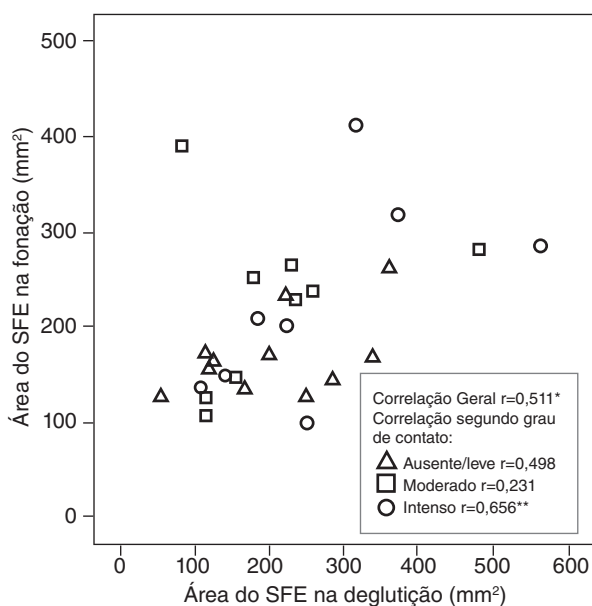
Tabela 2. Tipo de sinal na espectrografia acústica segundo os parâmetros da avaliação objetiva do segmento faringoesofágico

Parâmetros	Tipo de sinal na espectrografia acústica			Valor de p	Comparação entre dois grupos	Valor de p
	Tipo I-II	Tipo III	Tipo IV			
Avaliação objetiva do SFE	Média (DP)	Média (DP)	Média (DP)			
Distância AP do SFE ¹	19,09 (8,28)	16,35 (8,21)	10,65 (5,62)	0,034*	Tipo I-IIxIII	0,563
Comprimento do SFE ¹	12,51 (8,39)	12,90 (5,05)	21,49 (9,42)	0,044*	Tipo I-IIxIV	0,024*
					Tipo IIIxIV	0,052
					Tipo I-IIxIII	0,817
					Tipo I-IIxIV	0,024*
Tipo IIIxIV	0,073					
Área de projeção do SFE na fonação ²	211,89 (94,60)	256,06 (117,91)	186,51 (57,33)	0,464		
Área de projeção do reservatório ar ²	644,44 (370,32)	422,57 (274,57)	320,58 (193,85)	0,084		

Teste estatístico Kruskal-Wallis na comparação entre três grupos e Mann-Whitney na comparação entre dois grupos

*Valores estatisticamente significativos; 1 = medida em mm; 2 = medida em mm²

Legenda: SFE = segmento faringoesofágico; DP = desvio-padrão; AP = anteroposterior



*nível de significância 0,01; **nível de significância 0,05

Figura 4. Correlação entre as áreas da projeção do segmento faringoesofágico (SFE) na deglutição e na fonação, segundo o grau de contato da proeminência do SFE com sua parede anterior

Os resultados da Figura 5 indicam que: quanto maior a distância anteroposterior do SFE, menor o comprimento do SFE e maior a área da projeção do SFE na fonação. Quanto maior a área da projeção do reservatório de ar, menor o comprimento do SFE e maior a distância anteroposterior do SFE.

DISCUSSÃO

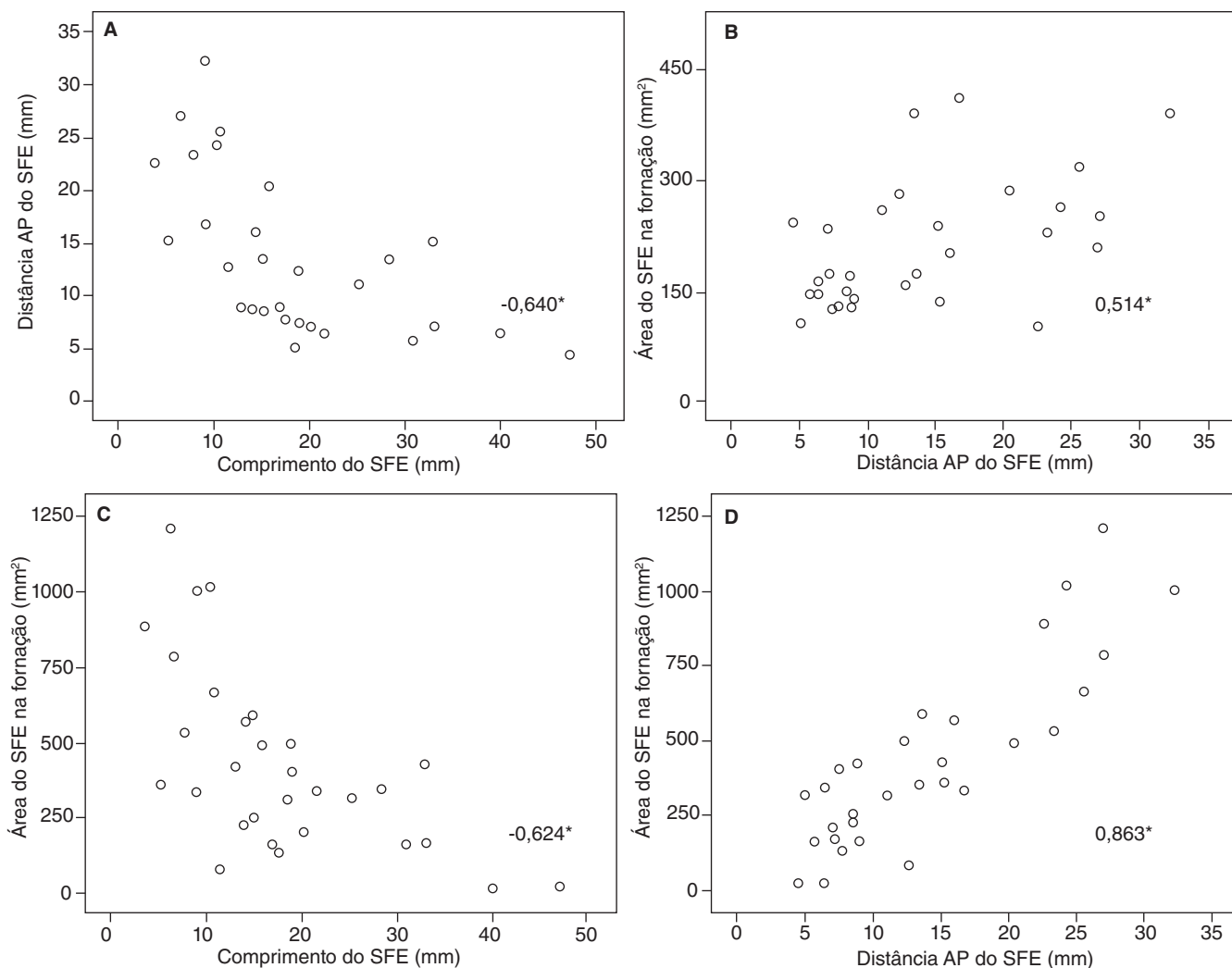
O objetivo do presente estudo foi avaliar a configuração do SFE e relacioná-la com os tipos de sinais na espectrografia acústica. Para atingir essa meta, o SFE foi avaliado por meio de exame VF. Na avaliação qualitativa, foi determinado o grau de contato da proeminência do SFE com sua parede anterior e, na avaliação objetiva, foram tomadas seis medidas: distância anteroposterior do SFE, comprimento do SFE, distância mínima

entre a proeminência do SFE e a parede anterior; áreas da projeção do SFE na deglutição, na fonação e do reservatório de ar. As espectrografias acústicas dos dois segundos centrais da melhor emissão da vogal /a/ de cada paciente foram avaliadas e categorizadas em sinal Tipo I, II, III e IV.

Foram incluídos no estudo pacientes laringectomizados totais com prótese vocal, independentemente da extensão cirúrgica ou reconstrução com retalho, manipulação do SFE (miotomia ou aplicação de toxina botulínica), sexo, idade, tratamento adjuvante, tempo de cirurgia e momento de colocação da prótese vocal. Esses dados foram avaliados, mas não houve relação significativa destes com os aspectos estudados neste trabalho.

Em geral, mede-se a distância mínima entre a proeminência do SFE e sua parede anterior^(15,17,18). Na amostra estudada, apenas seis pacientes apresentaram valores diferentes de zero, com distância máxima de 5 mm. Entretanto, durante a avaliação qualitativa da videofluoroscopia, percebeu-se que, apesar de 23 dos 29 pacientes com fonação apresentarem contato entre a proeminência do SFE e a parede anterior, havia grande variação do grau desse contato. Em estudo anterior, já foi citado que a medida de distância mínima entre a proeminência do SFE e sua parede anterior não reflete a extensão do contato do tecido nessa área⁽¹⁸⁾. A partir dessa observação, optou-se por avaliar qualitativamente o grau de contato entre a proeminência do SFE com sua parede anterior, além de medir a distância, classificando-o como contato ausente ou leve, moderado ou intenso (Figuras 1–3).

Acredita-se que possa ser feita uma analogia entre o grau de contato da proeminência do SFE com sua parede anterior e o grau do tônus do SFE, descrito amplamente na literatura como hipotônico, normotônico e hipertônico^(5,19,20). Entretanto, como não mensuramos o tônus do SFE, optamos por utilizar, neste artigo, os termos hipocontato, contato normal e hipercontato. Desta forma, o grupo com contato ausente ou leve (11 pacientes; 36,7%) apresentou SFE com hipocontato; aquele com contato moderado (nove pacientes; 30%) apresentou SFE com contato normal, e o grupo com contato intenso (10 pacientes; 33,3%), SFE com hipercontato. Do grupo com contato intenso, apenas um paciente apresentou ausência de fonação e foi classificado como tendo espasmo do SFE, segundo critério encontrado na literatura^(5,19,20). Os pacientes submetidos à miotomia foram mantidos na casuística sem a formação de um



* = nível de significância de 0,01

Legenda: SFE = segmento faringoesofágico; AP = anteroposterior

Figura 5. Correlações entre as medidas objetivas do segmento faringoesofágico: distância anteroposterior do segmento faringoesofágico, comprimento do segmento faringoesofágico, área de projeção do segmento faringoesofágico na fonação e área de projeção do reservatório de ar

grupo especial, pois foram encontrados pacientes com essas características nos três grupos de contato das paredes do SFE: hipocontato, contato normal e até no grupo com hipercontato.

Para caracterizar de forma mais objetiva o SFE e sua relação com a proficiência da voz TE, foram efetuadas algumas medidas. Foram mantidas duas das medidas realizadas em estudo anterior, como “distância anteroposterior do SFE” e “comprimento do SFE”, pois especialmente esta última apresentou correlação direta com a proficiência da voz⁽⁹⁾. Foi acrescentada a medida “distância mínima entre a proeminência do SFE e sua parede anterior”, cuja importância foi demonstrada em estudos anteriores⁽¹⁵⁻¹⁷⁾. Essa medida foi tomada somente para a fonação e não na situação de repouso, pois já foi demonstrado que existe correlação positiva entre essas medidas, sendo que bons falantes TE apresentaram valores menores dessa distância, tanto em repouso, quanto à fonação⁽¹⁵⁾. As medidas “área da projeção do SFE durante a

deglutição” e “área da projeção do SFE durante a fonação” foram avaliadas para verificar se poderia haver alguma relação entre a diferença de compressão do SFE da deglutição para a fonação e o grau de resistência oferecido pelo SFE à passagem do ar^(15,16,18). A medida “área da projeção do reservatório de ar” foi acrescentada, pois é possível que essa medida em área traduza de forma mais precisa o comportamento do ar no esôfago proximal do que a medida em linha reta da distância subneoglótica máxima⁽¹⁵⁾. Ocorreu uma grande variabilidade nas medidas objetivas dos pacientes com voz TE, principalmente nas medidas de área.

A extração de parâmetros acústicos objetivos, como frequência fundamental e índices de perturbação e ruído, nem sempre é confiável na voz TE, pois esta é bastante aperiódica e apresenta frequência muito grave^(11,23-25). Mediante essas dificuldades, as vozes foram classificadas em sinal Tipo I, II, III ou IV, por meio de avaliação visual dos traçados espectrográficos,

observando-se o traçado dos harmônicos durante a emissão, presença ou não de ruído espectral, estabilidade da frequência fundamental e amplitude do sinal⁽²³⁾. Nos resultados, observa-se que 14 pacientes (48,3%) apresentaram sinal Tipo IV, no qual ocorre ausência completa de harmônicos ou ausência de harmônicos por mais de um segundo. Portanto, quase a metade da amostra estudada apresentou predominância de ruído na espectrografia acústica, bem maior do que na população avaliada em estudo anterior, que descreve esse tipo de sinal apenas em oito dos 39 pacientes estudados (20,51%)⁽²³⁾.

Houve relação significativa ($p=0,009$) entre a espectrografia acústica e a avaliação qualitativa do SFE (Tabela 1). A maioria dos pacientes com sinal espectrográfico Tipo I-II apresentou contato moderado da proeminência do SFE com a parede anterior; os com sinal Tipo III, contato intenso; e os com sinal Tipo IV, contato ausente ou leve da proeminência do SFE com sua parede anterior. Esses achados indicam uma relação entre o grau de contato das paredes do SFE e o tipo de sinal acústico na espectrografia acústica. Portanto, pacientes com contato normal do SFE apresentam sinal Tipo I-II; hipercontato, sinal Tipo III; e hipocontato apresentam sinal Tipo IV. Um estudo anterior mostra achados comparáveis, e os autores afirmam que a hipertonicidade (hipercontato) influencia na instabilidade do sinal Tipo III, e a hipotonicidade (hipocontato) está relacionada com a falta de harmônicos do sinal Tipo IV⁽²⁶⁾.

Também foi observada relação significativa entre a espectrografia acústica e a avaliação objetiva do SFE (Tabela 2). O grupo de pacientes com sinal espectrográfico Tipo I-II apresentou comprimento do SFE menor do que o grupo de pacientes com sinal espectrográfico Tipo IV. Essa relação entre SFE mais curtos e vozes melhores, tanto esofágicas quanto TE, também foi descrita em trabalhos anteriores^(9,13). Provavelmente, SFE mais curtos vibram de forma mais regular e com um padrão mais próximo da vibração das pregas vocais, gerando um sinal com maior periodicidade do que SFE longo. Outros autores confirmam esse fato ao descreverem rugosidade maior quando os SFE são maiores⁽¹⁸⁾.

O grupo de pacientes com sinal espectrográfico Tipo I-II também apresentou distância anteroposterior do SFE maior do que o grupo de pacientes com sinal espectrográfico Tipo IV (Tabela 2). A qualidade da voz é superior quando a distância anteroposterior do SFE é maior^(8,13). Pode-se comparar essa medida com a distância subneoglótica máxima, que é medida logo abaixo do SFE; e um índice maior de aumento dessa distância, do repouso para a fonação, também foi relevante para uma melhor qualidade vocal⁽¹⁶⁾.

Ao estimar a correlação entre o grau de contato da proeminência do SFE com sua parede anterior e as áreas do SFE na deglutição e na fonação (Figura 4), observou-se que houve correlação moderada e positiva apenas no grupo com contato intenso da proeminência do SFE com sua parede anterior. Nesse grupo, quanto maior a área do SFE durante a deglutição, maior é a área do SFE durante a fonação. Possivelmente, isso ocorre por ser o SFE com hipercontato mais rígido, moldando-se menos à passagem do ar. Essa relação não foi avaliada em pesquisas anteriores.

Em relação às medidas objetivas do SFE durante a fonação, ressalta-se que foram mensuradas duas áreas específicas: 1) a

área da projeção do SFE na fonação; e 2) área da projeção do reservatório de ar. Essas áreas relacionam-se com o comprimento e distância anteroposterior do SFE, a saber:

Quanto maior a área do SFE na fonação, maior a distância anteroposterior do mesmo e, quanto maior a distância anteroposterior do SFE, menor o comprimento do SFE (Figura 5, diagramas A e B).

Quanto maior a área do reservatório de ar, maior a distância anteroposterior e menor o comprimento do SFE (Figura 5, diagramas C e D).

O melhor sinal espectrográfico (Tipo I-II) correlacionou-se com a distância anteroposterior maior e comprimento menor do SFE (Tabela 2). Pode-se afirmar, portanto, que a configuração ideal é: distância anteroposterior do SFE maior, comprimento do SFE menor e áreas da projeção – tanto do SFE à fonação, quanto do reservatório de ar – maiores.

Correlações entre as medidas objetivas do SFE também foram avaliadas anteriormente, e os autores indicam que o alargamento da área abaixo do SFE tem relação com uma medida maior da proeminência do SFE, mas não relacionam esses achados com uma melhor ou pior qualidade vocal⁽¹⁸⁾.

CONSIDERAÇÕES FINAIS

É possível distinguir um padrão de comportamento correlacionando o grau de contato do SFE com sua parede anterior. O grupo com contato moderado entre a proeminência do SFE e sua parede anterior apresenta a condição ideal, com contato normal, pois o contato não é nem leve, nem intenso demais (Figura 2). Esse grupo foi o que apresentou a maioria das espectrografias acústicas com sinal Tipo I-II (Tabela 1), caracterizado por sinal estável durante dois segundos e harmônicos visíveis até pelo menos a faixa de 1.000 Hz⁽²³⁾. Ao correlacionar a espectrografia acústica com as medidas objetivas do SFE, o grupo com sinal Tipo I-II apresentou distância anteroposterior do SFE maior e comprimento do SFE menor (Tabela 2). Além disso, constatou-se associação entre distância anteroposterior do SFE maior e comprimento do SFE menor, com área do SFE e área da projeção do reservatório de ar maiores durante a fonação (Figura 5).

Os pacientes com grau de contato intenso da proeminência do SFE com sua parede anterior podem ser classificados como pacientes com SFE com hipercontato, com exceção de um paciente, que apresentava ausência de fonação e foi classificado com espasmo do SFE. Dos nove pacientes restantes, a maioria apresentou sinal espectrográfico Tipo III (Tabela 1). As medidas do SFE e do esôfago dos pacientes com sinal espectrográfico Tipo III assemelham-se às medidas do grupo com sinal Tipo I-II, não havendo diferença estatística entre elas na distância anteroposterior e no comprimento do SFE (Tabela 2). Entretanto, quanto maior a área do SFE na deglutição, maior é a área do SFE na fonação no grupo com contato intenso (Figura 4), o que pode indicar um SFE mais rígido, que se molda menos com a passagem do ar.

Finalmente, o grupo de pacientes com contato da proeminência do SFE com sua parede anterior ausente ou leve pode ser classificado como possuidor de SFE com hipocontato, e a

maioria apresenta sinal espectrográfico do Tipo IV (Tabela 2). As medidas do SFE do grupo com sinal Tipo IV diferenciam-se do grupo com sinal Tipo I-II por uma distância anteroposterior do SFE menor, um comprimento do SFE maior (Tabela 2), área do SFE durante a fonação menor e área da projeção do reservatório de ar menor (Figura 5).

Constatou-se que as medidas do SFE distinguem claramente o grupo com contato moderado da proeminência do SFE com sua parede anterior e sinal Tipo I-II do grupo com contato ausente/leve e sinal Tipo IV, ou seja, distinguem SFE com contato normal dos SFE com hipocontato.

É possível inferir, ainda, que o contato intenso da proeminência do SFE com sua parede anterior é realmente o aspecto que diferencia o grupo com hipercontato e sinal Tipo III, pois as medidas do SFE são muito parecidas com as medidas do grupo com contato normal. Portanto, o paciente com hipercontato do SFE terá benefício com a manipulação do SFE, seja por miotomia ou aplicação de toxina botulínica^(4,28,29), para que ocorra uma suavização do contato de suas paredes durante a fonação, facilitando a emissão de voz TE mais adequada.

CONCLUSÃO

Na avaliação da configuração do SFE, por meio de videofluoroscopia, e da espectrografia acústica das vozes de 30 laringectomizados totais com prótese vocal, observou-se que: as melhores vozes TE são produzidas por SFE com contato normal de suas paredes, mais curtos e com distância anteroposterior maior, área do SFE maior e área do reservatório de ar maior durante a fonação; o tipo de traçado espectrográfico está relacionado com o grau de contato das paredes do SFE durante a fonação; e o que diferencia o paciente com SFE com hipercontato do contato normal, na configuração do SFE, é somente o grau de contato da proeminência do SFE com sua parede anterior.

**MLF foi responsável pela realização dos exames, coleta de dados, tabulação dos dados e elaboração e execução do manuscrito; MB foi responsável pelo delineamento do estudo, orientação na análise dos dados e na elaboração do manuscrito; AJG foi responsável pela orientação geral no delineamento do estudo e do manuscrito.*

REFERÊNCIAS

1. Stajner-Katusić S, Horga D, Musura M, Globlek D. Voice and speech after laryngectomy. *Clin Linguist Phon.* 2006;20(2-3):195-203.
2. Van Rossum MA, Quené H, Nootboom SG. Prosodic boundaries in alaryngeal speech. *Clin Linguist Phon.* 2008;22(3):215-31.
3. Op de Coul BMR, Hilgers FJM, Balm AJM, Tan IB, Van Tinteren H, Van den Hoogen FJA. A decade of postlaryngectomy vocal rehabilitation in 318 patients: a single institution's experience with consistent application of provox indwelling voice prostheses. *Arch Otolaryngol Head Neck Surg.* 2000;126(11):1320-8.
4. Van Weissenbruch R, Kunnen M, Van Cauwenberge PB, Albers FWJ, Sulter AM. Cineradiography of the pharyngoesophageal segment in postlaryngectomy patients. *Ann Otol Rhinol Laryngol.* 2000;109(3):311-9.
5. Singer MI, Blom ED. An endoscopic technique for restoration of voice after laryngectomy. *Ann Otol Rhinol Laryngol.* 1980;89(5):529-33.
6. Globlek D, Stajner-Katusić S, Musura M, Horga D, Liker M. Comparison of alaryngeal voice and speech. *Logoped Phoniatr Vocol.* 2004;29(2):87-91.

7. Van As-Brooks CJ, Fuller DP. Prosthetic tracheoesophageal voice restoration following total laryngectomy. In: Ward EC, Van As-Brooks CJ, editors. *Head and neck cancer: treatment, rehabilitation and outcomes.* San Diego: Plural Publishing; 2007. p. 229-66.
8. Vrticka K, Svoboda M. A clinical and X-ray study of 100 laryngectomized speakers. *Folia Phoniatr.* 1961;13:174-86.
9. Fouquet ML, Gonçalves AJ, Behlau M. Relation between videofluoroscopy of the esophagus and the quality of esophageal speech. *Folia Phoniatr Logop.* 2009;61(1):29-36.
10. Van As CJ, Hilgers FJM, Verdonck-de Leeuw IM, Koopmans-van Beinum FJ. Acoustical analysis and perceptual evaluation of tracheoesophageal prosthetic voice. *J Voice.* 1998;12(2):239-48.
11. Lundström E, Hammarberg B. Speech and voice after laryngectomy: perceptual and acoustical analyses of tracheoesophageal speech related to voice handicap index. *Folia Phoniatr Logop.* 2011;63(2):98-108.
12. Diedrich WM, Youngstrom KA. *Alaryngeal Speech.* v. III. Springfield: Charles C. Thomas; 1966.
13. Böhme G, Schneider HG. Die Pathophysiologie des Laryngektomierten in Zusammenhang mit der Güte der Sprechfunktion. *Z Laryngol Rhinol Otol.* 1960;39:512-20.
14. Nishizawa N, Mesuda Y, Kobashi M, Takahashi M, Inuyama Y. Identification of the opener and closer of the pharyngoesophagus in laryngectomees. *Auris Nasus Larynx.* 2001;28:S63-9.
15. Van As CJ, Op de Coul BMR, Van den Hoogen FJA, Koopmans-van Beinum FJ, Hilgers FJM. Quantitative videofluoroscopy: a new evaluation tool for tracheoesophageal voice production. *Arch Otolaryngol Head Neck Surg.* 2001;127(2):161-9.
16. Van As-Brooks CJ, Hilgers FJM, Koopmans-van Beinum FJ, Pols LCW. Anatomical and functional correlates of voice quality in tracheoesophageal speech. *J Voice.* 2005;19(3):360-72.
17. Kazi R, Singh A, Mullan GPJ, Venkitaraman R, Nutting CM, Clarke P, et al. Can objective parameters derived from videofluoroscopic assessment of post-laryngectomy valved speech replace current subjective measures? An e-tool-based analysis. *Clinical Otolaryngology.* 2006;31:518-24.
18. Lundström E, Hammarberg B, Munck-Wikland E, Edsberg N. The pharyngoesophageal segment in laryngectomees – videoradiographic, acoustic, and voice quality perceptual data. *Lobopedics Phoniatics Vocology.* 2008;33: 115-25.
19. Perry A, Cheesman AD, McIvor J, Chalton R. A British experience of surgical voice restoration techniques as a secondary procedure following total laryngectomy. *J Laryngol Otol.* 1987;101:155-63.
20. Singer MI, Blom ED. Selective myotomy for voice restoration after total laryngectomy. *Arch Otolaryngol.* 1981;107:670-3.
21. Gatenby RA, Rosenblum JS, Leonard CM, Moldofsky PJ, Broder GJ. Esophageal speech: double-contrast evaluation of the pharyngoesophageal segment. *Radiology.* 1985;157:127-31.
22. McIvor J, Evans PF, Perry A, Cheesman AD. Radiological assessment of post laryngectomy speech. *Clinical Radiology.* 1990;41:312-6.
23. Van As Brooks CJ, Koopmans-van Beinum FJ, Pols LCW, Hilgers FJM. Acoustic signal typing for evaluation of voice quality in tracheoesophageal speech. *J Voice.* 2006;20(3):355-68.
24. Haderlein T, Nöth E, Toy H, Batliner A, Schuster M, Eysholdt U, et al. Automatic evaluation of prosodic features of tracheoesophageal substitute voice. *Eur Arch Otorhinolaryngol.* 2007;264:1315-21.
25. Bocklet T, Toy H, Nöth E, Schuster M, Eysholdt U, Rosanowski F, et al. Automatic evaluation of tracheoesophageal substitute voice: sustained vowel versus standard text. *Folia Phoniatr Logop.* 2009;61(2):112-6.
26. Van As CJ, Koopmans-van Beinum FJ, Hilgers FJM, Pols LCW. Acoustic measures and signal typing of voice quality in tracheoesophageal speech, and their relations to perceptual evaluations. In: Van As CJ. *Tracheoesophageal speech – a multidimensional assessment of voice quality.* Netherlands: Budde Elinkwijk; 2001. p. 73-101.
27. Gonçalves MI, Leonard R. A hardware-software system for analysis of video images. *J Voice.* 1998;12(2):143-50.
28. Blom ED, Singer MI, Hamaker RC. An improved esophageal insufflation test. *Arch Otolaryngol.* 1985;111:211-2.
29. Crary MA, Glowasky AL. Using botulinum toxin A to improve speech and swallowing function following total laryngectomy. *Arch Otolaryngol Head Neck Surg.* 1996;122(7):760-3.