

Objective Acoustic Analysis and Comparison of Normal and Abnormal Voices

HT LATHADEVI¹, SURESH PUNDALIKAPPA GUGGARIGOUDAR²

ABSTRACT

Introduction: Acoustic analysis is commonly used to diagnose, document and treat voice disorders. Type I and Type II voices which are nearly periodic can be easily assessed with computerised acoustic analysers. Widely used voice parameters like Jitter, Shimmer and Harmonic Noise Ratio (HNR) are indices of voice pathology and indicate different diseases and help as an outcome measure.

Aim: To study whether objective acoustic analysis is able to differentiate between abnormal and normal voices.

Materials and Methods: In the present study, the patients were made to phonate sustained vowel /a/ and the voice was recorded to analyse parameters like Jitter (ddp), Shimmer (dda), HNR and Median pitch using the acoustic software,

PRAAT. These parameters were compared with the values of the institute's own normal voice database. They were analysed for Mean, Median, standard error of mean, and Kolmogorov-Smirnov values.

Results: The range of abnormal voice parameters like Jitter (ddp), shimmer (dda), median pitch and HNR measures were different from normal voices. The difference was significant in jitter (p-value of 0.026) in males, in Shimmer (p-value of 0.035) in females. HNR did not show any significance.

Conclusion: The traditional methods of perturbation measures like Jitter, Shimmer, Median pitch and HNR can help the clinicians for characterisation of voice into either normal or abnormal voices. But the comparison needs a local or regional normal voice database.

Keywords: Hoarseness of voice, Voice assessment, Voice disorder

INTRODUCTION

Assessment of vocal function has been a challenge to clinician since time immemorial. Two approaches are used for the same, the perceptual and objective measurement based analysis. The perceptual assessment involves listening to patient's voice production. This is performed by an expert jury and hence is a subjective measure of assessment. The objective assessment is done by using computerized software which requires acoustic, aerodynamic measures through complex medical equipments like laryngoscopy, stroboscopy, electroglottography etc. Visual methods like these are important tools in basic voice research but have several limitations. Acoustic analysis appears to have an advantage over others because of its noninvasive nature and its potential for providing quantitative data with reasonable expenditure of analysis time [1].

The development of simpler and portable instrumentation for acoustic analysis would lead to programs similar to audiometric testing in schools and industry. Such instrumentation can have large benefits in terms of overall health maintenance [2]. The acoustic parameters like fundamental frequency, Median pitch, jitter, shimmer and HNR are useful in describing vocal characteristics. These parameters change e.g., Jitter has a typical value of variation behaviour 0.5% to 1% for sustained phonation in young adults [3]. Shimmer changes in breathiness and mass lesions of vocal cord. The values become altered in adults upto 3% and 1% in children [4]. The HNR ratio depends on periodic and non-periodic components of a segment of voiced speech and represented by dB. A value less than 7 dB is considered abnormal [5].

Usually, these voice parameters are measured at a well-equipped voice lab maintained by speech pathologists. But here, we have showed the way of assessing the voice objectively using easily accessible recording equipments and PRAAT software unlike other studies where they use customised voice lab equipments (Dr. Speech, MDVP CSL Speech etc.,) which are of high cost.

The normal voice database from these voice labs may not correlate with local population as the voices are influenced not only by the person, age, sex, time of the day, emotions of the person, disease but also by locality and region [6]. Most of them have felt that more regional databases and abnormal voice comparisons are necessary [7]. Besides, each health institution should have its own voice assessment protocol with its recording equipment and software for patient services and research. As there are no local databases or prior studies from our region, we felt the need of voice database and created the same. Therefore, the present study was conducted and abnormal voices were compared with the normal voices.

MATERIALS AND METHODS

This was a prospective comparative study wherein normal voices and abnormal voices were collected at Department of ENT, BLDE (Deemed to be University) Shri BM Patil Medical College, Hospital and Research Centre, Vijaypur, Karnataka, India. The study was conducted from March 2016 to Feb 2017. Ethical clearance was taken from the Institutional Ethical Clearance Committee (ECR/383/INST/KA/2013/RR-16). After getting informed consent from each person, they underwent thorough clinical evaluation by ENT surgeon.

With anticipated least difference of means of study variable between the cases and control groups as 0.0185 and anticipated SD (standard deviation) as 0.0236, with 90% power and 5% level of significance, the minimum sample of 43 under each group was calculated.

Collection of Normal Voice Samples

A total of 43 normal voices were collected from the database of young healthy adults between age of 18 to 28 years, of which 23 were males and 20 were females. Any person having history of smoking or any other pathological conditions which made them unfit for the study were excluded.

Collection of Abnormal (Pathological) Voices

Forty-three patients (23 males and 20 females) between 18 to 55 years presenting with hoarseness and lesions of larynx like vocal nodules, vocal polyps, chronic laryngitis, voice abuse, early glottis carcinoma etc., were selected during the study.

Recording of Voices

Voices were recorded in a sound treated chamber using a unidirectional microphone (Sony Audio technical 250*L). The sustained /a/ vowel signal was recorded for three seconds, using PRAAT version 5.4.04 [8]. The intensity was controlled using VU meter built into PRAAT. From the recorded sample voice of three seconds duration, sampling frequency of 44100 Hz, spectrograph was extracted and the more stable and uniform one second was selected and saved as a sound file. A PRAAT script was written which can batch process all files into a folder to extract four parameters, Jitter (ddp), Shimmer (dda), Harmonic Noise Ratio (HNR) and Median pitch.

STATISTICAL ANALYSIS

These parameters were analysed for different variations like Mean, standard error of mean and Kolmogorov-Smirnov ZH values. Median Pitch, Jitter, Shimmer, and HNR between normal and abnormal voice groups of males and females were compared. The level of significance was set at 5% ($p < 0.05$) with SPSS Version 23.

RESULTS

The values of Mean, standard error of mean, and Kolmogorov-Smirnov ZH values of Median Pitch, Jitter, Shimmer, and HNR between normal and abnormal voice groups of males are given. The Jitter had a p-value of 0.026 which was significant but Median pitch, HNR and shimmer did not show the significance [Table/Fig-1].

When the Mean, standard error of mean, and Kolmogorov-Smirnov ZH values of Median Pitch, Jitter, Shimmer, and HNR of females were compared between normal and abnormal voice groups, the shimmer had a p-value of 0.035 which was significant but Jitter, Median pitch and HNR did not show the significance [Table/Fig-2].

DISCUSSION

Voice is a multidimensional measure which requires various methods for its assessment like Perceptual measures, acoustic

analysis, aerodynamic measures, video-stroboscopy, non-linear dynamic measures etc. The objective of the present study was to compare voice parameters of abnormal voices with normal voices. This study also shows that with the availability of freely downloadable software like PRAAT, ENT surgeons and Speech pathologists can easily use the software to analyse the voice effectively.

As early as 1987, Baken RJ had given details about few valid techniques for acoustic assessment of vocal dysfunction which are easily accomplished with current instrumentation [9]. Leiberman P recorded 23 voices from the speakers who had pathologic growths on their vocal cords. It was found that they had larger perturbations than did normal speakers with the same median fundamental periods. This may be related to size of pathological growth [10]. This showed that study of perturbation measures is important for assessment of voice. Hence voice parameters like Jitter (Leiberman P), shimmer (Horii Y), and harmonics-to-noise ratio (Yumoto E et al.) were extensively studied [10-12]. A study by Hillman RE et al., demonstrated that acoustic measures alone could be highly accurate in determining the presence/absence of a voice disorder [13]. According to Eskenazi L et al., the two most useful parameters for predicting vocal quality were the Pitch Amplitude (PA) and the HNR [14].

Many authors studied both normal and pathological voices using the parameters like jitter, shimmer, and HNR. Di Nicola V et al., studied 208 subjects (148 with dysphonia and 60 normal) using computerised digital sonography [15]. HNR developed by Yumoto E et al., was analysed and found to be highly sensitive as values were different in dysphonic patients. The comparison between the average HNR recorded in those patients (1.697 dB) is significantly different from that recorded in the normal subjects (11.169 dB) ($p < 0.001$ [12].

A similar study was done for voice analysis by Bielamowicz S et al., using C Speech, Computerized Speech Laboratory, SoundScope, and a hand marking voice analysis system. Sustained vowels from 29 male and 21 female speakers with mild to severe dysphonia were used. They felt that measures of perturbation in the various analysis packages use different algorithms, provide results in different units, and often yield values for voices that violate the assumption of quasi-periodicity. As a result, poor rank order correlations between programs using similar measures of perturbation were noted [16]. Their sample size is almost same as this study.

Variables	Normal (N=23)				Abnormal (N=23)				Kolmogorov-Smirnov ZH	p-value
	Min	Max	Mean	SE of Mean	Min	Max	Mean	SE of Mean		
Median Pitch	94.41	616.32	181.59	29.34	90.07	469.49	178.64	18.21	0.74	0.649
Jitter (ddp)	0.002	0.029	0.014	0.002	0.002	0.153	0.033	0.008	1.47	0.026*
Shimmer (dda)	0.011	0.244	0.115	0.013	0.017	0.311	0.141	0.019	0.89	0.414
H/N Ratio	1.41	27.79	14.14	2.13	1.36	21.70	13.06	1.41	1.18	0.124

[Table/Fig-1]: Mean, standard error of mean, and Kolmogorov-Smirnov ZH values of Median. Pitch, Jitter, Shimmer, and H/N ratio between normal and abnormal voice groups among males.

*significant at 5% level of significance ($p < 0.05$)

Variables	Normal (N=20)				Abnormal (N=20)				Kolmogorov-Smirnov ZH	p-value
	Min	Max	Mean	SE of Mean	Min	Max	Mean	SE of Mean		
Median Pitch	100.00	516.32	216.93	23.61	146.72	249.62	202.97	6.77	1.11	0.172
Jitter (ddp)	0.002	0.040	0.014	0.002	0.004	0.119	0.031	0.007	0.95	0.329
Shimmer (dda)	0.014	0.149	0.077	0.009	0.038	0.386	0.148	0.023	1.42	0.035*
H/N Ratio	1.36	27.79	14.64	1.94	1.73	27.79	14.22	1.47	0.791	0.56

[Table/Fig-2]: Mean, standard error of mean, and Kolmogorov-Smirnov ZH values of Median Pitch, Jitter, Shimmer, and H/N ratio between normal and abnormal voice groups among females.

*significant at 5% level of significance ($p < 0.05$)

In this study Mean, standard error of mean, and Kolmogorov-Smirnov ZH values of Median Pitch, Jitter, Shimmer, and HNR were compared between normal and abnormal voice groups.

When the above parameters were compared among normal and abnormal male voices, the value of Jitter (ddp) was significant (p -value was 0.026, i.e., $p < 0.05$). Median pitch, Shimmer (dda) and HNR values were different but not significant. (p -values were 0.649, 0.414 and 0.124 respectively).

When these parameters were compared among female normal and abnormal voices, shimmer values were significant. (p -value was 0.035 i.e., $p < 0.05$).

Limitations are present for acoustic analysis when it comes to assess aperiodic voices. Titze IR discusses three types of voices depending on the periodicity of the voices. While types 1 and 2 have relatively better periodicity, the type 3 voices have aperiodic waves [17]. Type 3 voices pose problems on correctly analysing the voice parameters. In fact recently, the study by Núñez Batalla F et al., showed that there is no difference between using PRAAT and Dr. Speech for analysis of type 3 voices. Even PRAAT can analyse effectively type 3 voices [18].

Many authors felt other methods like nonlinear dynamic measures are more indicative of differentiation in such cases particularly in chaotic voices. Jacqueline BF et al., studied voices of Indian population comprising adults and elderly. They concluded that parameters like correlation dimension D2 (a voice measure of nonlinear dynamic analysis) are better assessors. The anatomical alterations in the vocal mechanism that occur for any pathological conditions result in higher values of correlation dimension. Thus, it can be considered as a useful tool in the assessment of voice. However, they also felt that the existing voice analysis techniques available to the voice clinician cannot be replaced but nonlinear measures can be added to the existing battery of tests [19].

In this study, comparison of the recorded abnormal voice parameters were done with normal voices and added to a voice database. This normal database is continuously added with voices recorded at our institution which naturally reflects the local sample population effectively. The same authors also proposed an institutional voice database, standard voice analysis tool, and method of voice measurement [7].

According to Stemple JC et al., normal database should report demographic details of their local sample populations to account for factors that influence the instrumental results such as gender, age, health history and local database [20]. Recording techniques and sample tasks may vary across the studies as well as equipment and analysis routines [21]. These factors limit the ability to compare different findings. A practical solution is to collect local norms by measuring a large group of normal speakers as a separate sample.

Di Nicola V et al., on comparison of normal and abnormal voices, felt that forensic evaluation of dysphonia needs application of strict, precise, correct sampling and analysis method following well-defined rules. Further comparison of normal and abnormal voices, HNR was analysed and found to be highly sensitive [15].

Jiang et al., examined objective acoustic analysis methods nonlinear dynamic and traditional perturbation measures like jitter and shimmer, to assess voices of patients with vocal nodules and polyps [22]. The jitter or shimmer showed no significant changes but correlation dimension, a parameter of nonlinear dynamic measure of voice showed significance. They concluded that the combination of traditional perturbation and nonlinear dynamic measures may improve our ability to provide objective clinical analysis of voices with vocal mass lesions.

The cost of these analysis programs is undoubtedly high and routine use by clinicians is not possible. However, development of free computer application with widest diffusion like PRAAT has greater capabilities of analysing acoustic signals [18]. The authors also recommend PRAAT program as valid, reliable, and easily manageable and has minimum equipment requirements.

More and more normal and abnormal voices from local area are to be added to institutional voice database. This helps in overcoming bias and adds to further research on voice assessment.

LIMITATION

The age range of normal voice study cases were 18-28 years whereas that of abnormal voice were between 18-55 years. This happened because abnormal voices were more common in older age group. The sample size is small. Hence, age difference was not taken into consideration. Though the collected sample size was 63 abnormal voices (32 males and 31 females) only 43 voice files could be assessed by the program. Remaining 20 voice files became corrupted and were not analysed by PRAAT Program. The future studies are to be done on a relatively larger number of subjects from a variety of dysphonia population so that changes in parameters in each specific pathology can be defined.

CONCLUSION

Acoustic voice analysis is still a valuable technique which enables voice clinicians to compare voices to differentiate them into normal and abnormal. But this requires a robust normal database of the local and regional demographic voice samples as the recording techniques, equipments, age, sex and analysis programs differ from one another. This method can provide a non-invasive and objective tool to identify and document abnormal voices.

REFERENCES

- [1] Davis SB. Acoustic characteristics of normal and pathological voices. Elsevier. 1979;1:271-335. Available at <https://doi.org/10.1016/B978-0-12-608601-0.50010-3>.
- [2] Moore GP. Terminal Report for a conference and early detection of laryngeal pathology, Gainesville; University of Florida, Department of Speech, Communication Sciences Laboratories. 1973.
- [3] Teixeira JP, Oliveira C, Lopes C. Vocal acoustic analysis-jitter, shimmer and HNR parameters. *Procedia Technology*. Elsevier. 2013;9:1112-22.
- [4] Guimaraes, I. A. *Ciencia e a Arte da Voz Humana*. Escola Superior de saude de Alcoitao, 2007*.
- [5] Boersma P. Accurate short-term analysis of the fundamental frequency and the harmonics to noise ratio of a sampled sound. *IFA proceedings*. 1993;17:97-110.
- [6] Bonzi EV, Grad GB, Maggi AM, Munoz MR. Study of the characteristic parameters of normal voices of Argentinian speakers. *Papers in Physics*. 2014;6:art. 060002.
- [7] Lathadevi HT, Malipatil SR, Guggarigoudar SP. Creation of voice database, acoustic analysis and standardisation of normal Indian voices. *Int J Pharma & Biosciences*. 2017;8(2):B349-55.
- [8] Paul Boerasma, David weenick. Praat : Doing phonetics by computer [Home page internet] No date. Available from <http://www.fon.hum.uva.nl/praat/>
- [9] Baken RJ, Orrikloff RF. *Clinical examination of speech and voice*. San Diego. Singular Thomson Learning ; 2000.
- [10] Leiberman P. Some acoustic measures of fundamental periodicity of normal and pathologic larynges. *The Journal of the Acoustical Society of America*. 1963;35:344. <https://doi.org/10.1121/1.1918465>
- [11] Horii Y. Fundamental frequency perturbation observed in sustained phonation. *J of Speech and Hearing Res*. 1979;22(1):5-19.
- [12] Yumoto E, Sasaki Y, Okamura H. Harmonics-to-noise ratio and psychophysical measurement of the degree of hoarseness. *J Speech Hear Res*. 1984;27(1):2-6.
- [13] Hillman RE, Montgomery WW, Zeitels SM. Appropriate use of objective measures of vocal function in the multidisciplinary management of voice disorders. *Current Opinion in Otolaryngology & Head and Neck Surgery*. 1997;5:172-75.
- [14] Eskenazi L, Childers DG, Hicks DM. Acoustic correlates of voice quality. *J Speech Hear Res*. 1990;33(2):298-306.
- [15] Di Nicola V, Fiorella ML, Luperto P, Staffieri A, Fiorella R. Objective evaluation of dysphonia. Possibilities and limitations. *Acta Otorhinolaryngol Ital*. 2001;21(1):10-21.

- [16] Bielamowicz S, Kreiman J, Gerratt BR, Dauer MS, Berke GS. Comparison of voice analysis systems for perturbation measurement. *J Speech Hear Res.* 1996;39(1):126-34.
- [17] Titze IR. Workshop on acoustic voice analysis: Summary statement. National Centre for Voice and Speech. 1995.
- [18] Núñez Batalla F, González Márquez R, Peláez González MB, González Laborda I, Fernández Fernández M, Morato Galán M. Acoustic voice analysis using the praat program: comparative study with the dr. speech program. *Acta Otorrinolaringol Esp.* 2014;65(3):170-76.
- [19] Jacqueline BF, Balasubramaniam RK, Pitchaimuthu AN, Bhat JS. Nonlinear dynamic analysis of voice: A normative study in the Indian population. *Int J Med Res & Health Sci.* 2014;3(1):128-32.
- [20] Stemple JC, Roy N, Klaben B. *Clinical voice pathology: Theory and management*, 5th edn. Plural Publishing, 28-Jan-2014-468.
- [21] Vogel AP, Morgan AT. Factors affecting the quality of sound recording for speech and voice analysis. *International Journal of Speech-Language Pathology.* 2009;11(6):431-37.
- [22] Jiang JJ, Zhang Y, MacCallum J, Sprecher A, Zhou L. Objective acoustic analysis of pathological voices from patients with vocal nodules and polyps. *Folia Phoniatr Log op.* 2009;61(6):342-49.

PARTICULARS OF CONTRIBUTORS:

1. Professor, Department of ENT, BLDE University's (DT) Sri BM Patil's Medical College, Hospital and Research Centre, Vijayapur, Karnataka, India.
2. Professor, Department of ENT, BLDE University's (DT) Sri BM Patil's Medical College, Hospital and Research Centre, Vijayapur, Karnataka, India.

NAME, ADDRESS, E-MAIL ID OF THE CORRESPONDING AUTHOR:

Dr. HT Lathadevi,
Professor, Department of ENT, BLDE University's (DT) Sri BM Patil's Medical College, Hospital and Research Centre,
Vijayapur-586103, Karnataka, India.
E-mail: lathadevi45@gmail.com

Date of Submission: **Apr 14, 2018**Date of Peer Review: **May 08, 2018**Date of Acceptance: **Sep 22, 2018**Date of Publishing: **Dec 01, 2018****FINANCIAL OR OTHER COMPETING INTERESTS:** None.