

Marina Lang Fouquet¹
Mara Behlau²
Antônio José Gonçalves¹

Keywords

Voice
Tracheoesophageal fistula
Spectrography
Acoustics
Laryngectomy/rehabilitation
Fluoroscopy/methods

Descritores

Voz
Fístula traqueoesofágica
Espectrografia
Acústica
Laringectomia/reabilitação
Fluoroscopia/métodos

Correspondence address:

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

Received: 03/06/2012

Accepted: 02/18/2013

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

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

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.

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.

Study carried out at the School of Medical Sciences from Santa Casa de São Paulo – São Paulo (SP), Brazil.
(1) Discipline of Head and Neck Surgery, Hospital Irmandade Santa Casa de Misericórdia de São Paulo – São Paulo (SP), Brazil.

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

Conflict of interests: nothing to declare.

INTRODUCTION

In oral communication rehabilitation of patients with total laryngectomies, the tracheoesophageal (TE) voice is considered a superior option to the esophageal voice^(1,2), primarily because of the high percentage of success achieved (88% of success) with this type of rehabilitation^(3,4). The TE voice is possible with the insertion of a silicone prosthesis, with a unidirectional valve, in a surgically made TE fistula⁽⁵⁾. Lung air is directed toward the esophagus, through the vocal prosthesis, and the sound source vibrates, which is labeled as pharyngoesophageal segment (PES). Because this makes it possible to use the flow of the lung air, the TE voice approximates the speech fluency and speed of the laryngeal voice^(6,7). In the meantime, the irregular vibration of the PES, which is evident, generates a voice quality quite rough, with low pitch^(8,9) and which is different, when compared with the laryngeal voice^(10,11).

Because the PES is the sound source of the TE voice, it was the subject of study in diverse scientific activities that used video fluoroscopy (VF) to help comprehend its physiology. The PES is located in the transition region between the hypopharynx and esophagus, at the level of the cervix vertebrae C5 and C6, and is part of the constricting muscle fibers of the inferior part of the pharynx, cricopharyngeal muscles, and the striated fibers of the proximal esophagus. A growth of a mucosa on the posterior wall of the hypopharynx that actively projects in its lumen, in contact with its anterior wall, vibrating when air passes through, and consequently, participating as a sound source, was observed⁽¹²⁾. This mucosa was positively associated with the success of the esophageal and the TE voices and was labeled "cricopharyngeal bar," "neoglottal bar," "prominence of the neoglottis," or "prominence of the PES"^(8,9,13-16).

In an attempt to characterize more objectively the ideal PES for the esophageal voice and the TE voice, the investigators measured the images generated through VF, evaluating the anteroposterior distance (APD) of the PES and the air reserves, length of the PES, and its surface area^(8,9,13,15-18). The best TE voices presented with shorter distances between the PES prominence and the anterior wall^(5,19,20); nevertheless, that measurement did not reflect the extent of contact of the tissue in that area⁽¹⁸⁾. The contact extent, either larger or smaller, can produce a more tense or breathy voice quality, respectively.

The degree of contact between the prominence of the PES and its anterior wall may reflect what the world literature has described as the tone differences of the PES. This varies from hypotonia, where there is little approximation between the walls and the voice is weak and breathy, passing as normotonia, hypertonia, and culminating with spasms, characterized by the intense constriction of the PES and absence of emissions^(5,19,20). Several studies associated the lack of success of esophageal voice and TE voice with hypertonia, primarily, to spasms in the PES^(4,9,21,22).

There is a limitation in the acoustic evaluation of TE voices, because these are quite aperiodic and present with very low frequencies, not always permitting the reliable extraction of the fundamental frequencies and other objective parameters⁽²³⁻²⁵⁾, thereby limiting the use of objective parameters in the acoustic

evaluation in this population. In such cases, narrow-band spectrography is quite interesting, because it is a visual evaluation of the whole TE voice, which is gamma independent from the emission qualities, permitting the visualization of the presence of harmonic and noise components. Acoustic spectrographs of the TE voices are classified into Type I, II, III, and IV signals, in accordance with the presence and continuity, or not, of harmonics during all of the emission, presence or not of spectral noise, and stability of the frequency and signal amplitude^(23,26).

The objective of the present study was to evaluate the PES configuration, by evaluating the degree of contact of the PES prominence with its anterior wall and objective measures, and to relate them with the TE voice acoustic spectrogram signal types, with the aim to comprehend the TE voice production better.

METHODS

This was a cross-sectional study with total of 30 laryngectomy patients with TE vocal prosthesis, including 23 male patients and seven female patients. The mean age of the group was 61.3 years (age ranging from 37 to 79 years) and the post-operation time varied between 19 and 179 months, the mean being equal to 59 months. Twenty five patients had total laryngectomies, three patients had total laryngectomies with partial pharyngectomies, and two subjects had total pharyngectomies with reconstruction flap. Twelve patients were subjected to myotonia of the PES, two patients were applied Botox to the PES, and two patients did not have muscle tone intervention of the PES. Of the remaining 14 patients, there were no data in their files. With reference to cervical lymphadenectomy, seventeen patients had selective, 11 patients had radical, and for two patients there were no data. Radiotherapy was experienced by 26 patients and chemotherapy was experienced by six patients. The TE fistula was primary in 13 patients and secondary in seventeen patients; the time of use of the vocal prosthesis varied between 2 and 153 months (average 44 months). All the patients were with their vocal prosthesis appropriately placed and used digital occlusion of the tracheostomia to communicate with the TE voice. The patients with total laryngectomies with vocal prosthesis were included in this study, independent from the extent of their surgery or reconstruction with flap, PES manipulation (myotomy or application of toxina botulinica), gender, age, assistive treatment, time from surgery, moment of prosthesis application, and quality of residual voice.

The investigation was approved by the Ethics Committee of Human Beings of the Institution (Project 302/09). All the patients accepted to participate in this study by signing the Terms of Consent and Clarification.

To carry out the objective of this study, the quality of emission of the vowel /a/ was evaluated using acoustic spectrography, and the evaluation of the PES was performed through visual and objective analysis of VF while swallowing and phonation.

Evaluation of the pharyngoesophageal segment

Thirty patients were part of a VF, using a Phillips Optimus CDUO diagnostic-pele, of the lateral view (right profile) of

the head and neck, with a ten cent coin located on the end of a rigid 30 cm ruler, held to a base, forming a 90° angle. The coin was positioned in front of the patient's neck, in the midline, and served as a parameter for the measures.

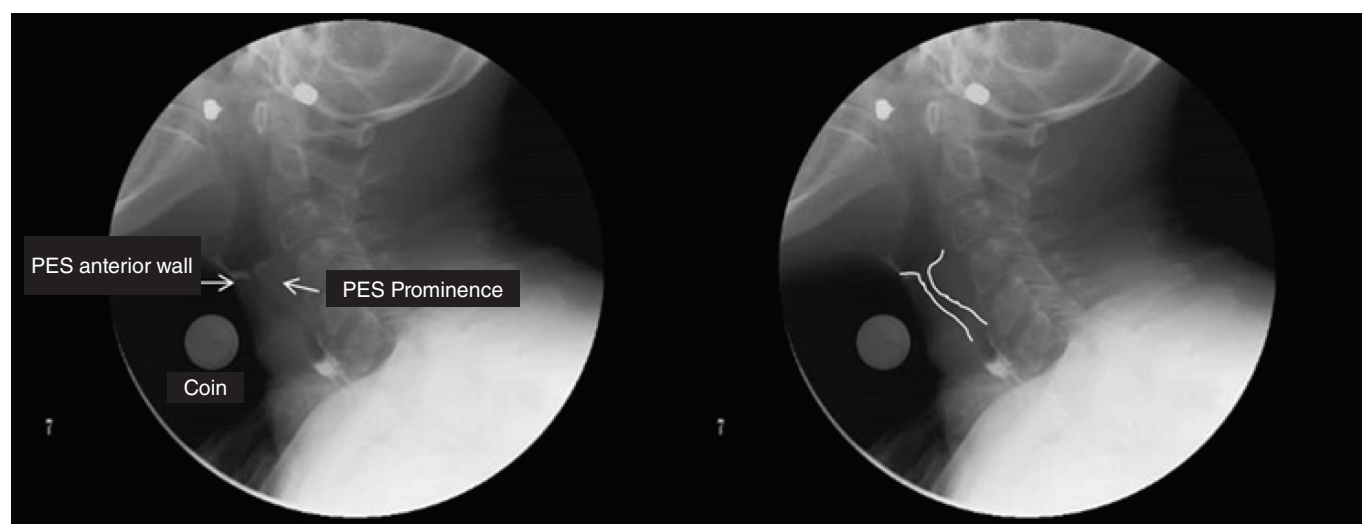
First, the patients were instructed to swallow two doughy food portions, mixed with barium, to better visualize the structure being evaluated. Following this, the patients were asked to emit and sustain five /a/ vowels. Direct radioscropy and serial images of the phonation were used. In the direct radioscropy, the image is constant, the dosage is less, and the definition is better. In the serial image, with three images per second, the contrast is better, which permits the observation of greater details, where one structure begins and the next begins, facilitating the observations of the PES movements during phonation. The tests were edited with the software Windows Live Movie Maker. The images for later measurements were selected using the Snapshot application. The

computer program ImageJ 1.44 was used to measure the lengths and areas of the images⁽²⁷⁾.

The VF tests were qualitatively evaluated through visual analysis and were objectively evaluated through the use of various measures.

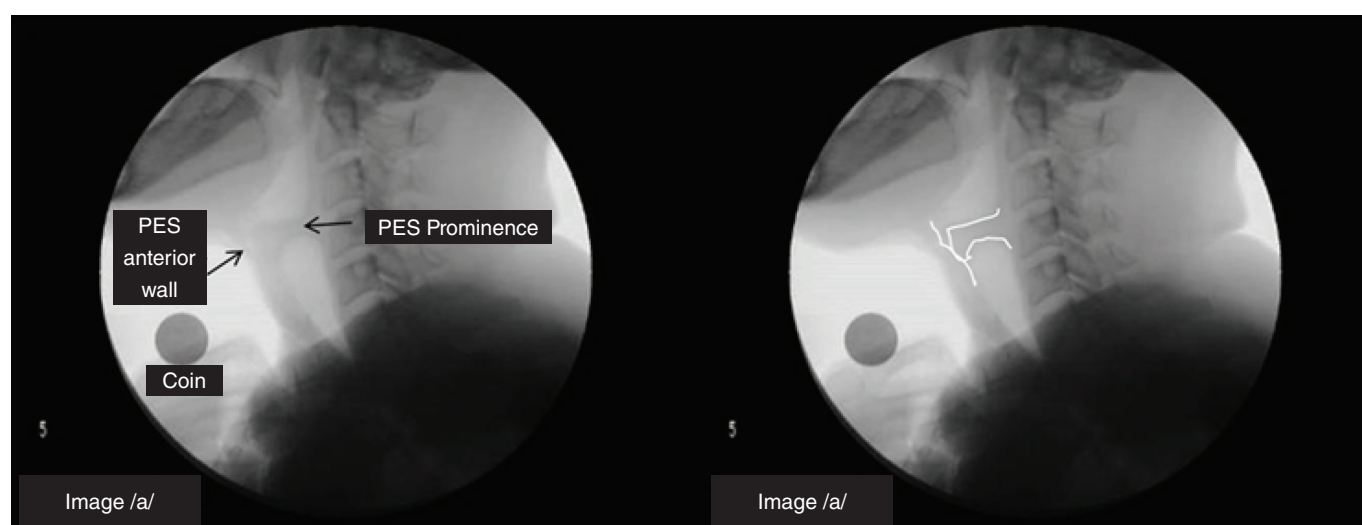
Qualitative evaluation of the pharyngoesophageal segment

In the evaluation of the phonation VF images, the degree of contact of the prominence of the PES with its anterior wall, distinguishing between absent or mild, and moderate or intense contact was observed. Absence or mild contact was defined when there was contact up to approximately $\frac{1}{4}$ of the prominence of the PES with its anterior wall (Figure 1), moderate contact was defined when there was contact approximately $\frac{1}{2}$ of the prominence (Figure 2), and intense contact was defined



Caption: PES = pharyngoesophageal segment

Figure 1. Videofluoroscopic images with example of contact degree absent / mild prominence of the pharyngoesophageal segment to the anterior wall. In the right image, the region is marked with lines



Caption: PES = pharyngoesophageal segment

Figure 2. Videofluoroscopic images with example of contact degree absent / mild prominence of the pharyngoesophageal segment to the anterior wall. In the right image, the region is marked with lines

when there was contact beyond $\frac{3}{4}$ of the prominence from its anterior wall (Figure 3).

Objective evaluation of the pharyngoesophageal segment

For the objective evaluation, several parameters were selected based on the previous studies^(9,15,17,18). As a measure of reference, we used the measurement of the coin position in the midline of the patient with the intent to reduce magnification errors⁽¹⁷⁾. The following were measured: PES APD during phonation; PES length during phonation; minimum distance between PES prominence and its anterior wall during phonation; PES projection area during swallowing; PES projection area during phonation, and we increased the measure of the projection of the air reserve, during phonation, of the vocal prosthesis until immediately below of the PES.

The qualitative and objective evaluations of the PES were carried out by the first investigator of the present study, without her being able to know or identify the patients, because each exam was identified by number. With the intent to verify the consistency of the responses, 20% of the images were re-evaluated. In the qualitative evaluation, the degree of contact of the PES prominence with its anterior was presented with an 83.3% agreement (Kappa 0.571). To verify the intra-rater reliability of the PES objective evaluations, an interclass correlation coefficient was calculated. All the parameters resulted in values above 0.892.

Acoustic spectrography

The voices of 29 patients were recorded (one patient was unable to emit TE voice) in an acoustic chamber. The Hewlett Packard computer was used with the program Sound Forge 4.5c, window of 44.000 Hz, and microphone Headset Plantronics – Sound Innovation, model Audio 320, positioned laterally and at two centimeters from the mouth of the patient. The volume was adjusted for each patient using a Sound Forge

4.5c calibrator, such that the sound, which entered stayed between -18 and -6 dB.

The speech material comprised three sustained /a/ vowels. The patients were asked to talk as natural as possible.

For the acoustic spectrograph analysis, the vowels that had been recorded were transferred into the program Fono View. Plates for analysis were prepared by selecting the two 2 s central and most stable vowel /a/ emissions for each patient⁽²³⁾. The plates presented narrow band acoustic spectrography, with 40 Hz filter, window of trace presentation at 7.3 kHz, and gray-scaled image. The images were presented in random order in Power Point, with only the gender and age of the patient, for the evaluation by four phonoaudiologists specialized in voice.

The phonoaudiologists were instructed to observe the spectrographic trace and select the type of trace as follows⁽²³⁾:

Type I – Stable and harmonic:

- Stable signal for complete 2 s; and
- Defined presence of harmonics up to at least 1000 Hz.

Type II – Stable with at least one harmonic:

- Stable signal for complete 2 s; and
- Presence of at least one harmonic in the fundamental frequency for 2 s complete.

Type III – Stable and partially harmonic:

- Stable signal for complete 2 s; or,
- Absence of harmonics for less than 1 s.

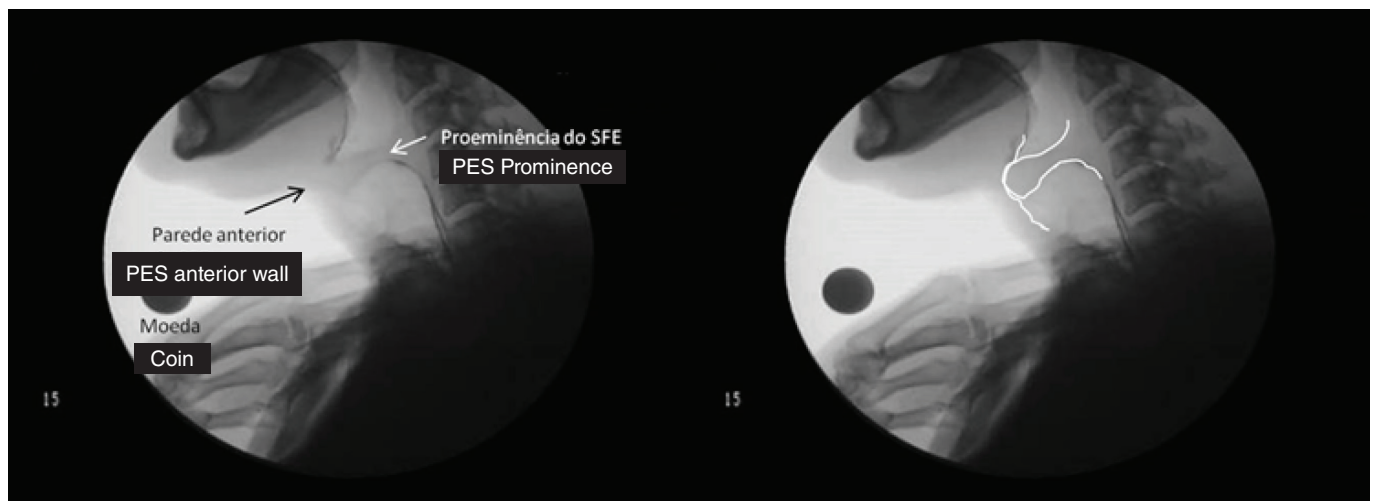
Type IV – carce harmonics:

- Complete absence of harmonics; or
- Absence of harmonics for more than 1 s.

Twenty percent of the spectrographic images were re-evaluated and presented with 100% intra-rater reliability and 93.1% inter-rater reliability.

Statistical analysis

The analysis of the described variables was carried out in this study and for the qualitative variables, absolute and



Caption: PES = pharyngoesophageal segment

Figure 3. Videofluoroscopic images with example of degree of prominence of intense contact of the pharyngoesophageal segment to the anterior wall. In the right image, the region is marked with lines

relative (i.e., percentages) frequencies, were calculated. For the quantitative variables, descriptive calculations (mean, median, and standard deviation [SD]) were used and scattergrams were prepared.

The comparisons of the quantitative variables, among the qualitative variables, were carried out using the Kruskal-Wallis or Mann-Whitney tests. The analysis of the quantitative variables were made using Pearson (r) or Spearman correlation coefficients, which served to compare the variables against each other. The degree of significance used was 5% (0.05). The Statistical Package for the Social Sciences (SPSS) v. 13 was used.

RESULTS

For better comprehension, the results are presented in the form of descriptive and correlational statistics.

In the descriptive analysis, the data for each evaluation carried out are presented as follows: for the qualitative variables, the absolute and relative frequencies (i.e., percentages) are presented and for the quantitative variables, means and standard deviation descriptive statistics, are presented. Correlations are presented as analogies of the evaluations carried out in this study. The results are presented in scattergrams for two-variable correlations, and in tables for correlations of more than two variables.

Descriptive analysis

In the qualitative evaluation of the PES VF, 11 (36.7%) patients presented prominence contact of the PES with the anterior wall in absent or mild degrees; nine (30%) patients presented prominence contact with moderate degree; and ten (33.3%) patients presented prominence contact with intense degree.

With regard to the objective evaluation of the PES VF of patients with total laryngectomies with vocal prosthesis, the APD of the PES resulted in a mean of 14.02 mm (SD=7.85 mm); the average PES length resulted in 17.95 mm (SD=10.57 mm); and the minimum distance of the PES prominence with its anterior wall was found to be 0.65 mm (SD=1.38 mm). The measurements with respect to the projection area, the PES during swallowing presented an average value of 229.78 mm² (SD=120.86 mm²), PES during phonation was 211.38 mm²

(SD=85.58 mm²), and the projection area of the air reservoir presented was 430.79 mm² (SD=303.74 mm²). Much variability in the objective measures was observed, primarily in the projection area measures.

In the signal type evaluation using the acoustic spectrography with patients who had total laryngectomies with vocal prosthesis, Type I signal was observed in two (6.9%) of the patients, Type II signal was observed in six (20.7%) patients, Type III signal was observed in seven (24.1%) patients, and Type IV signal was observed in 14 (48.3%) patients. Almost half of the sample presented with noise preponderance in the acoustic spectrography (Signal Type IV).

Correlations

For the study of correlations between the qualitative and objective evaluations of the PES and the type of signal in the acoustic spectrography, the patients with signal types I and II were merged into one group, because of the low counts of patients with Type I, which would have made the statistical analysis of that group in isolation impossible.

The values for the signal type in the acoustic spectrography with regard to the degree of contact of PES prominence with the anterior wall (qualitative evaluation of the PES) are found in Table 1.

The majority of the patients with the best spectrographic signal (Types I – II) present with moderate contact of the PES prominence with the anterior wall, those with signal Type III present with intense contact, and those with signal Type IV present with an absence or mild contact.

The results for the signal type in the acoustic spectrography, with regard to the objective evaluation parameters of the PES may be viewed in Table 2.

The patients with the best spectrographic signal (Types I – II) present with a larger PES APD and shorter PES length.

Figure 4 presents the correlation between the PES projection area during swallowing and the PES projection area during phonation, in relation to the degree of contact of the PES prominence with its anterior wall.

In the group with intense contact of the PES prominence with the anterior wall, it was found that the larger the PES projection area during swallowing, the larger the PES projection area during phonation.

Table 1. Type of signal in acoustic spectrography according to the degree of contact with the prominence of pharyngoesophageal segment with its anterior wall

	Type of signal in acoustic spectrography			p-value
	Type I-II	Type III	Type IV	
	n (%)	n (%)	n (%)	
Degree of contact between prominence of PES and its interior wall				
Non-existent/slight	1 (12.5)	1 (14.3)	9 (64.3)	0.009*
Moderate	5 (62.5)	1 (14.3)	3 (21.4)	
Intense	2 (25.0)	5 (71.4)	2 (14.3)	

* Significant value (≤ 0.05)

Caption: PES = pharyngoesophageal segment

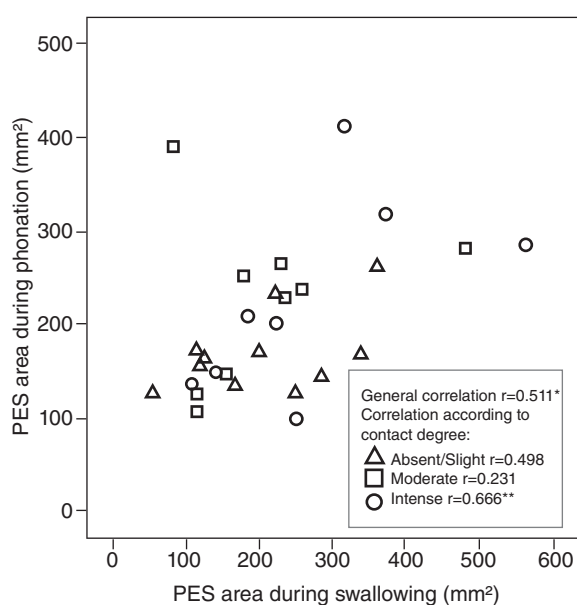
Table 2. Type of signal in acoustic spectrography according to the parameters of the objective evaluation of the pharyngoesophageal segment

Parameters	Type of signal in acoustic spectrography			p-value	Comparison	
	Type I-II Mean SD	Type III Mean SD	Type IV Mean SD		between two groups	p-value
PES objective evaluation						
AP distance from PES1	19.09 (8.28)	16.35 (8.21)	10.65 (5.62)	0.034*	Type I-II × III	0.563
					Type I-II × IV	0.024*
					Type III × IV	0.052
PES1 length	12.51 (8.39)	12.90 (5.05)	21.49 (9.42)	0.044*	Type I-II × III	0.817
					Type I-II × IV	0.024*
					Type III × IV	0.073
Projection area of the PES in phonation ²	211.89 (94.60)	256.06 (117.91)	186.51 (57.33)	0.464		
Projection area of the reservoir ar ²	644.44 (370.32)	422.57 (274.57)	320.58 (193.85)	0.084		

Kruskal-Wallis statistical test when comparing three groups and Mann-Whitney test when comparing two groups.

*Statistically significant values; 1 = measured in mm; 2 = measured in mm²

Caption: PES = pharyngoesophageal segment; SD = standard deviation; AP = anteroposterior



* 0.01 level of significance; ** 0.05 significance level.

Figure 4. Correlation between the projection of the pharyngoesophageal segment (PES) in swallowing and phonation, according to the degree of contact with the prominence of the PES and its anterior wall

In Figure 5, scattergram A illustrates a moderate and negative correlation between the PES length and the APD ($r=-0.640$); scattergram B illustrates a moderate and positive correlation between PES APD and the SPE projection area during phonation ($r=0.514$); scattergram C illustrates a moderate and negative correlation between the PES length and the projection area of the air reservoir ($r=-0.624$); and scattergram D illustrates a strong and positive correlation between the PES APD and the projection area of the air reservoir ($r=0.863$).

The results in Figure 5 indicate that the longer the PES APD, the shorter the PES length and the larger the PES projection area during phonation. The larger the projection area of the air reservoir, the shorter is the PES length and the longer is the PES APD.

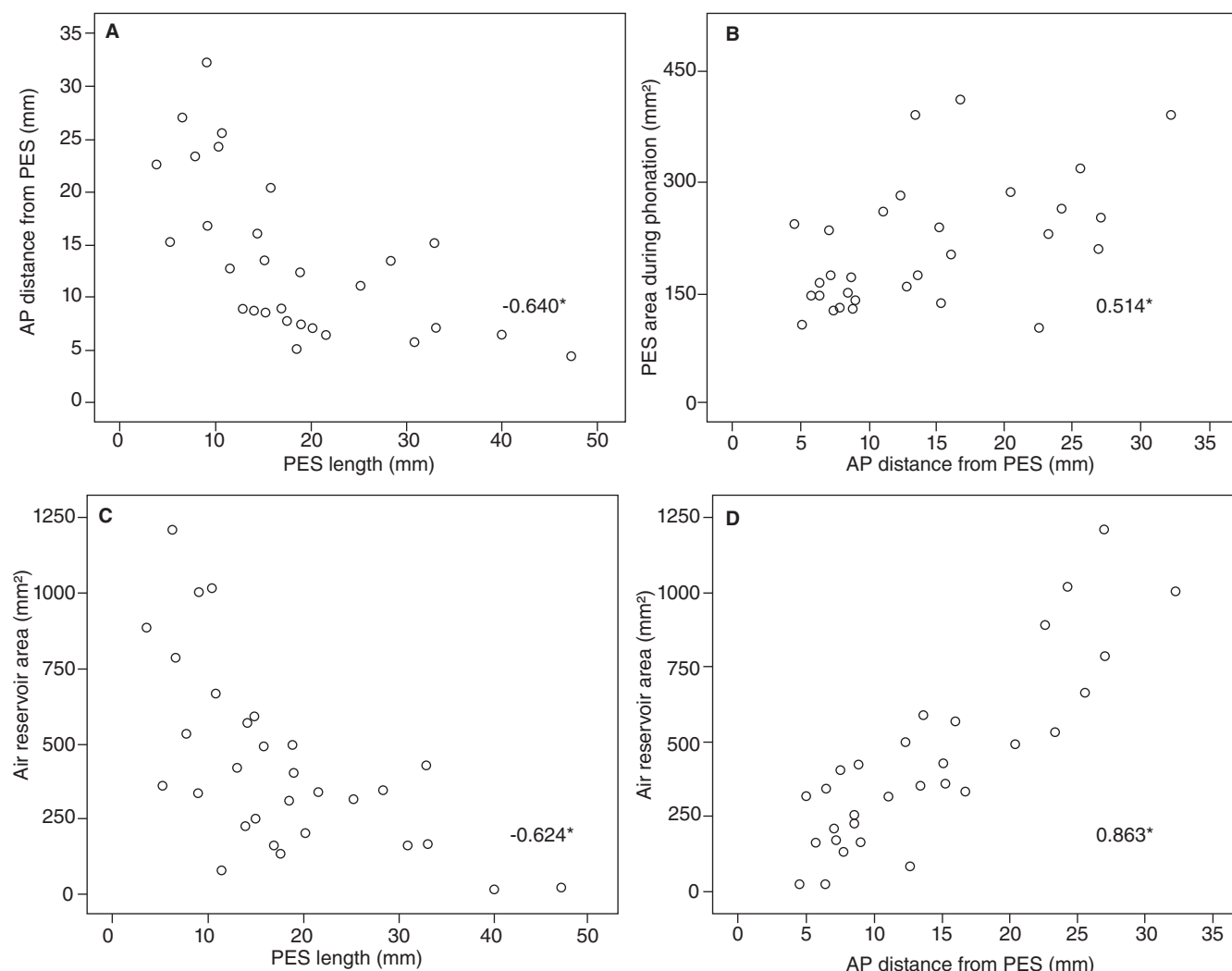
DISCUSSION

The aim of the present study was to evaluate the configuration of the PES and its relationship to acoustic spectrography signal types. To achieve this aim, PES was evaluated using VF. In the qualitative evaluation, the degree of PES prominence contact with its anterior wall was identified, and in the objective evaluation six measures were obtained, including PES APD, PES length, minimum distance between PES prominence and the anterior wall, PES projection areas during swallowing, during phonation, and of the air reservoir. The acoustic spectrographs of the two central seconds of the best emissions of the /a/ vowel for each patient were evaluated and categorized by signal Types I, II, III, and IV.

The patients with total laryngectomies and vocal prosthesis, independent for the extent of surgery or reconstruction with flap, PES manipulation (myotomy or application of toxina botulinica), gender, age, assistive treatment, time from surgery, and date of vocal prosthesis insertion were included in this study. Those data were evaluated, but there were no significant relationships of these with the aspects studied in this research.

In general, the minimum distance between the PES prominence and its anterior wall was measured. In the sample studies, only six patients presented with values different from zero, with a maximum distance of 5 mm. Also, during the qualitative evaluation of the VF, it was observed that, in spite of 23 (out of 29 patients) patients with phonation presenting contact between the PES prominence and the anterior wall, there was great variation in the degree of contact. Earlier studies have cited that the average minimum distance between the PES prominence and its anterior wall does not reflect the extension of contact of the tissue in that area. From this observation, it was decided to qualitatively evaluate the degree of contact between the PES prominence and its anterior wall, beyond measuring the distance, classifying the contact as absent or mild contact, or moderate or intense contact (Figures 1–3).

An analogy may be made between the degree of the PES prominence contact with its anterior wall and the PES degree of tone, which has been amply described in the literature as hypotonic, normotonic, and hypertonic. In the meantime,



*significance level 0,01

Caption: PES = pharyngoesophageal segment; AP = anteroposterior

Figure 5. Correlations between objective measures of pharyngoesophageal segment: anteroposterior distance of the pharyngoesophageal segment, pharyngoesophageal segment length, projection area of the pharyngoesophageal segment in phonation, and projection area of the air reservoir

because we did not measure the PES tones, we opted to use, in this article, the terminology hypocontact, normal contact, and hypercontact. This way, the group with no contact or mild contact (11 patients; 36.7%) presented with PES with hypocontact; the group with moderate contact (nine patients; 30%) presented with PES with normal contact, and the group with intense contact (10 patients; 33.3%) presented with PES with hypercontact. Of the group with intense contact, only one patient presented with absence of phonation and was classified as presenting with PES spasms, following the criteria found in the literature. The patients who had experienced myotomy were kept in the case series without forming a special group, since patients with these characteristics were found in the three types of PES wall contact groups: hypocontact, normal contact, and even in the group with hypercontact.

To better objectively characterize the PES and the relationship to TE voice proficiency, several measures were carried out. Two measures of previous studies were kept, such as “PES anteroposterior distance” and “PES length”, because especially

the latter, presented with direct correlations with voice proficiency. The other measure added “minimum distance between PES prominence and its anterior wall” has been demonstrated in other studies as being important. This measure was taken only during phonation and not during rest, because it has already been demonstrated that there is a positive correlation between these measures, such that good TE speakers present lower distance values during rest and during phonation. The measures “PES projection area during swallowing” and “PES projection area during phonation,” were evaluated to verify if a relationship exists between the difference of PES compression of swallowing during phonation and the degree of resistance offered by the PES during air flow. The measure “projection area of air reservoir” was added because it is possible that this measure of area may be more precise of the air behavior in the proximal esophagus than a straight line measure of the maximum subneoglottal distance. There was great variability among the objective measures of the patients with TE voice, primarily in the measures in this area.

The extraction of objective acoustic parameters, such as fundamental frequency, and perturbation and noise indices, are not always reliable with TE voice, because this voice is quite aperiodic and presents itself very low. Upon these difficulties, the voices were classified into signal Types I, II, III, or IV, using visual evaluation of the spectrographic traces, observing the harmonic traces during emissions, with and without spectral noise, fundamental frequency stability and signal amplitude. In the results, it was observed that 14 patients (48.3%) presented with signal Type IV, where there is a complete absence of harmonics or absence of harmonics for more than one second. Therefore, most of the sample studied presented a preponderance of noise in the acoustic spectrography, much larger than the population previously studied, that described this type of signal in only eight (out of the 39 patients) patients being studied (20.51%).

There was a significant difference ($p=0.009$) between the acoustic spectrography and the PES qualitative evaluation (Table 1). The majority of the patients with spectrography signal Types I-II presented moderate contact of the PES prominence and the anterior wall, those with signal Type III presented intense contact, and those with signal Type IV presented absence of contact or a mild contact of the PES with its anterior wall. These findings indicate a relationship between the degree of contact of the PES walls and the type of acoustic signal in the acoustic spectrograph. Therefore, patients with normal PES contact present with signal Types I-II; hypercontact with signal Type III, and hypocontact present with signal Type IV. A previous study presented similar findings, and the authors affirm that hypertonicity (hypercontact) influences the stability of signal Type III, and hypotonicity (hypocontact) is related to the lack of harmonics in the signal Type IV.

A significant relationship was also observed between the acoustic spectrography and the objective evaluation of the PES (Table 2). The group of patients with spectrographic signal Types I-II presented a PES length shorter than the group of patients with spectrographic signal Type IV. This relationship between shorter PES and better voices, both esophageal and TE voices, was also described in previous studies. Probably, shorter PES vibrates in a more regular fashion and with a pattern closer to the vibration of the vocal folds, generating a signal with more periodicity than a longer PES. Other authors confirm this fact and describe more jitter when PES is longer.

The group of patients with spectrographic signal Types I-II also presented with PES APD longer than the group of patients with spectrographic signal Type IV (Table 2). The voice quality is superior when the PES APD is longer. This measure may be compared to the subneoglottal maximum distance, which is therefore, a measure below the PES; and is a major indicator of increase of that distance, of rest for phonation, was also relevant for a better vocal quality.

When estimating the correlation between the degree of contact of the PES prominence with its anterior wall and the PES areas during swallowing and phonation (Figure 4), it was observed that there was a moderate and positive correlation only in the group with intense contact of the PES prominence and its anterior wall. In this group, it was observed that the larger

the PES area during swallowing, the larger the PES area during phonation. It may possibly occur because the PES with hypercontact is more rigid, shaping itself less during air flow. This relationship was not evaluated in the previous research work.

With regard to the objective measure of the PES during phonation, following are the two specific areas measured: 1) the PES projection area during phonation; and 2) the projection area of air reservoir. These areas are related to the PES length and APD, as follows:

The larger the PES area during phonation, the longer is the APD and the longer is the PES APD, and the shorter is the PES length (Figure 5, Scattergrams A and B).

The larger the area of air reservoir, the longer is the APD and shorter is the PES length (Figure 5, Scattergrams C and D).

The best spectrographic signal (Types I-II) correlated with the longer PES APD and the shorter PES length (Table 2). It can be affirmed, therefore, that an ideal configuration includes longer PES APD, shorter PES length, and larger projection areas of both phonation and air reservoir.

Correlations between PES objective measures have also been previously evaluated, and the authors indicate that increased length of the area below the PES has a relationship to increased measures of the PES prominence, but does not relate these findings to a better or worse vocal quality.

FINAL CONSIDERATIONS

It is possible to distinguish a pattern of behavior correlated to the degree of PES contact with its anterior wall. The group with moderate contact between the PES prominence and the anterior wall presents an ideal condition, with normal contact, because the contact is neither mild nor too intense (Figure 2). This group was the one that presented with the majority of acoustic spectrographic signals Type I-II (Table 1), characterized by a stable signal during two second and harmonics, observed up to at least the 1000 Hz band. When correlating the acoustic spectrography with the PES objective measures, the group with signal Types I-II presented with longer PES APD and shorter PES length (Table 2). Also, an association was found between longer PES APD and shorter PES length, with the PES larger area and larger projection area of the air reservoir during phonation (Figure 5).

The patients with degree of intense contact of the PES prominence with its wall may be classified as patients with hypercontact PES, with the exception of one patient, who presented with an absence of phonation and was classified as having PES spasms. Of the nine remaining patients, the majority presented with spectrographic signal Type III (Table 1). The PES and esophageal measures with patients with spectrographic signal Type III are similar to measures Type I-II group, where no statistical differences were found between the PES APD and length (Table 2). Also, the larger the PES area during swallowing, the larger the PES area during phonation in the group with intense contact (Figure 4), which may indicate a PES more rigid lending itself to be less shapeable during air flow.

Finally, the group of patients with PES prominence with absent or mild contact with its anterior wall may be classified as belonging to the class of PES with hypocontact, and the

majority present with spectrographic signal Type IV (Table 2). The PES measurements of the Type IV group are differentiated from the Type I-II groups by shorter PES APD, a longer PES length (Table 2), smaller PES area during phonation, and smaller projection area of air reserves (Figure 5).

It was found that the PES measurements clearly distinguish the group with moderate contact of the PES prominence with its anterior wall and signals Type I-II, from the group with absent or mild contact and signal Type IV. In other words, they distinguish between PES with normal contact from PES with hypocontact.

It is even possible to infer that PES intense contact of the prominence with its anterior wall is really an aspect that differentiates the group with hypercontact and signal Type III, because the PES measurements are very similar to the measurements of the group with normal contact. Therefore, a patient with PES hypercontact will benefit from the manipulation of the PES, either by myotomy or the application of toxina botulinica, so that there is a softening of the wall contact during phonation, facilitating the emission of more adequate TE voice.

CONCLUSION

In the evaluation of the PES configuration via VF, and of the acoustic spectrography of the voices of 30 patients with total laryngectomies and vocal prosthesis, it was observed that the best TE voices are produced by PES with normal contact with its wall, shorter in length, and with longer APD. The type of spectrographic trace is associated to the degree of contact of the PES wall during phonation and this differentiates the patient with PES with hypercontact from normal contact in the PES configuration, which is only the degree of contact of the PES prominence with its anterior wall.

**MLF was responsible for the elaboration of the tests, data collection, data tabulation, and the elaboration and write-up of the manuscript; MB was responsible for the delineation of the study, orientation with respect to the data analysis, and the elaboration of the manuscript; and AJG was responsible for the general orientation of the delineation of the study and the manuscript.*

REFERENCES

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 Phoniatri.* 1961;13:174-86.
9. Fouquet ML, Gonçalves AJ, Behlau M. Relation between videofluoroscopy of the esophagus and the quality of esophageal speech. *Folia Phoniatri 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 Phoniatri 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 Phoniatrics 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 Phoniatri 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.