Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 164 (2019) 654–661

Procedia Computer Science

www.elsevier.com/locate/procedia

CENTERIS - International Conference on ENTERprise Information Systems / ProjMAN - International Conference on Project MANagement / HCist - International Conference on Health and Social Care Information Systems and Technologies, CENTERIS/ProjMAN/HCist 201

Parameters for Vocal Acoustic Analysis - Cured Database

Joana Fernandes[a], Leticia Silva[a,b], Felipe Teixeira[a], Victor Guedes[a,c], Juliana Santos[a,d], João P. Teixeira[a,e] *

[a] *"Instituto Politécnico de Bragança, Bragança 5300, Portugal"*
[b] *"Universidade Tecnológica Federal do Paraná, Câmpus Cornélio Procópio, Brasil"*
[c] *"Universidade Tecnológica Federa do Paraná, Câmpus Medianeira, Brasil"*
[d] *"Universidade Tecnológica Federal do Paraná, Câmpus Campo Mourão, Brasil"*
[e] *" Research Centre in Digitalization and Intelligent Robotics (CEDRI),Applied Management Research Unit (UNIAG), Bragança 5300, Portugal"*

Abstract

This paper describes the construction and organization of a database of speech parameters extracted from a speech database. This article intends to inform the community about the existence of this database for future research. The database includes parameters extracted from sounds produced by patients distributed among 19 diseases and control subjects. The set of parameters of this database consists of the jitter, shimmer, Harmonic to Noise Ratio (HNR), Noise to Harmonic Ratio (NHR), autocorrelation and Mel Frequency Cepstral Coefficients (MFCC) extracted from the sound of sustained vowels /a/, /i/ and /u/ at the high, low and normal tones, and a short German sentence. The cured database has a total number of 707 pathological subjects (distributed by the various diseases) and 194 control subjects, in a total of 901 subjects.

Peer-review under responsibility of the scientific committee of the CENTERIS -International Conference on ENTERprise Information Systems / ProjMAN - International Conference on Project MANagement / HCist - International Conference on Health and Social Care Information Systems and Technologies.

Keywords: Laryngeal Speech Features; MFCC; Speech Database; HNR; NHR; Autocorrelation.

* Corresponding author. Tel.: +351 273 30 3129; fax: +351 273 30 3051.
E-mail address: joaopt@ipb.pt

1877-0509
Peer-review under responsibility of the scientific committee of the CENTERIS -International Conference on ENTERprise Information Systems / ProjMAN - International Conference on Project MANagement / HCist - International Conference on Health and Social Care Information Systems and Technologies.
10.1016/j.procs.2019.12.232

1. Introduction

A disturbance in voice has implications for a person's social and professional life. In patients with progressive pathologies, it is important to have access to a rapid diagnosis so that a better treatment and prognosis is possible [1].

There are a number of tests like laryngoscopy, endoscopy, stroboscopy or electroglottography, that can be done to detect voice-related conditions, but they are invasive and may be uncomfortable for patients, and may even induce vomiting and the auditory examination depend on the doctor's experience [2] [3] [4].

Vocal acoustic analysis is receiving increasingly attention from scientific researchers as an advantageous alternative to diagnose pathologies. Vocal pathologies were the first receiving contributions from the acoustic analysis for diagnose pathologies, since vocal apparatus is used to the production of speech and damages can directly interfere in the speech quality [4].

The acoustic analysis allows to measure acoustic signal properties of a recorded voice, whether it is continuous speech or a sustained vowel [4] [5]. This analysis is able to provide the waveform format allowing the evaluation of certain characteristics such as frequency disturbance measurements, amplitude disturbance measurements and noise parameters.

In this article, it is described the set of pathologies and parameters. These parameters were extracted from the speech of patients with pathologies and speech of the control group. It is also described the organization of the cured database. It aims to publicize the availability of the cured database to the scientific community.

In this work, it is intended to complete the cured database, initiated in a previous research adding more parameters and more diseases. This database includes now Mel Frequency Cepstral Coefficients (MFCC) parameters with 13 cepstral coefficients, Harmonic to Noise Ratio (HNR), Noise to Harmonic Ratio (NHR), absolute jitter, relative jitter, absolute shimmer, relative shimmer, extracted from 9 voice records corresponding to 3 vowels in 3 tones and one sentence, for subjects with 19 pathologies, plus the control subjects.

The database is available in .xls format, through the link: *http://www.ipb.pt/~joaopt/* ; menu: *products*; link: '*Base de Dados Curada de Parâmetros de Fala Patológica*'.

2. Pathologies in the Database

Lesions in the vocal chords change the process of phonation, since the patterns of vibration during the opening and closing phases of the vocal chords are irregular [3]. Taking into account this principle, this database provides parameters for the following list of laryngeal diseases:

- Chronic Laryngitis - corresponds to a persistent inflammation of the laryngeal mucosa, sometimes with many years of evolution, usually provoked by repeated acute infections [3];
- Dysphonia - is a communication disorder and makes vocal production difficult, with an impediment to voice production [2];
- Vocal Cord Paralysis - occurs when the laryngeal muscles cannot perform their function; paralysis of one of the vocal cords or of the two [3];
- Tumor of the larynx and hypopharynx – being both a tumor, what distinguishes them is the location, being one in the larynx and the other located in the back of the pharynx (hypopharynx) [6];
- Carcinoma of vocal cords - is a malignant cancer located in the vocal cords;
- Cysts - situated in the larynx or pharynx;
- Spasmodic dysphonia - is a neurological disease involving the vocal chords, where the muscles contract in an intense and irregular way (spasms), making the voice "tense" and "fragmented", conditioning the quality and fluency of the voice [7];
- Functional dysphonia - derives from the use of voice itself, i.e. when there is a disturbance in vocal behavior [8], this disturbance is considered if no anatomic change are known [2];
- Hyperfunctional dysphonia - occurs when there is excessive involuntary contraction of the phonatory musculature as a consequence of inappropriate voice use [9]. Results in an hoarse or strained voice [2];
- Hypofunctional dysphonia - there is a weakness in the laryngeal musculature, with incomplete closure of the glottis due to weak generalized or laryngeal musculature [10];

- Hypotonic dysphonia - occurs when there is a weak intrinsic musculature of the larynx (lack of strength of the vocal cords), which usually occurs after long periods of silence, causing a difficult onset of speech, improving as one goes about speaking [11];
- Psychogenic dysphonia - is a disorder of a psychological nature that is characterized by voice changes without a laryngeal structural injury or neurological disease [12];
- Reinke's edema - causes changes in the elasticity of the vocal folds and as a consequence, the voice becomes hoarser, since the space of Reinke is occupied by thick mucus. The space increases and the vocal folds increase of thickness and go into the larynx [13] [14];
- Fibroma - is a benign tumors located in the throat area. The place where it occurs most is in the vocal cords, since it is a disease that arises essentially in people who use their voice very frequently in their daily lives, such as the teachers and singers [15];
- Granuloma - relatively rare organic affections, which present a well-defined clinical and pathological picture, and may sometimes be confused with laryngeal cancer. These are benign tumors that present well-defined granulation and are characterized by lesions that generally appear as a localized mass of variable size, whitish, yellowish or reddish color [16];
- Intubation granuloma - is the same as described previously, but this situation occurs due to long periods of intubation, since the orotracheal cannula is supported in the vocal processes of the arytenoid cartilages, exerting certain pressure on the mucosa of the region [17];
- Laryngeal Dysplasia – consists most in laryngeal and hypopharyngeal cell cancers begin as a pre-cancerous condition. Although these cells appear abnormal, they do not resemble cancer cells [6];
- Vocal cords polyps - is a non-cancerous mass that arises as a consequence of excessive voice use, chronic allergic reactions in the larynx or chronic inhalation of irritants such as industrial pollutants or tobacco. This mass, as a rule, dilates during growth, until it is united to the surface by means of a pendulum [18].

This work intends to continue the work started in [19]. In the first work, only the chronic laryngitis subjects were included. Now the database was extended with other pathologies listed above, and with additional parameters, described below.

3. Speech Parameters

The parameters were extracted from speech files from the German Saarbrucken Voice Database (SVD) of the sustained speech sounds of vowels /a/, /i/ and /u/ at low, high and normal tones, and from the German sentence "Guten Morgen, wie geht es Ihnen?" ("Good morning, how are you?"). Sounds corresponding to the vowels consist of sustained speech of the steady state of the respective vowel and tone.

The SVD database speech signals have a sampling frequency of 50 kHz and a 16-bit resolution.

Parameters related with the variations of the periodicity such as jitter and shimmer were registered, namely the absolute and relative measures of jitter and also the absolute and relative measures of shimmer. The parameters that measure the harmonic and anharmonic components were also extracted and registered, namely the harmonic to noise ratio (HNR), noise to harmonic ratio (NHR) and Autocorrelation.

The parameters jitter and shimmer were extracted through the algorithm developed by Teixeira and Gonçalves, 2016 [20]. The parameters HNR, NHR and Autocorrelation were determined using the algorithm developed by Fernandes, 2018 [21] and the studies performed in [22]. For the MFCCs the mfcc Matlab function [23] was used, which returns the coefficients in the Mel Frequency Scale. However, before the extraction of this parameter, it was necessary to carry out a signal processing in order to select only the part of the signal where speech occurs. This processing was the same as the one used by Fernandes, 2018 [21].

3.1. Jitter Parameters

Jitter is defined as a measure of glottal variation between cycles of vocal cord vibration. Subjects who fail to control vocal cord vibration tend to have higher jitter values. Jitter can be measured in four different ways [24]. However, in this work, only two of these forms are used: relative jitter (jitter) and absolute jitter (jitta). The other two measures are not used, since previous statistical analysis [2] showed that the relative jitter has similar results.

Relative jitter (jitter) is the mean absolute difference between the consecutive glottal periods divided by the mean period and expressed as a percentage [4].

Absolute jitter (jitta) is the variation of the glottal period between cycles, that is, the mean absolute difference between consecutive periods [4].

3.2. Shimmer Parameters

The shimmer is related to the variation of magnitude along the glottal periods. A reduction in glottal resistance and lesions can cause variations in glottal magnitude correlated with air resonance in the vocal tract and noise emission, resulting in higher shimmer values. This can be measured in four different ways [24], however, in this work, only two will be used, Relative Shimmer (Shim) and Absolute Shimmer (ShdB). The other two measures will not be used, since previous statistical analysis [2] showed that the relative shimmer has similar results.

Relative Shimmer (Shim) is defined as the mean absolute difference between the magnitudes of consecutive periods, divided by the mean magnitude, expressed as a percentage [4].

Absolute Shimmer (ShdB) is expressed as the peak-to-peak magnitude variation in decibel, i.e. the base 10 algorithm of the absolute mean of the ratio of magnitude between consecutive periods multiplied by 20. It is expressed in decibel [4].

3.3. Harmonic Parameters

The harmonic characteristics of the voice can be measured into three parameters: HNR (Harmonic to Noise Ratio), NHR (Noise to Harmonic Ratio) and Autocorrelation. The HNR measures the relationship between harmonic and noise components. It provides an indication of overall periodicity of the speech signal by quantifying the relation between the periodic component (harmonic part) and aperiodic components (noise). The overall HNR value of a signal varies because different vocal tract configurations imply different amplitudes for harmonics. Therefore, different vowel can have different HNR. The /u/ vowel has higher frequency components, then, it is expected to have lower HNR. Since the HNR is sensitive to the vowel, HNR for different vowels should not be compared.

HNR is to measure the energy of the first peak of the normalized autocorrelation and consider that this is the energy of the harmonic component of the signal, and consider the remaining energy as the noise energy given by the difference between 1 and the harmonic energy, as represented in equation 1 [25]. In this equation, H is the harmonic component given by the energy of the first peak of the normalized autocorrelation of the signal. The final value of HNR is the average HNR along the successive segments.

$$HNR(dB) = 10 \times \log_{10} \frac{H}{1-H} \tag{1}$$

The NHR tends to be the inverse of the HNR, anyhow once the measure is made at the logarithmic domain (dB), their values tend to move in opposite directions but the values are not exactly neither the inverse nor the symmetric, because the final measure is an average along the signal [26]. The NHR can be calculated by equation 2.

$$NHR = 1 - Autocorrelation \tag{2}$$

The Autocorrelation function gives a measure of the similar parts of speech repeated along the signal. The higher the autocorrelation value, the greater are the repetitions of similar events along the signal [26]. To calculate the autocorrelation of a signal it is necessary to multiply the normalized autocorrelation of a segment of the speech signal by the normalized autocorrelation of a window (hanning). Therefore, the first peak of the signal segment is considered as the autocorrelation. This process is repeated segment by segment until the end of the signal. Finally, the average of all autocorrelations of the segments gives the autocorrelation of the signal.

3.4. Mel Frequency Cepstral Coefficientes - MFCCs

The Mel Frequency Cepstral Coefficients (MFCC) are calculated through the frequency spectrum of small windows of the speech signal, which is obtained by the Fast Fourier Transform (FFT) of the signal. In order to obtain an approximation of the perception of the human auditory system to the frequencies of sound, the frequency spectrum is subjected to a bank of triangular filters, equally spaced in the Mel frequency scale. Through the discrete cosine transform applied to the output of the filters the Coefficients in the Mel Frequency are determined. Finally, it is necessary to calculate the energy and a factor called the delta. The delta MFCC, and delta^2 MFCC intend to represent the dynamics of the signal along the time [1] [13].

For the extraction of this parameter a pre-processing was done, to select the speech beginning and end, removing the initial and final silence parts from the signal. For vowel speech signals the whole sound was considered as only one segment. For the speech signals of a sentence the window segments were overlapped in order to get exactly 50 segments with 35 ms long each. For both cases, the following parameters were used: pre-emphasis coefficient (0.97), frequency range to be considered in the analysis [300, 3700], number of filter bank channels (20), number of cepstral coefficients (13) and cepstral sine lifter parameter (22). The first MFCC corresponds to the energy of the segment.

4. Organizations of Cured Database

The cured database was built taking into account the sounds available in the Saarbrucken Voice Database (SVD).

19 diseases and control subjects were included in the cured database. It has a total number of samples of 707 pathological subjects (distributed by the various diseases) and 194 control subjects, and for each subject there is a sentence and three vowels /a/, /i/, and /u/, each vowel with three tones: high, low and normal. Table 1 shows the various speech pathologies, the number of patients by gender, for each pathology, mean and standard deviation age. In the SVD database the number of subjects for some groups, such as control, is greater than the number in table 1, since in the SVD database there are subjects who do not have all the sound files that are used. Another reason for this difference is due to the fact that there are subjects with more than one pathology and only those with only the pathology that identifies them are in the database, i.e. subjects with dysphonia have only dysphonia.

Table 1 – Groups used for the study, sample size, mean and standard deviation of the ages.

Test Groups	Sample size		Average Ages	Standard Deviation Age
	Female	Male		
Control	123	71	38,06	14,36
Dysphonia	40	29	47,38	16,27
Chronic Laryngitis	16	25	49,69	13,47
Vocal Cord Paralysis	102	67	57,75	13,77
Cyst	2	1	47,5	15,56
Vocal Cords Polyps	10	17	52,28	13,41
Carcinoma of Vocal Cords	1	18	57,00	6,60
Laryngeal Tumor	1	3	53,50	8,17
Granuloma	1	1	44,50	4,50
Intubation granuloma	-	3	53,00	11,22
Hypopharyngeal Tumor	-	5	59,50	9,29
Fibroma	1	-	46,00	0
Laryngeal Dysplasia	-	1	69,00	0
Reinke's edema	29	5	56,10	11,37
Functional Dysphonia	51	24	47,12	14,54
Hypofunctional Dysphonia	4	8	41,63	15,07
Hyperfunctional Dysphonia	95	32	42,32	13,62
Hypotonic Dysphonia	-	2	49,50	12,50

Psychogenic Dysphonia	38	13	51,40	9,40
Spasmodic Dysphonia	40	22	57,15	15,75
Total	554	347	-	-

In figure 1 it is possible to observe the distribution of the number of subjects by the different groups to which they belong.

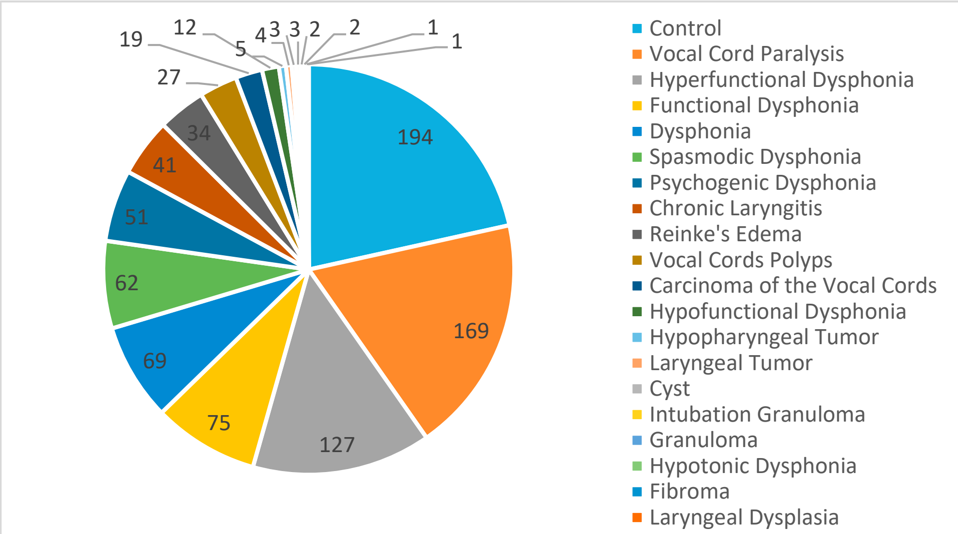

Fig. 1 - Number of individuals in each group

In the cured database there is an initial page containing the identification (id) of each subject, the age, gender, characterization group that corresponds to the subject and the abbreviation that is used for this group. In Fig. 2, it is possible to observe how this first part is constructed.

ID N.	Age	Gender	Group Characterization	Abbreviation
1714	51	F	Carcinoma of the Vocal Cords	CVC
110	62	M	Carcinoma of the Vocal Cords	CVC

Fig. 2 – First page of the cured database

For each group of pathologies there are 3 pages to organize the parameters. In Fig. 3, it can be seen that, for example, for the control group of subjects there are 3 pages with different parameters, which will be described next.

Fig. 3 – Parameter organization for each group of subject

The first page (Fig 4) of each group includes the parameters: absolute jitter, relative jitter, absolute shimmer, relative shimmer, HNR, NHR and Autocorrelation

Patient			Recorded Sound		Jitter		Shimmer		Harmonicity		
ID N.	Age	Gender	Vowel	Tone	Absolute	Relative (%)	Relative (%)	Absolute(dB)	HNR (dB)	NHR	Autocorrelation
1	20	F	a	High	15,792	0,589	1,370	0,120	31,135	0,001	0,999
				Low	13,314	0,246	1,036	0,091	27,652	0,002	0,998
				Normal	27,519	0,755	2,509	0,219	23,208	0,007	0,993
			i	High	11,171	0,419	0,372	0,033	34,232	0,001	0,999
				Low	18,765	0,346	2,268	0,198	17,776	0,020	0,980
				Normal	11,004	0,304	0,587	0,051	26,066	0,003	0,997
			u	High	11,348	0,432	0,611	0,054	36,249	0,000	1,000
				Low	22,683	0,422	1,840	0,160	28,280	0,002	0,998
				Normal	76,443	2,124	2,380	0,207	28,979	0,002	0,998

Fig. 4 – First page of the group with the set of 7 parameters

The second page lists the MFCCs of the recorded sentences, in this case there are 13 cepstral coefficients throughout the sentence, making a total of 50 sentence segments. In figure 5, it is possible to observe part of the distribution of this parameter. The 13 row contains the 13 MFCC along 50 column for the speech segments along the sentence.

Patient			Number of cepstral coefficients	Phrase Segments													
ID N.	Age	Gender		1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	20	F	1	40,629	50,285	58,260	63,675	41,252	58,353	63,249	62,148	59,934	59,439	66,060	67,152	67,268	67,510
			2	-0,672	9,484	4,680	-3,285	-2,223	-11,515	-4,380	-1,345	-3,054	-2,947	7,390	6,063	6,553	7,213
			3	0,393	8,230	1,826	-4,654	-0,011	2,150	0,169	3,506	4,801	2,086	2,000	-2,043	-0,624	0,261
			4	-6,836	-1,818	-10,142	3,021	-1,581	1,944	2,727	3,157	-5,813	-3,158	-11,373	-19,286	-19,049	-19,522
			5	-2,941	-5,460	-9,148	4,340	3,676	-0,609	-0,817	-3,957	-8,561	-3,041	-6,162	-0,060	-2,319	-1,453
			6	-7,544	-0,525	2,056	0,667	-0,785	-5,166	-4,726	-5,668	0,254	7,639	2,170	-0,455	1,143	-3,411
			7	-1,584	-1,276	-0,419	-10,280	-3,542	4,758	-12,052	-13,217	-8,392	-15,090	-5,179	1,514	-0,285	0,040
			8	-2,809	6,426	-5,786	-5,515	-4,471	3,795	0,087	7,684	7,285	-2,200	-2,454	-5,356	-8,342	-5,719
			9	3,873	12,630	-0,923	9,335	4,524	-5,939	-16,727	-3,228	-12,773	-10,075	-2,709	3,444	1,985	-0,851
			10	3,497	-0,425	-11,409	-9,383	-0,013	3,586	-8,101	-9,065	-14,634	-8,903	-2,047	3,627	4,210	8,777
			11	4,385	3,783	-5,179	-4,754	3,721	-2,481	-5,617	-6,734	-14,264	-12,203	2,787	0,970	-1,988	-2,123
			12	0,471	-2,334	-17,603	-5,302	-4,489	0,186	-2,307	-6,639	-5,656	-17,876	2,609	7,318	3,660	3,842
			13	-1,544	-0,326	-14,060	-2,354	0,463	1,256	11,778	-2,941	-12,571	-20,269	-3,364	-0,048	3,103	-2,344

Fig. 5 – Second page of the group with MFCC of the sentence

Finally, the third page lists the MFCCs for the vowels. This parameter was also extracted with 13 cepstral coefficients presented in the 13 columns, for each vowel and tone (Fig. 6).

Patient			Recorded Sound		Number of cepstral coefficients												
ID N.	Age	Gender	Vowel	Tone	1	2	3	4	5	6	7	8	9	10	11	12	13
1	20	F	a	High	67,965	1,349	-1,319	-16,046	5,959	-2,746	2,348	10,535	14,577	31,756	20,841	13,264	-9,660
				Low	70,221	6,675	-3,596	-15,870	-2,339	-2,356	-3,699	-0,234	4,346	17,483	-2,171	-1,152	-2,138
				Normal	68,907	0,177	-6,134	-18,689	-2,976	-3,239	-7,146	-3,722	-0,534	-0,666	-8,800	-9,654	-8,978
			i	High	65,039	-4,670	14,529	9,859	2,031	15,785	8,172	9,832	19,350	15,203	9,480	8,682	-4,401
				Low	58,543	-12,696	20,092	2,520	-0,820	-1,971	4,439	-11,153	9,117	7,770	0,258	5,430	-3,198
				Normal	61,472	-10,027	14,225	-2,513	-4,384	7,333	-7,956	-2,145	0,990	-5,933	-0,142	3,845	4,935
			u	High	60,170	5,780	12,087	4,268	3,904	8,561	10,458	24,627	19,265	15,055	5,075	-7,813	-7,507
				Low	54,649	9,062	4,293	2,642	0,409	-1,514	2,009	-6,877	10,762	0,152	-0,115	4,106	0,939
				Normal	56,375	5,787	-0,133	-8,789	-6,062	1,358	-1,196	0,642	4,806	-1,760	-1,035	-2,048	4,233

Fig. 6 - Third page of the group with 13 MFCC for the vowels

5. Conclusion

The paper presents the organization of the database of speech parameters extracted from the German Saarbrucken Voice Database for the vowels /a/, /i/ and /u/ spelled at high, low and normal tones and one short sentence, for subjects with 19 diseases and control subjects in a total of 901 subjects.

This database consists of the speech parameters of jitta, jitter, Shim, ShdB, Autocorrelation, HNR, NHR and MFCCs extracted by the previously mentioned algorithms of the speech sounds.

The database created has its organizational structure described in section 3. It is divided into 4 parts, containing the characterization of the subject, the parameters of the 3 vowels and 3 tones, the MFCC for the phrase, and the MFCC for the vowels and tones. This organization allows easy reading of acoustic parameters for future research.

The database is available in .xls format, through the link: *http://www.ipb.pt/~joaopt/* ; menu: *products*; link: '*Base de Dados Curada de Parâmetros de Fala Patológica*'.

Acknowledgements

The work is supported by the Fundação para a Ciência e Tecnologia under the project number UID/GES/4752/2019.

References

[1] N. Alves, 2016. "Diagnóstico Inteligente de Patologias da Laringe," MSc Thesis, Instituto Polit. de Bragança.

[2] J. P. Teixeira e P. Fernandes, 2015. "Acoustic Analysis of Vocal Dysphonia," *Procedia Computer Science* - Elsevier 64; 466 – 473.

[3] J. P. Teixeira, D. B. Ferreira e S. M. Carneiro, 2011. "Análise Acústica Vocal - Determinação do Jitter e Shimmer para Diagnóstico de Patologias da Fala," Proceedings of VI Congresso Luso Moçambicano de Engenharia; Maputo, Moçambique, ISBN: 978-972-8826-24-6.

[4] J. P. Teixeira, F. Teixeira, J. Fernandes e P. O. Fernandes, 2018. "Acoustic Analysis of Chronic Laryngitis - Satistical Analysis of Sustained Speech Parameters" In proceedings of 11th International Joint Conference on Biomedical Engineering Systems and Technologies, pp 168-175.

[5] J. P. Teixeira e N. A. Paula Odete Fernandes, 2017. "Vocal Acoustic Analysis – Classification of Dysphonic Voices with Artificial Neural Networks," *Procedia Computer Science,* vol. 121, pp. 19-26.

[6] E. Oncoguia, 2018. “A Laringe e a Hipofaringe,” [Online]. Available: http://www.oncoguia.org.br/conteudo/a-laringe/671/134/. [Accessed on 14 April 2018].

[7] “Disfonia Espasmódica,” [Online]. Available: https://www.saudecuf.pt/mais-saude/doencas-a-z/disfonia-espasmodica. [Accessed on 12 April 2018].

[8] A. Baena, 2013. “Disfonias,” [Online]. Available: http://www.sulms.saudeatual.com.br/especialidades/fonoaudiologia/disfonias. [Accessed on 12 April 2018].

[9] “Disfonia hiperfuncional,” 2015. [Online]. Available: https://www.cun.es/diccionario-medico/terminos/disfonia-hiperfuncional. [Accessed on 12 April 2018].

[10] “Disfonía hipofuncional,” 2015. [Online]. Available: https://www.cun.es/diccionario-medico/terminos/disfonia-hipofuncional. [Accessed on 12 April 2018].

[11] “Trastornos de la voz,” [Online]. Available: http://www.logopedas-castellon.com/patologias/trastornos-de-la-voz_disfonias_disfuncionales.html. [Accessed on 12 April 2018].

[12] M. Bergamini, M. Englert, L. Ribeiro e R. Azevedo, 2015. “Estudo de caso: disfonia psicogênica,” *CEFAC,* vol. 17.

[13] H. Cordeiro, 2016. “Reconhecimento de Patologias da Voz usando Técnicas de Processamento da Fala,” PhD Thesis, UniNOVA - Lisboa.

[14] R. Martins, M. Domingues, A. Fabro, N. Dias e M. Santana, 2009. “Edema de Reinke: estudo da imunoexpressão da fibronectina, da lamina e do colágeno IV em 60 casos por meio de técnicas imunoistoquímicas,” *SCIELO,* vol. 75.

[15] “El fibroma de las cuerdas vocales, el fibroma de la laringe – la causa, los síntomas, la diagnosis y el tratamiento,” 2015. [Online]. Available: http://terapiaherbal.com/el-fibroma-de-las-cuerdas-vocales-el-fibroma-de-la-laringe-la-causa-los-s%EF%BF%BD-ntomas-la-diagnosis-y-el-tratamiento/. [Accessed on 13 April 2018].

[16] F. Dieguez, M. Barbosa, A. Almeida, N. RaMaciel e V. Souza, 2010. “Granuloma Laríngeo: Relato de caso,” *Revista Cientifica da FMC,* vol. 5.

[17] R. Martins e N. Dias, *Complicações das vias aéreas relacionadas à intubação endotraqueal,* Botucatu: Faculdade de Medicina de Botucatu - Unesp.

[18] D. Merdau, “Pólipo nas cordas vocais,” [Online]. Available: https://www.infoescola.com/saude/polipos-nas-cordas-vocais/. [Accessed on 14 April 2018].

[19] J. Fernandes, F. Teixeira, P. Odete e J. P. Teixeira, 2018. “Cured Database of Speech Parameters for Chronic Laryngitis Pathology” In Proceedings of 31st International Business Information Management (IBIMA) Conference. Milan.

[20] J. P. Teixeira e A. Gonçalves, 2016. “Algorithm for Jitter and Shimmer Measurement in Pathologic Voices” Procedia Computer Science - Elsevier 100; 271 – 279.

[21] J. Fernandes, 2018. “Determinação da Autocorrelação, HNR e NHR para Análise Acústica Vocal,” MSc Thesis, Instituto Politécnico de Bragança.

[22] J. Fernandes, F. Teixeira e A. J. J. P. T. Victor Guedes, 2018. “Harmonic to Noise Ratio Measurement - Selection of Window and Length,” *Procedia Computer Science ,* vol. 138, pp. 280-285.

[23] MathWorks, “HTK MFCC MATLAB,” 2011. [Online]. Available: https://www.mathworks.com/matlabcentral/fileexchange/32849-htk-mfcc-matlab. [Accessed on 7 November 2017].

[24] J. P. Teixeira e A. Gonçalves, 2014. “Accuracy of Jitter and Shimmer Measurements” Procedia Technology - Elsevier, Volume 16, Pag. 1190-1199.

[25] P. Boersma, 1993. “Accurate short-term analysis of the fundamental frequency and the harmonic-to-noise ratio of a sample sound,” *IFA Proceeding,* vol. 17, pp. 97-110.

[26] P. Boersma, 2004. “Stemmen meten met Praat,” *Universiteit van Amsterdam.*