On the Development of an Acoustic Pattern for Esophageal Speech in European Portuguese

Susana Vaz-Freitas[1 2 3], Elisabete Cardoso [4 5], Luís Coelho[4 6], Ricardo Sousa[3], Diogo Coelho[4]

[1]Serviço Otorrinolaringologia do Centro Hospitalar Universitário do Porto – Hospital de Santo António (CHUP-HSA), [2]Escola Superior de Saúde – Universidade Fernando Pessoa. [3]Laboratório de Inteligência Artificial e Apoio à Decisão- INESCTEC [4]Instituto Superior de Engenharia do Porto, [5]Hospital-Escola da UFP, [6] Centro de Inovação em Engenharia e Tecnologia Industrial,

Introduction

Voice rehabilitation in laryngectomized is a challenge, with the goal of restoring voice to patients unable to perform laryngeal phonation. However, current available methods can lead to the production of poorly intelligible and naturally impaired voices. In addition, most don`t reflect the patient's pre-morbid vocal identity.

On the other hand, from a speech language pathologist point of view, it`s important to know what acoustic patterns can be expected when using a esophageal voice (ES) or a voice with a tracheoesophageal prosthesis (TES), in order to define and adjust the therapy plan. To the European Portuguese language, which has phonetic-acoustic specificities, there are still insufficient references to the study of alaryngeal voices.

Methodology

This study characterizes ES and TES of European Portuguese speakers.

- ✓ 10 male cases who had a total laryngectomy and did Speech Therapy at CHUP-HSA.
- ✓ European Portuguese Oral Vowels (/a/, /i/ and /u/) were used, first in isolated production (for Maximum Phonation Time (MPT) measurement); then presented as a word, in written form and with the combination of occlusive consonants /p/, /t/ and /k/ (for acoustic analysis).

Results and Discussion

MPT values are found in table 1. The differences are evident and significant when comparing the type of voice with the duration of vowels emission, with superiority for TES. This values are in accordance with the previous investigations developed in Portugal [1] and abroad [2] [3] [4].

Table1: Values of the maximum phonation time according to the type of speech used

TYPES OF SPEECH						
Vowel	ES			TES		
	Duration (seconds)± SD	Min	Máx	Duration (seconds) ± SD	Min	Máx
/a/	1,87±1,04	0,40	4,95	**4,32±1,43**	0,55	9,62
/i/	1,80±0,64	0,41	3,33	**3,49±1,12**	0,65	7,03
/u/	1,14±0,62	0,27	2,33	**2,49±0,81**	0,50	6,87

The parameters f_0, 1st and 2nd Formant were analyzed according to the type of speech, ES or TES (table 2). The values found for f_0 in the TES speech are in line with studies already carried out. We found significant differences between the two speech types for /i/, in all of the analyzed parameters. Regarding the 1st Formant, in the esophageal voice it is in agreement with the values reported in investigations of 2018 [5]. Concerning the TES voice, Formants have lower values than what is described in the literature [5].

Table2: Values of f_0, 1st and 2nd formant according to the type of speech used

TYPE OF SPEECH									
	ES			TES			p-value		
	/a/	/i/	/u/	/a/	/i/	/u/	/a/	/i/	/u/
f_0 -Mean±SD	565,74±504,62	617,55±564,13	213,45±151,15	271,40±286,62	137,23±63,56	353,07±2,90	**,004**	**,000**	,582
1st Formant (Hz) ±SD	792,16±50,26	777,87±507,49	512,62±38,91	719,84±110,07	391,35±221,87	527,33±149,56	**,001**	**,016**	**,002**
2nd Formant (Hz) ±SD	1592,54±102,50	2554,02±152,69	1103,03±212,96	1689,53 ±108,28	2288,17±319,01	1404,68±531,38	,077	**,000**	,670

Regarding the isolated vowels' analysis of the parameters 1st and 2nd Formant, Jitter, Shimmer and HNR (table 3), the null hypothesis for the 1st Formant, 2nd Formant and the HNR was rejected, which indicates that these parameters are statistically different, according to the studied vowel.

Table3: Acoustic Parameters 1st, 2nd formant, Jitter, Shimmer and HNR according to the vowel /a/, /i/ or /u/

PARAMETERS	/a/	/i/	/u/	p-value
1st Formant (Hz) ±SD	750,58±15,15	545,96±60,60	522,65±18,82	**,000**
2nd Formant (Hz)±SD	1648,31±18,22	2394,51±43,84	1308,70±71,40	**,000**
Jitter (%) ±SD	4,15±0,42	5,69±0,56	5,34±0,44	,133
Shimmer (%) ±SD	18,26±1,17	18,11±1,53	15,60±1,06	,356
HNR (dB) ±SD	3,22±0,41	4,90±0,51	4,36±0,56	**,049**

This data heightens the similar level of perturbation that crosses over the three analyzed vowels [6].

Conclusion

The obtained results allow the first sketch of alaryngeal voice patterns for European Portuguese speakers. Sample size must be enlarged and acoustic analysis algorithms need to be calibrated to this specific alaryngeal sound production, to overcome methodological bias. These results and the produced data contributed to a software development – LxReahb – with interfaces that support a distance consultation. A reality in nowadays clinical settings.

THE VOICE FOUNDATION

References

[1] Vaz-Freitas, S., Cunha, D., Tavares, D., Dias, I., Neto, H., Cabral, P., Teixeira, P., "Tempos máximos de fonação na voz alaríngea," in 60° Congresso Nacional da SPORL, 2014.

[2] Siric, L., Rosso, M., Sos, D., Stevanovic, S., "Objective Assessment of Tracheoesophageal and Esophageal Speech Using Acoustic Analysis of Voice," Coll Antropol., vol. 36, pp. 111-114, 2012.

[3] Schindler, A., Mozzanica, F., Ginocchio, D., Invernizzi, A., Peri, A., Ottaviani, F.,Voice-related quality of live in patients after total and parcial laringectomy, Italy: University of Milan, 2012.

[4].Villarín, M., "Evaluación acústica y análisis prosódico de la voz esofágica, Universidad de Sevilla., 2006.

[5] Ng, L., Yan, N., Chan, V., Chen, Y., Lam, P., "A Volumetric Analysis of the Vocal Tract Associated with Laryngectomees Using Acoustic Reflection Technology," Folia Phoniatrica et Logopaedica, vol. 28, pp. 44-49, 20.

[6]]Eugenia, A., La Mantia, I., Bianco, M., Drago, G., Le Fosse, M., Azzolina, A., Grillo, G., Saita, V., "Verbal performance of total laryngectomized patients rehabilitated with esophageal speech and tracheoesophageal speech: impacts on patient quality of life," Psychology Research and Behavior Management, vol. 12, pp. 675-681, 2019.