RESEARCHERID
THOMSON REUTERS

[ICRTCET-2018] ISSN 2348 – 8034
Impact Factor- 5.070

GLOBAL JOURNAL OF ENGINEERING SCIENCE AND RESEARCHES

NON-INVASIVE APPROACH: VOICE ACOUSTIC ANALYSIS OF LARYNGEAL CANCER PATIENTS

Mr.G.B.Gour[*1], Dr. V. Udayashankara[2], Dr. Vijaya Sorganvi[3], Ms. Sahana Choudri[4] & Ms. Soumya Alabnoor[5]

[*1]Assistant Professor, Dept. of ECE, V.P.Dr.P.G.Halakattit College of Engineering & Technology, Vijayapur, India

[2]Professor and Head, Dept of I.T, Sri Jayachamarajendra College of Engineering, JSS Science and Technology University, Mysore, India

[3]Dept. of Community Medicine, Shri B.M.Patil Medical College, Vijayapur, India

[4&5]UG.student, Dept. of ECE, V.P.Dr.P.G.Halakattit College of Engineering & Technology, Vijayapur, India

ABSTRACT

Voice analysis is one of the emerging non invasive approach in the medical field. Voice dysfunction is an important complication in laryngeal cancer patients. The main objective of the present study is to investigate the features of voice changes in vocal cord carcinoma (VCC) patients. Non operative VCC patients are selected from freely available online Saarbruecken Voice Database (SVD). Acoustic voice analyses of VCC patients was performed using GNU-Octave, an open source simulation tool equivalent to Matlab on Linux platform, which makes the present work cost effective. The eight out of thirteen patients were reported with pronounced dysponia and phonatory discomfort. Preoperative voice is slightly abnormal with fundamental frequency (Fo). Absolute jitter, relative jitter, absolute shimmer and relative shimmer are found to be significantly different from those of control group ($p<0.05$). Positive significant correlation is observed between perturbation measurements with $p<0.001$. The present study emphasizes the development of a non invasive diagnostic tool for cancer patients by considering more number of acoustic parameters with their potential to index affective physiological changes in the voice production.

Keywords: *Voice acoustic analysis, Vocal cord carcinoma, Voice disorder analysis, Laryngeal cancer, Postoperative dysphonia, Vocal tract steadiness, Biostatistics*.

I. INTRODUCTION

Ninety percent of laryngeal cancer is squamous cell carcinoma. Smoking, alcohol abuse, lower socioeconomic status and for above 60 years, increases the risk. Early diagnosis is common with vocal cord tumors because of early development of hoarseness. Among three common sites of laryngeal cancer, vocal cord carcinoma is more predominant. Most of the tests and procedures used to detect, diagnose and stage the laryngeal cancer are invasive with other health side effects like observed in Radiotherapy and Chemotherapy. These evaluation methods includes biopsy, laryngoscopy, endoscopy and Radiologic imaging methods. Their costs varies from few lakhs to thousands and more over is not a suitable for routine use for a common Indian (In 2014, as per [1], 21.9% of the population who live below the national poverty line) patient. According to NCRP (National Cancer Registry Program) report, by the year 2020, the cases of head & neck cancers are estimated to be around 218,421 and the projection for Laryngeal cancers is from 25,172 in 2010 to 33,885 by 2020(15.5%). It shows that the Laryngeal cancer is one of the emerging cancer in India. The negative impact of a voice disorder is social, psychological, professional, and economic (as in the case of a singer or actor)[2]. Pathological diagnosis of vocal tract is a field, which demands further investigation due to difficulty in standardizing diagnosis process of the speech pathologists [3].

Voice analysis with sustained vowels is more efficient as it avoids articulatory effects and useful for the estimation of vocal performance. The voice production system uses most part of its mechanisms (e.g., glottal flux of constant air, vibration of the vocal folds in a continuous way, etc.) in the phonation of vowels and is useful in detecting anomalies of such mechanisms[4].

Section II will discussing about materials and methods used in the present study. Section III is discussing about results and section IV concludes the present work.

II. MATERIALS AND METHODS

A. Materials

First, Saarbruecken Voice Database (SVD), version 2.0 is used for the present study. The SVD is a freely available online voice database containing a collection of voice recordings of different pathologies, including both functional and organic[5]. Thirteen male vocal cord carcinoma subjects without any surgery are selected and treated as experimental group with age (56.92 ±9.37) and control group containing thirteen male healthy subjects with age (56.62± 8.94) are considered for the present study. The sustained vowels /a/ are recorded at sampling frequency of 50kHz with 16 bits

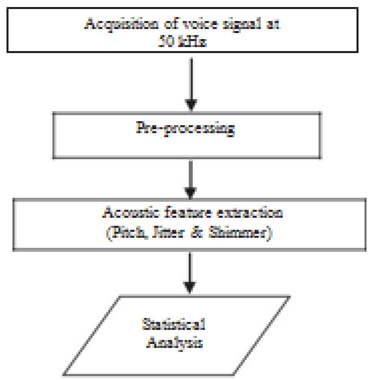

Fig. 1

resolution. All patients had no head and neck surgery, or head and neck trauma.

B. Methods

In the medical field the voice acoustic analysis is used in diagnostic and treatment of voice disorders. As the patients are normally reluctant to video laryngoscope (VLS), because of irritation to patient, examination also requires topical anesthesia in many patients. The patient's mouth should be open with anterior traction on the tongue which also prevents examination of the larynx during speech or swallowing [6].

Whereas, the stroboscope which provides a flashing light to simulate a slowed vocal-cord vibration, and has a microphone to sense the fundamental frequency of the vibrations and to coordinate the flashing light to the same frequency. If the light flashes at the fundamental frequency, the stroboscope will freeze one image in the vibratory cycle. Thus, it will not document the status of vocal function with cycle to cycle information. The stroboscopic parameters include closure pattern, amplitude of vibration, mucosal wave, symmetry, and regularity, and these objective data are visual perceptual ratings and subject to reliability and validity errors. Besides, the existing diagnostic systems available for laryngeal cancer patients are prone to radiation and are costlier. In regard to above discussion the voice analysis will help the clinicians todocument the status of vocal function with cycle to cycle information that is not available with stroboscope [6]. The methodology of the present work is as shown in the Fig.1. The methodology is implemented in Linux platform and using GNU-Octave, an open source simulation software similar to Matlab[7]. GNU-Octave uses scriptic language similar to the Matlab. It is one of the major step included for the cost effective implementation of the present work.

As speech signals fall off at higher frequencies, the higher frequencies are compensated by using pre emphasizing filter as in (1) with Į = 0.9.

ISSN 2348 – 8034
Impact Factor- 5.070

y(n) = x(n) – Į x(n-1) (1)

This pre emphasis helps to make stationary behavior within frames. Then the speech signal is converted into frames of length 25 ms with frame step of 8 ms. The unwanted noise due to framing is minimized by using hamming window which is tapered at beginning and end edges of the frame to give a less distorted and smoother spectrum[8].

In the feature extraction step the analysis indicators included are average fundamental frequency (Fo), the fundamental frequency perturbation (Jitter), the amplitude perturbation (Shimmer) variations, as follows:

A *Fundamental frequency (Fo):* Pitch or fundamental frequency is calculated by using autocorrelation method, searching for a peak in normal pitch range in speech of 50Hz to 500 Hz [9,10].

B Jitter and Shimmer: Jitter and Shimmer are used as indicators of harshness of voice and relative jitter and shimmer are given as a percentage that represents the maximum deviation from a nominal frequency or amplitude [11, 12, 13]. If N is the total number of Pitch periods, T be the fundamental frequency and the amplitude A of particular period then, relative/absolute jitter and shimmer are calculated using (2),(3), *(4)*and (5).

*N*í 1
1
™ |*T*i*í T*i+1|
Jitter (absolute) *N* í 1
= *i*í 1 (2)

N í 1
1
™ |*T*i*í T*i+ 1|
*N*í 1 *i*í 1
N
1
™ *T*i
Jitter(relative) *N*
= *i*= 1 (3)

*N*í 1
1
™ |20 log (*A*i+1 / *A*i)|
Shimmer(dB) *N* í 1
= *i*=1 (4)

Table.1

Analysis Indicators	Fo	Absolute Jitter	Absolute Shimmer (dB)	Relative Jitter	Relative Shimmer
Experimental Group (n =13)	202.51±46.06	0.0004±0.000 5	1.50± 1.35	0.07±0.0 9	0.02±0.011
Control Group (n=13)	183.99±27.89	0.0001±0.000 03	0.61±0.23	0.02±0.00 4	0.01±0.004

$$\text{Shimmer(relative)} = \frac{\frac{1}{N-1}\sum_{i=1}^{N-1}|A_i - A_{i+1}|}{\frac{1}{N}\sum_{i=1}^{N} A_i} \qquad (5)$$

Here N indicates the number of frames. These acoustic features are computed in linux environment using GNU-Octave simulation tool. Script based language has been used while extracting voice acoustic features. The computed parameters are then subjected to statistical analysis in linux [14].

III. RESULTS

The further statistical analysis is carried out by using statistical analysis tool in the linux platform. The descriptive statistical results for the comparison of Voice Acoustic Anlyses of study and Control groups shown in Table.1. The absolute jitter and relative shimmer shows significant difference where as, relative jitter has near significant difference between the experimental and control groups with $p<0.05$. Then all the features are subjected to statistical students 2-tailed 't' test with $p<0.05$ and 95% confidence interval as in Table.2.[15].

The student's 't' test is performed between study and control groups and the results are shown in Table.2. It is clear from Table.2 that amplitude and frequency perturbations like relative jitter and shimmer, absolute shimmer have shown significant difference between study and control groups, whereas absolute jitter is slightly significant with $p<0.05$. This signifies the presence of hoarseness in the VCC patients, which also matches with their diagnostic reports. This hoarseness is due to the asymmetrical structure of affected vocal cords. However, the pitch has not shown any significant variation between the groups.

As Cricothyroid muscle is associated with high degree of pitch and volume, remain not much affected due to insignificant variation of pitch between the groups. The thyroaretynoid and lateral cricoaretynoid muscles are related to fading of sustained phonation found to be under tension due to VCC impact. This is also clear from significant difference of perturbations between the groups. In vowel articulation hyoid bone movement is affected by the change in the activity of the tongue muscles and the function of the genioglossus posterior is to raise the tongue dorsum for high vowels by drawing the tongue root forward. Hence, the high vowels are associated with forward position and low vowels with back position of the hyoid bone. The more activity of genioglossus is observed only in high-pitched vowel /a/ and there is a direct correlation between the intrinsic pitch of the vowels and the activity of the posterior fibers of genioglosssus. In summary, the hyoid bone which is mainly concerned with pitch control, remains not much affected during sustained phonation, which is also clear from insignificant variation shown by the pitch [16,17].

Table.3 shows the computation of Pearson's correlation coefficients between significant perturbations. It is clear that there is a positive significant correlation between variations of jitter and shimmers values with $p<0.001$. But correlations of perturbations with pitch are not significant $p>0.001$.

IV. CONCLUSION

The prospective study has been made with VCC patients using SVD. The use of free open source tools like Linux platform and GNU-Octave simulation tools are useful for cost effective work. In view of costlier, invasive diagnostic methods available to VCC patients, Voice analysis is one of the emerging non invasive approach in the medical field. The statistical analysis of voice acoustic parameters of VCC shows significant variations of perturbations with $p<0.05$. However the pitch has not shown any significant variation in the groups.

ISSN 2348 – 8034
Impact Factor- 5.070

Table.2

Analysis Indicators		Student's 't'	'p' value	Difference mean	95% Confidence Interval of Difference (Lower limit to Upper limit)
Between Experimental Group (n=13) and Control Group (n=13)	**Fo**	0.227	>0.05	-18.52	-49.34 to 12.30
	Absolute Jitter	0.072	>0.05	-0.0002	-0.00057 to 0.00003
	Absolute Shimmer	0.026	<0.05	-0.8985	-1.6825 to -0.11446
	Relative Jitter	0.037	<0.05	-0.05403	-0.10454 to -0.00352
	Relative Shimmer	0.010	<0.05	-0.0094	-0.01633 to -0.00249

TABLE.3

Correlation between	Correlation coefficient	p value	Remarks
Fo & Absolute Jitter	r = -0.119	0.564	Insignificant
Fo & Absolute Shimmer	r = -0.191	0.349	Insignificant
Fo & Relative Jitter	r = -0.036	0.861	Insignificant
Fo & Relative Shimmer	r = -0.128	0.534	Insignificant
Absolute Jitter & Absolute Shimmer	r = 0.947	0.001	Significant
Absolute Jitter & Relative Jitter	r = 0.976	0.001	Significant
Absolute Jitter & Relative Shimmer	r = 0.890	0.001	Significant
Absolute Shimmer & Relative Jitter	r = 0.937	0.001	Significant
Absolute Shimmer & Relative	r = 0.952	0.001	Significant

Shimmer			
Relative Jitter & Relative Shimmer	r = 0.916	0.001	Significant

The Pearson's correlation has correlations between variations of perturbations with p<0.001. This indicates that much of the perturbations related study has to be done for more insight into the voice disorder analysis of VCC patients. However, by the discussion of results it is clear that selected acoustic features have physiological changes during voice production. By including spectral and non linear acoustic parameters may put more light on dynamical behavior of vocal cords. There is a scope for the development of cost effective, non invasive, alternative diagnostic tool to assist clinicians routinely.

REFERENCES

1. *"Poverty In India" Asian Development Bank. Retrieved 16 August 2017.*
2. *Ramnath Takiar, Deenu Nadayil, A Nandakumar,"Projections of Number of Cancer Cases in India (2010-2020) by Cancer Groups",*
3. *Asian Pacific Journal of Cancer Prevention, Vol 11, 2010. Quality Assessment Based on Higher Order Statistics", School of engineering, 2008.*
4. *Athanasios Tsanas, Matías Zañartu, Max A. Little, Cynthia Fox, Lorraine O. Ramig, Gari D. Clifford," Robust fundamental frequency estimation in sustained vowels: detailed algorithmic comparisons and information fusion with adaptive Kalman filtering", Journal of Acoustical Society of America, 2014.*
5. *David Mart´Õnez, Eduardo Lleida, Alfonso Ortega, Antonio Miguel, and Jes´us Villalba, " Voice Pathology Detection on the Saarbr¨ucken Voice Database with Calibration and Fusion of Scores Using MultiFocal Toolkit", Aragon Institute for Engineering Research (I3A), University of Zaragoza, Spain.*
6. *Cummings: Otolaryngology:" Head & Neck Surgery", 4th ed.Mosby.Inc,2005*
7. *Matthew W. Brewster and Matthias K. Gobbert, " A Comparative Evaluation of Matlab, Octave, FreeMat, and Scilab on Tara", Department of Mathematics and Statistics, University of Maryland, Baltimore County*
8. *Kinnunen T, Li HZ. "An overview of text-independent speaker recognition: From features to super vectors. J. of Speech Comm. 2010;52(1): 1-30.*
9. *Lawrence R. Rabiner, "On the use of autocorrelation Analysis for pitch detection", IEEE Transactions On Acoustics, Speech, and Signal Processing, Vol. ASSP-25, No. 1, Feb. 1977.*
10. *Kavita Kasi, "Yet another algorithm for Pitch detection", Old Dominion University, 2002.*
11. *Mireia Farrús, Javier Hernando, Pascual Ejarque," Jitter and Shimmer Measurements for Speaker Recognition", TALP Research Center, Department of Signal Theory and Communications,Universitat Politècnica de Catalunya, Barcelona, Spain.*
12. *João Paulo Teixeira and André Gonçalves, " Algorithm for jitter and shimmer measurement in pathological voices", Procedia Computer Science 100 (2016) 271 – 279.*
13. *Steven Bielamowicz,Jody Kreiman,Bruce R. Gerratt, Marc S. Dauer,Gerald S. Berke, "Comparison of Voice Analysis Systems for Perturbation Measurement", Journal of Speech and Hearing Research, Volume 39, 126-134, February 1996.*
14. *Paul Cobbaut,"Linux Fundamentals".*
15. *HarveyM otulsk, "Intuitive Biostatistics", Oxford Press, 1995.*
16. *Rosen .C.A, Simpson .B, "Operative Techniques in Laryngology- Anatomy and Physiology of Larynx" , Springer, http://www.springer.com/978-3-540-25806-3.*
17. *Relationship between the Pitch control and the Vowel articulations... by Kiyoshi Honda, 1983.*