Effectuation of Acoustic Characterisics in Young Adults

Prarthana Karunaimathi V., Gladis D., Balakrishnan D.

Abstract: ***Each human being has a voice that is distinct and unique. Extraction and analysis of the features of a person's voice plays a vital role in diagnosis of diseases in the recent years. The basic parameters used for the voice analysis are Fundamental Frequency (F0), Jitter, Shimmer and Harmonics-to-Noise Ratio. Most of the prevailing Voice Analysis Software are highly commercial with complicated User Interface, requiring expertise to operate them. Hence the proposed study is carried out, which led to the development of a Voice Analysis Application called 'Ephphatha'. A novel algorithm with improvised measures for voice analysis is used in Ephphatha. The acoustic parameters of duly collected voice samples were analyzed by Ephphatha and compared with the existing algorithms like PRAAT and Multi-Dimensional Voice Program (MDVP) of CSL(Computerized Speech Lab), with F-test analysis. The result showed no significant difference in F0 (p>0.05), except for its Standard Deviation values (p<0.001). The Ratio Of Variance (ROV) is less than or equal to 1 for F0. The other parameters are not significantly correlated. Also it showed that the shimmer values are moderately correlated (p<0.05), when Ephphatha is compared with CSL. It was also seen that all the parameters of PRAAT and Ephphatha have no significance difference (P > 0.05), except for the Shimmer values (p<0.001). Thus Ephphatha would undoubtedly equip the experienced clinicians with its novel approach and better User Interface.***

Keywords: ***Fundamental Frequency, Jitter, Shimmer, Harmonics-to-Noise Ratio.***

I. INTRODUCTION

Voice disorders which are most commonly caused by the vocal folds and the laryngeal musculature require voice analysis for providing objective evidence in perceptual diagnosis. Dynamic analysis of the vocal folds and their movement is physically very difficult. Less invasive methods like imaging is not effective due to the cartilaginous structure of the vocal cord. Indirect methods like inverse filtering (IF) and electroglottography (EGG) are currently in use. IF estimates the glottal flow and indirectly gives the measure of laryngeal behavior or reflects the movement of the vocal fold. EGG gives the information of the closure of the vocal folds by measuring the electrical resistance between two electrodes placed on either side of the subject's throat at the level of vocal fold. Neither inverse filtering nor EGG is sufficient to completely describe the complex pattern of vocal fold movement.

Hence Voice Analysis process is vital, which is carried out by examining the voice characteristics, which include phonation, pitch, loudness, and rate the speech pathologists to monitor and give the right course

Voice disorders include laryngitis, vocal cord paresis and Spasmodic Dysphonia. Analysis of these voice characteristics plays a major role in studying and treating the progressiveness of the neurological disorders such as Stroke [1], Parkinson's disease [2], multiple sclerosis [3], myasthenia gravis [4] and Amyotrophic lateral sclerosis [5] and also equips and guides of treatment according to the severity of the impairments such as loss of hearing [6][7][8], autism [9] and down syndrome [10].

Several commercial systems are available for acoustic voice analysis, which are designed to perform complex manipulations of speech signals. The signals in the form of waveforms are processed and manipulated in order to extract the voice parameters. In general, values derived from most of the existing programs include: (i) *fundamental frequency* of voice which gives the patient's habitual pitch, (ii) *jitter* which gives the value of the degree of pitch instability or perturbations in the speed of vocal fold vibrations, (iii) *shimmer* which shows the degree of loudness instability or perturbations in the amplitude of vocal fold vibrations, and (iv) *harmonic-to-noise ratio*, which shows the overall amount of noise in the voice signal. These values are used to quantify the levels of voice impairment by comparing them with their respective standardized values. Voice Analysis, when combined with perceptual diagnosis, provides objective evidence and hence can improve the treatment of patients. Also these data play a vital role in providing the benefits of alternative treatments for different types of vocal pathologies [11].

II. MATERIALS AND METHOD

The data for the analysis includes voice signals, which were obtained by recording using the Computerized Speech Lab (CSL) with a microphone placed at a distance of 10 cm from the volunteer's lips in a sound-attenuated room. The respondents included 27 female and 24 male volunteers, with normal voices belonging to the age group of 18 – 25 years. Sustained vowel phonation //a// was recorded for each subject in their comfortable pitch without any break for a minimum of 4 seconds. The voice signals were sampled at 44.1 kHz with 16 bits of resolution.

Manuscript published on 30 September 2019
* Correspondence Author
V. Prarthana Karunaimathi*, Department of Computer Science, Presidency College, University of Madras, Chennai, Tamil Nadu, India. Email: prarthanakamal@gmail.com
D. Gladis, Principal, Bharathi Women's College, University of Madras, Chennai, Tamil Nadu, India. Email: gladischristopher@gmail.com
D. Balakrishnan, Head of the Department of Audiology and Speech, SRM Medical College, Chennai, Tamil Nadu, India. Email: balaent@gmail.com

The data collected were then analyzed with MDVP (Multi-Dimensional Voice Program) module of CSL, KayPentax, USA, PRAAT and the newly developed system Ephphatha.

III. ACOUSTIC ANALYSIS PROGRAMS

Acoustic vocal assessment consists of a noninvasive process of obtaining objective measures from a voice signal. The exceptional accuracy of these measures allows evaluators to recognize voice alterations quite early, thereby helping them to compare the efficiency and outcome of other vocal techniques and therapeutic procedures.

A. Ephphatha

Commonly used vocal acoustic parameters such as Fundamental Frequency (F0), Jitter, Shimmer, and Harmonics to noise ratio (HNR) were estimated using Ephphatha. The aim of this application is to increase the accuracy of the acoustic measurements and in turn to equip the physicians make a better decision. Most of the existing systems are highly commercial and can be used only by experts. This motivated the proposed work and development of the application Ephphatha with the following acoustic measures:

- *Fundamental Frequency (F0):* The main technique used by Ephphatha is Normalized Cross-Correlation Function (NCCF) [12], which is used to estimate the fundamental frequency. It is designed to work with any sampling frequency and frame rate over a wide range of possible F0 and different noise conditions. The computational complexity is slightly increased in NCCF, in order to overcome the shortcomings of other candidate generators like auto-correlation. To reduce the computational load, the NCCF is used in two stages. In the first stage, the signal with low sample rate is used to generate the set of candidates, whereas high sample-rate signal is used in the second stage. The NCCF of the low sample rate signal is computed for all the lags in the F0 range and the locations of the local maxima are recorded in the first pass. The NCCF of the high-sample rate signal is computed only in the vicinity of the promising peaks found in the first pass. It then searches for the local maxima in the refined NCCF to obtain improved peak location and amplitude estimates. Finally, dynamic programming is used to select the set of NCCF peaks or unvoiced hypotheses across all frames. NCCF function is shown in equation (1)

$$\text{NCCF (m)} = \frac{\sum_{n=0}^{N-m-1} x(n).x(n+m)}{\sqrt{\sum_{n=0}^{N-m-1} x^2(n).\sum_{n=0}^{N-m-1} x^2(n+m)}} \qquad (1)$$

  where, N is the analyzed frame length, and m is a lag; M0 is the number of autocorrelation points to be computed. The pitch period peaks are more prominent than autocorrelation.
- *Jitter Calculation:* Jitter, which is also known as frequency perturbation is the cycle-to-cycle variation of fundamental frequency. It is a well known parameter to investigate minute changes in the vocal folds. It represents the average absolute difference between two consecutive periods [13] as shown in equation (2):

$$\text{Jitter Absolute} = \frac{1}{N-1}\sum_{i-1}^{N-1}|P_i - P_{i-1}| \qquad (2)$$

  Jitter Percentage is calculated by dividing the Jitter Absolute by the average period as shown in equation (3):

$$\text{Jitter Percent} = \frac{Jitter\ Absolute}{\frac{1}{N}\sum_{i-1}^{N} P_i} \qquad (3)$$

High Jitter is caused by various conditions that affect the vocal cords, including nodules, polyps, and weakness of the laryngeal muscles. Few studies show that the value of the jitter is directly dependent on the size of the vocal polyps [14].

- *Shimmer Calculation:* Shimmer (dB) is expressed as the variability of the peak-to-peak amplitude in decibels, which is also called as amplitude perturbation. Shimmer value is formulated as the average absolute difference in amplitude between adjacent periods [15] as shown in equation (4)

$$\text{Shimmer (dB)} = \frac{1}{N-1}\sum_{i=1}^{N-1}\left|20log\left({A_{i+1}}/{A_i}\right)\right| \qquad (4)$$

where, A_i are the extracted peak-to-peak amplitude data and N is the number of extracted fundamental frequency periods. Shimmer (relative) is defined as the average absolute difference between the amplitudes of consecutive periods, divided by the average amplitude, expressed as a percentage as shown in equation (5):

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

- *Hormonicity:* Noise is an aperiodic energy of a signal. Normal voices have low levels of noise, whereas abnormal voices show higher noise levels. Sustained vowel phonation in this case, can be assumed to be periodic and the signal-to-noise ratio (SNR) will be equal to the harmonics-to-noise ratio (HNR). The Harmonics-to-Noise (HNR) ratio is a measure of the proportion of harmonic sound to noise in the voice measured in decibels (dB). It quantifies the amount of additive noise in the voice signal. HNR is indirectly dependent on the voice of a person and it acts as an index to show the level of hoarseness and vocal aging [16]. It is also useful in quantitatively assessing the results of treatment for hoarseness and vocal fold disorders [17]. The estimation of HNR is shown in equation (6).

$$HNR = 10 * \log_{10} \frac{AC_v(T)}{AC_v(0) - AC_v(T)} \qquad (6)$$

It is used to estimate the level of noise in human voice signals. It can be estimated in two ways: (i) on a time-domain basis, in which HNR is computed directly from the acoustic waveform; and (ii) on a frequency-domain basis, in which HNR is computed from a transformed representation of the waveform.

B. CSL

The CSL (Computerized Speech Lab) tool is compatible with several voice analysis applications including MDVP, which is a standard software for the complete analysis of voices [18]. It is widely used by researchers and leading clinicians all over the world. It provides a robust multi-dimensional analysis of voice with graphic and numerical representation of analysis results. The sustained vowel phonations were captured and completely analyzed by choosing the MDVP option in CSL main window.

In MDVP, six windows consisting of voice analysis results including the radial graph of the parameters were used. In this study, the following parameters of MDVP are used:

- *Fhi (Highest Fundamental Frequency [Hz])*: Maximum of all extracted period-to-period fundamental frequency values. Voice break areas are excluded. Fhi and $F0_{(i)}$ are shown in equation (6) and (7) respectively.

$$\text{Fhi} = \max(F0_{(i)}) \qquad (6)$$

where, F0 → Fundamental Frequency ; i=1,2….N

$$F0_{(i)} = \frac{1}{P0_{(i)}} \qquad (7)$$

where, P0 → Pitch Period; i=1,2….N

N is the number of extracted pitch periods

- *Flo (Lowest Fundamental Frequency [Hz]):* Minimum of all extracted period-to-period fundamental frequency values which is shown in equation (8). Voice break areas are excluded.

$$\text{Flo} = \min(F0_{(i)}) \qquad (8)$$

- *F0 (Average Fundamental Frequency [Hz]):* It gives the average value of all extracted period-to-period fundamental frequency values which is shown in equation (9). Voice break areas are excluded.

$$\text{Flo} = \text{avg}(F0_{(i)}) \qquad (9)$$

- *Jita (Jitter Absolute [usec]:* It gives an evaluation of the period-to-period variability of the pitch period within the analyzed voice sample. Voice break areas were excluded. It was computed from the extracted period-to-period pitch data. It is an absolute measure and an average fundamental frequency dependent parameter. The pitch may vary for several reasons and hence a minute change in the pitch value will affect the jitter values.
- *Jitt (Jitter Percent [%]):* It gives the relative evaluation of the period-to-period (very short-term) variability of the pitch within the analyzed voice sample. Voice break areas are excluded. It is a relative measure and hence the influence of the fundamental frequency value is comparatively less than the Jita.
- *NHR (Noise-to-Harmonics Ratio):* It gives the average ratio of the inharmonic spectral energy to the harmonic spectral energy in the frequency range 70-4200 Hz. This is a common parameter used to evaluate the noise present in the analyzed signal. It was computed using a pitch-synchronous frequency-domain method. It is a global measure used to denote the spectral noise and it can be highly affected by the amplitude and frequency variation.
- *Shim (Shimmer Percent [%])*: It gives the relative evaluation of the period-to-period (very short term) variability of the peak-to-peak amplitude within the analyzed voice sample. Voice break areas are excluded. It was computed from the extracted peak-to-peak amplitude data.
- *ShdB (Shimmer in dB)*: It gives an evaluation of the period-to-period variability of the peak-to-peak amplitude of the signal in dB. Both Shim and ShdB were very sensitive to the amplitude variation between consecutive periods and hence both were influenced by the same.
- *STD (Standard Deviation of the Fundamental Frequency [Hz]):* It gives the standard deviation of all the extracted period-to-period fundamental frequency values which is shown in equation (10). Voice break areas are excluded.

$$\text{STD} = \sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(F0 - F0^{(i)})^2} \qquad (10)$$

C. PRAAT

PRAAT is a tool for phonetic speech analysis developed by Paul Boersma [19] and David Weenink in the Institute of Phonetic Sciences of the University of Amsterdam. It can be executed in various operating systems. It is open source software that is continually updated to accommodate new developments and new analysis methods.

It provides graphical visualization of sound such as waveform, spectrogram, pitch contour, amplitude curve and formant tracks. The main pitch analysis algorithms available are auto-correlation and cross-correlation, in which one of them can be selected. The default algorithm is based on the accurate auto-correlation function. The signal is divided into frames and F0 is estimated for each frame. PRAAT then normalizes the autocorrelation of the signal by dividing the autocorrelation of the window function. PRAAT uses post-processing to reduce large changes in successive estimates. The estimated parameters can be obtained by viewing the voice report for the selected portion of the signal. It consists of quantitative information of pulses, voicing, F0, Jitter, Shimmer and Harmonicity of the voiced parts. In PRAAT, vowel phonation /a/ or /i/ should have a HNR of 20 for healthy voices, and 40 for the phonation /u/. Hence, a HNR < 20 is considered to be a measure of noticeable hoarseness [20].

IV. RESULTS AND DISCUSSION

The parameters were extracted for a segment of signal in Ephphatha and the same segment was analyzed in MDVP in order to correlate the corresponding variables in both systems. Statistical analyses of the parameters were done using R programming.

Table-I contains statistical analysis results for the voice parameters extracted using Ephphatha and CSL. The results obtained using f-test shows that there is no significance difference in all the fundamental frequency (F0) values ($p>0.05$) except for its standard deviation values ($p<0.001$) and the ratio of variance is closer to 1 for F0. The range F0 standard deviation value of CSL (2.815±1.31) is relatively high when compared to that of Ephphatha (1.596±0.45). The other parameters were not significantly correlated. The Jitter values and the NHR values are not comparable. Also it shows that the shimmer values are moderately correlated ($p<0.05$). In case of gender voice comparison, the correlations between the variables are almost similar for all the parameters except for the standard deviation of fundamental frequency. Ratio of Variance is slightly higher in Female voice than the Male voice. The deviation is lesser in Male voice than in Female voice.

Table- I: F- test results of Voice Parameters extracted using CSL Vs Ephphatha (Age 18-25 years

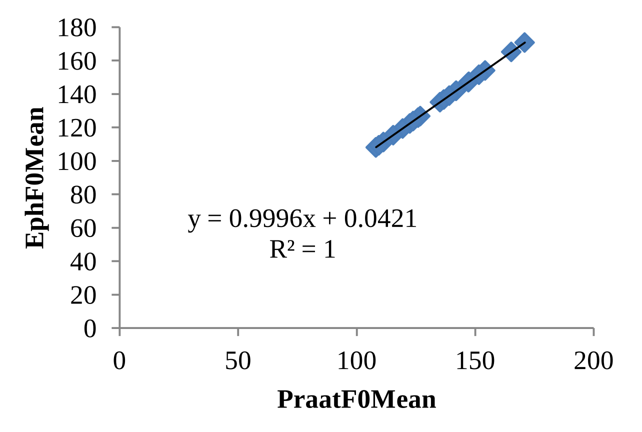
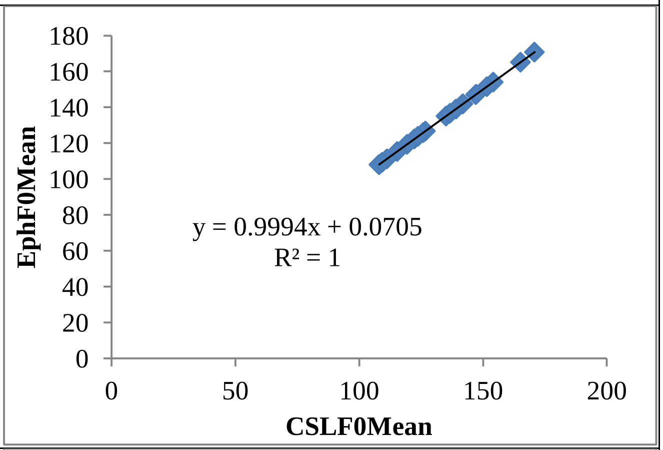
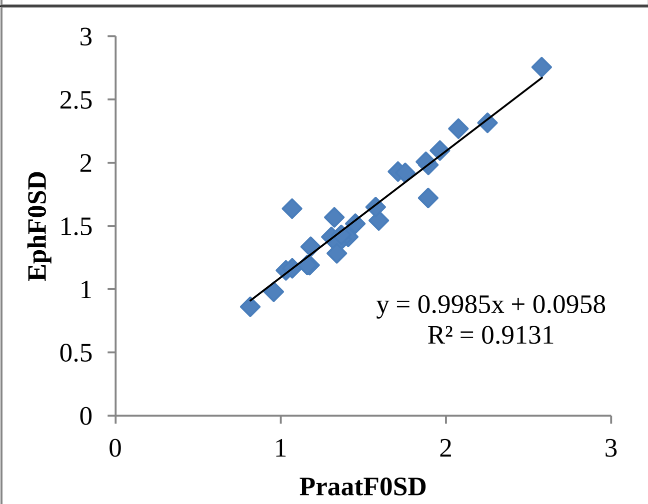
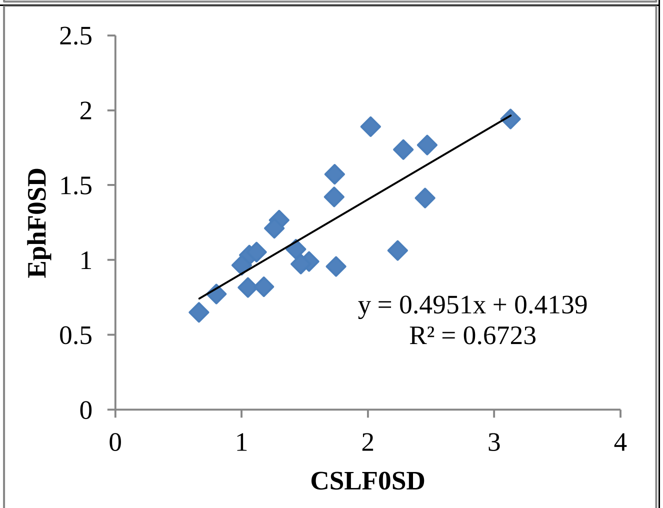
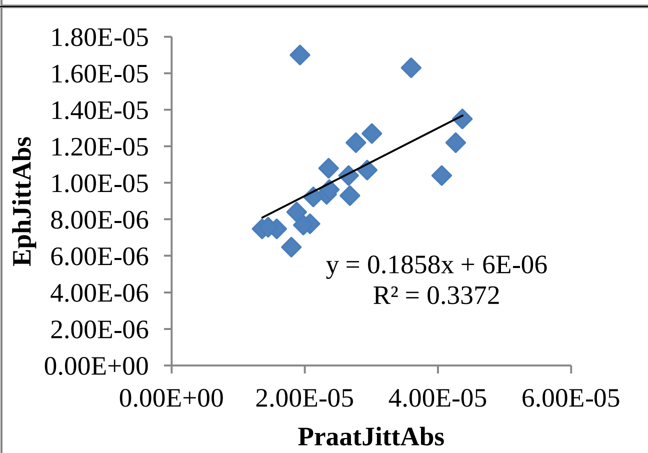
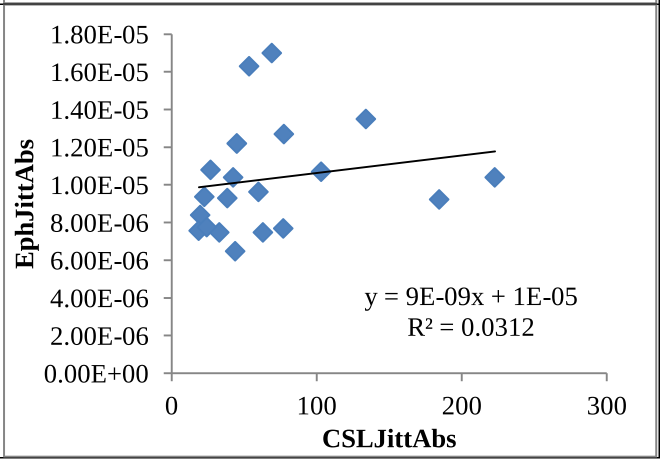
Parameters	Programs	Female			Male		
		Mean±SD	*P-value*	*ROV*	*Mean±SD*	*P-value*	*ROV*
F0Mean	CSL	232.6±17.77	0.9887	1.005602	133.4±18.62	0.9978	1.001268
	Ephphatha	232.6±17.77	>0.05		133.4±18.61	>0.05	
F0SD	CSL	2.815±1.31	4.66E-07	8.695512	1.605±0.64	0.02897	2.742755
	Ephphatha	1.596±0.45	<0.001		1.2086±0.38	>0.01	
F0Max	CSL	243.8±19.84	0.6365	1.206001	141.2±21.41	0.5605	1.302087
	Ephphatha	236.9±18.06	>0.05		136.7±18.76	>0.05	
F0Min	CSL	222.7±16.86	0.9529	1.02382	127.69±19.6	0.7981	1.122862
	Ephphatha	228.2±17.06	>0.05		130.2±18.5	>0.05	
JittAbs	CSL	50.99±34.25	2.20E-16	4.54E+13	66.97±54.02	< 2.2e-16	3.58E+14
	Ephphatha	1.52E-05 ±5.09E-06	<0.001		1.03E-05±2.85E-06	<0.001	
Jitter%	CSL	1.186±0.8	2.20E-16	73.6377	0.8477±0.58	3.70E-11	35.30157
	Ephphatha	0.2878±0.09	<0.001		0.3455±0.1	<0.001	
Shimm%	CSL	4.161±1.07	0.1328	1.821518	3.121±1.23	0.3019	1.599325
	Ephphatha	4.874±1.44	<0.05		3.773±1.55	>0.05	
ShimmdB	CSL	0.3633±0.09	0.1521	1.823965	0.272±0.11	0.2976	1.606184
	Ephphatha	0.4256±0.12	>0.05		0.33±0.14	>0.05	
NHR	CSL	0.1234 ±0.14	8.16E-06	6.567055	0.1362±0.02	6.16E-05	6.951473
	Ephphatha	0.0119±0.01	<0.001		0.0127±0.01	<0.001	

Table-II shows that all the parameters of PRAAT and Ephphatha have no significance difference (P > 0.05) between almost all the parameters, except for the shimmer values (p<0.001). The shimmer is an amplitude dependent variable. There may be minute instability in the amplitude perturbation measure during sustained vowel phonation of the normal voice. The result in Tabl-II shows that the variation in the newly developed algorithm is slightly higher (4.874±1.44, 0.425±0.12) than the existing one (3.123±0.77, 0.273±0.07) for both the shimmer measures. The right tailed f-test is preferred, which gives the ratio of variance >3 for the shimmer values. The ratio of variance for Jitter absolute is slightly higher for males when compared to female voice. On the other hand, Shimmer values of female and male voice are moderately and positively correlated.

Table- II. F- test results of Voice Parameters extracted using PRAAT Vs Ephphatha (Age 18-25 years)

Parameters	Programs	Female			Male		
		Mean±SD	*P-value*	*ROV*	*Mean±SD*	*P-value*	*ROV*
F0Mean	PRAAT	232.6±17.72	0.9888	1.005579	133.4±18.61	0.9986	1.000774
	Ephphatha	232.6±17.77	>0.05		133.4±18.61	>0.05	
F0SD	PRAAT	1.502±0.42	0.8048	1.091794	1.156±0.4	0.894	1.062203
	Ephphatha	1.596±0.45	>0.05		1.2086±0.38	>0.05	
F0Max	PRAAT	236.3±18.13	0.9858	1.007051	136.4±18.79	0.995	1.002824
	Ephphatha	236.9±18.06	>0.05		136.7±18.76	>0.05	
F0Min	PRAAT	228.5±17.14	0.9803	1.009807	130.5±18.49	0.9997	1.000165
	Ephphatha	228.2±17.06	>0.05		130.2±18.5	>0.05	
JittAbs	PRAAT	1.4E-05±5.69E-06	0.5713	1.251595	2.55E-05±8.91E-06	4.01E-06	9.76E+00
	Ephphatha	1.52E-05±5.09E-06	>0.05		1.03E-05±2.85E-06	<0.001	
Jitter%	PRAAT	0.3285±0.14	0.05944	2.127748	0.3321±0.09	0.8718	1.075842
	Ephphatha	0.2878±0.09	>0.05		0.3455±0.1	>0.05	

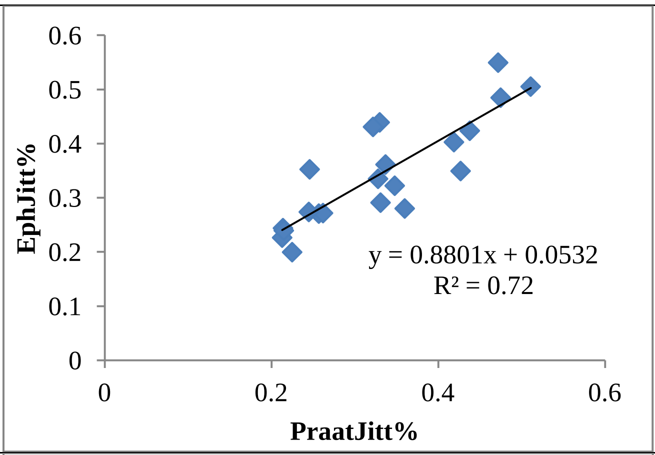
Shimm%	PRAAT	3.123±0.77	0.00205	3.521176	2.476±1.46	0.7947	1.125101
	Ephphatha	4.874±1.44	>0.001		3.773±1.55	>0.05	
ShimmdB	PRAAT	0.273±0.07	0.00364	3.264505	0.2181±0.13	0.8402	1.095654
	Ephphatha	0.425±0.12	>0.001		0.33±0.14	>0.05	
NHR	PRAAT	20.48±2.47	0.441	1.357639	22.3±3.66	0.6205	1.251599
	Ephphatha	19.96±2.12	>0.05		20.84±3.36	>0.05	

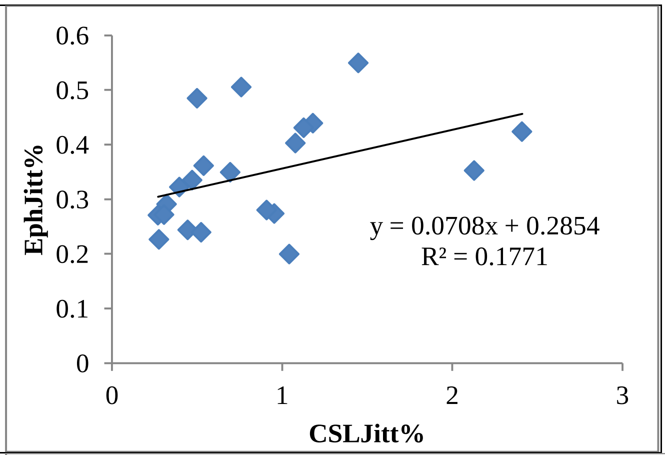
<table>
<tr><th>Parameters</th><th>Praat Vs Ephphatha</th><th>CSL Vs Ephphatha</th></tr>
<tr><td>F0Mean</td><td>
</td><td>
</td></tr>
<tr><td>F0 SD</td><td>
</td><td>
</td></tr>
<tr><td>Jitter</td><td>
</td><td>
</td></tr>
</table>

Retrieval Number: C6039098319/2019©BEIESP
DOI:10.35940/ijrte.C6039.098319
Journal Website: www.ijrte.org

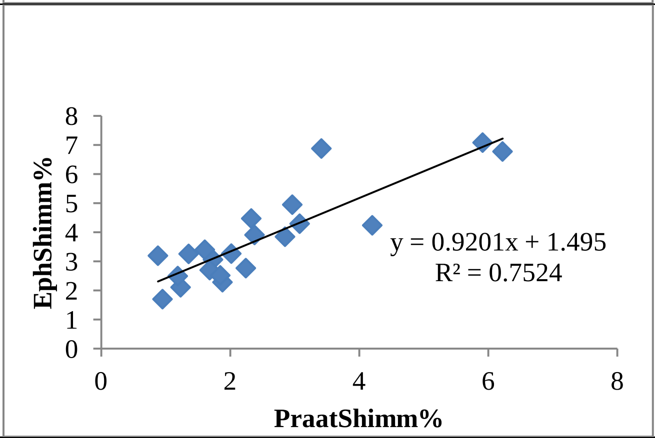
Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

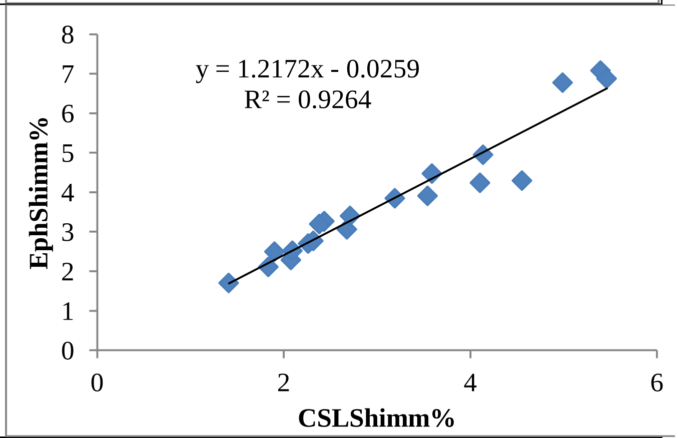
Jitter %



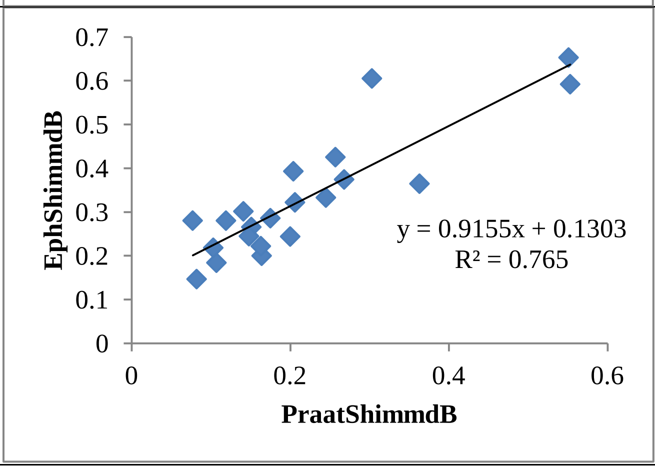


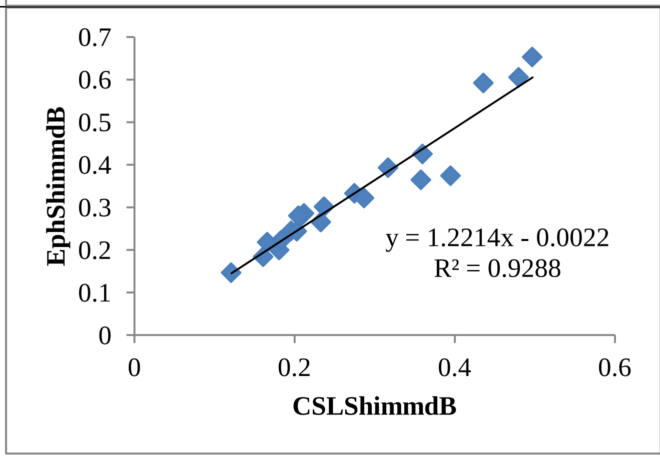
Shimmer %





Shimmer (dB)





Retrieval Number: C6039098319/2019©BEIESP
DOI:10.35940/ijrte.C6039.098319
Journal Website: www.ijrte.org

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

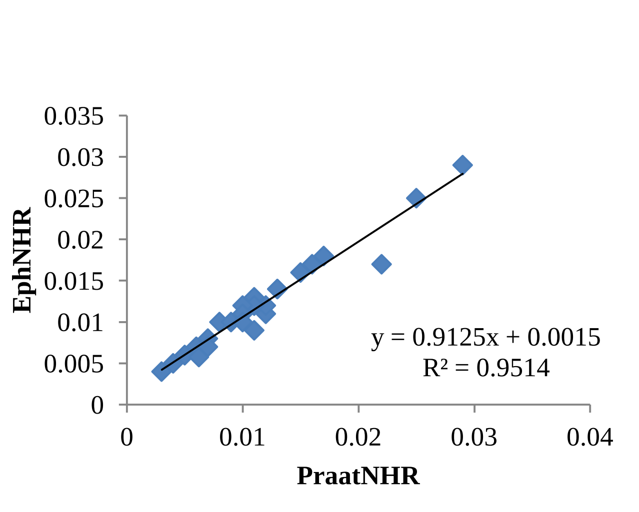
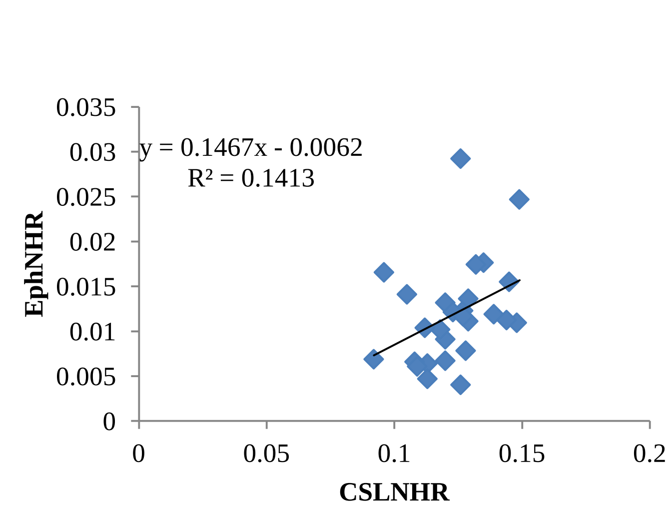
NHR		

Fig. 1. Correlation between the parameters, a comparison with Ephphatha

Fig.1 shows the relationship between the voice parameters determined using PRAAT, CSL and Ephphatha. R-Squared or coefficient of determination (R^2) gives the goodness of fit in the above regression model. R-Squared is given as (11)

$$\text{R-Squared} = 1 - \frac{SS_{regression}}{SS_{total}} \qquad (11)$$

where, $SS_{regression}$ is the sum of squares due to regression and SS_{total} is the total sum of squares. The parameter F0 Mean values are perfectly and positively correlated both in PRAAT and CSL, when compared with Ephphatha, since their R-Squared value is 1. In case of F0SD measure, it is noted that the values obtained in Praat is strongly correlated with Ephphatha than the values obtained in CSL since the deviation in pitch sigma is higher in CSL than in PRAAT which is shown in Fig. 2. Jitter values are weakly correlated in CSL, since its R^2 is 0.03, which is closer to 0. Also the regression line is almost parallel to the x-axis. Hence this measure is incomparable. On the other hand, the coefficient of determination values of Jitter absolute and Jitter % of PRAAT are 0.34 and 0.72 respectively, when it is compared with Ephphatha. Shimmer values of CSL are positively and strongly correlated with Ephphatha than the values of PRAAT. The coefficient of determination values of Shimmer % and Shimmer(dB) values of CSL are closer to 1. That is, the shimmer values of CSL changes with corresponding shimmer values of Ephphatha. But in case of PRAAT, the shimmer values are moderately correlated with the shimmer values of Ephphatha. The Noise to harmonic ratio values of both PRAAT and CSL are positively correlated. But in CSL, NHR values are weakly correlated with NHR values of Ephphatha.

Fig. 2 shows that the deviation of pitch sigma is higher in CSL than the values of PRAAT program and Ephphatha. The deviation can be reduced by using a post-processing filter. The post-processing method namely de-step filter was applied to the F0 contour in the developed application in order to remove pitch doubling and halving errors [21]. Thus it is clear that adding an additional post-processing filter to the raw pitch contour improves the performance of the algorithm.

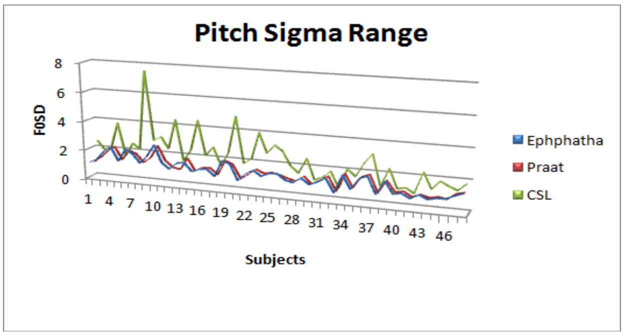


Fig. 2. Standard Deviation of Fundamental Frequency of the three algorithms

V. Conclusion

The voice parameters were extracted from the voice samples of 50 subjects with perceptually normal subjects, using the newly developed algorithm viz. Ephphatha. The values extracted from the existing software CSL and PRAAT were compared with the same values, extracted with Ephphatha. The statistical analysis of F_0 values showed no significant difference between the values in all the three algorithms. However, the range of F_0 as calculated in the three different algorithms was found to be widely different. The shimmer values are moderately correlated when Ephphatha is compared with CSL. The jitter values of CSL are weakly correlated with the jitter values of Ephphatha, that is, it was incomparable. All the parameters of PRAAT and Ephphatha have no significance difference except for the shimmer values. The observation shows that the existing software PRAAT is positively correlated with the proposed system. As various literature shows, the values are needed to be standardized not only due to the differences in the systems but also the nature of the data.

ACKNOWLEDGMENT

The authors would like to thank the Department of Audiology and Speech Language Pathology, SRM University for their support and for providing them with voice samples for this study.

REFERENNCE

1. Guranski K, Podemski R. “Emotional prosody expression in acoustic analysis in patients with right hemisphere ischemic stroke.” *Neurologia i neurochirurgia polska*. 2015;49(2):113-20.
2. J. Holmes R, M. Oates J, J. Phyland D, J. Hughes A. “Voice characteristics in the progression of Parkinson's disease.”, *International Journal of Language & Communication Disorders*, 2000 Jul 9, 35(3), 407-18.
3. Vizza P, Mirarchi D, Tradigo G, Redavide M, Bossio RB, Veltri P. “Vocal signal analysis in patients affected by Multiple Sclerosis.”, *Procedia Computer Science*. 2017 Jan 1;108:1205-14.
4. Konstantopoulos K, Christou YP, Vogazianos P, Zamba-Papanicolaou E, Kleopa KA. “A quantitative method for the assessment of dysarthrophonia in myasthenia gravis.”, *Journal of the neurological sciences*. 2017 Jun 15;377:42-6.
5. Wang J, Kothalkar PV, Cao B, Heitzman D. “Towards Automatic Detection of Amyotrophic Lateral Sclerosis from Speech Acoustic and Articulatory Samples.” Interspeech 2016 Sep 12 (pp. 1195-1199).
6. Coelho AC, Medved DM, Brasolotto AG. “Hearing loss and the voice.”, *InUpdate On Hearing Loss*, 2015 Dec 2. IntechOpen.
7. Higgins MB, Carney AE, Schulte L. “Physiological assessment of speech and voice production of adults with hearing loss.”, *Journal of Speech, Language, and Hearing Research*. 1994 Jun;37(3):510-21.
8. Bolfan-Stosic N, Simunjak B. “Effects of hearing loss on the voice in children.”, *Journal of otolaryngology*. 2007 Apr 1; 36(2):120-3.
9. Shriberg LD, Paul R, Black LM, Van Santen JP. “The hypothesis of apraxia of speech in children with autism spectrum disorder.”, *Journal of autism and developmental disorders*, 2011 Apr 1;41(4):405-26.
10. Albertini G, Bonassi S, Dall’Armi V, Giachetti I, Giaquinto S, Mignano M. “Spectral analysis of the voice in Down syndrome.”, Research in developmental disabilities. 2010 Sep 1;31(5):995-1001.
11. Teixeira, J.P. and Fernandes, P.O.,”Acoustic analysis of vocal dysphonia”, *procedia computer science*, *64*,2015, pp.466-473.
12. Talkin, D.,” A robust algorithm for pitch tracking (RAPT)”. *Speech coding and synthesis*, *495*, 1995 p.518.
13. Farrús, M., Hernando, J. and Ejarque, P.,”Jitter and shimmer measurements for speaker recognition.”, In *Eighth annual conference of the international speech communication association*. 2007
14. Akbari, E., Seifpanahi, S., Ghorbani, A., Izadi, F., & Torabinezhad, F. (2018). “The effects of size and type of vocal fold polyp on some acoustic voice parameters.”, *Iranian journal of medical sciences*, *43*(2), 158.
15. Teixeira, J.P., Oliveira, C. and Lopes, C.,”Vocal acoustic analysis–jitter, shimmer and hnr parameters”, *Procedia Technology*, *9*, 2013, pp.1112-1122.
16. Ferrand, C.T., 2002. “Harmonics-to-noise ratio: an index of vocal aging.”, *Journal of voice*, *16*(4), pp.480-487.
17. Yumoto, E., 1983. “The quantitative evaluation of hoarseness: A new harmonics to noise ratio method.”, *Archives of Otolaryngology*, *109*(1), pp.48-52.
18. MDVP manual. Model Model 5105, Kay PENTAX, Lincoln Park, New Jersey, USA.
19. P. Boersma, “Praat, a system for doing phonetics by computer”, *Glot International*, vol. 5, no. 9/10, pp. 341–345, 2001.
20. Zakkula, S., Sreedevi, B., Anne, G., Prakash Manne, S.H.B.O. and Jyothi Atla, S.D., “Evaluation of palatal plate thickness of maxillary prosthesis on phonation-a comparative clinical study.”, *Journal of Clinical and Diagnostic Research: JCDR*, 2014, *8*(4), p.ZC11.
21. Gladis, D. and Dalvi, U., “ A study of F0 estimation based on RAPT framework using sustained vowel.”, *International Conference on Advances in Computing, Communications and Informatics (ICACCI),2015* (pp. 2290-2295). IEEE.

AUTHORS PROFILE

Mrs Prarthana Karunaimathi, is a Full-time Research Scholar, Department of Computer Science, Presidency College, Chennai. She completed Master of Computer Applications in Madras Christian College, Chennai. Her area of research is on Clinical Voice Analysis and her earlier works have been published in IEEE and ACM publications.

Dr. D. Gladis is the Principal of Bharathi Women’s College, Chennai. She has headed PG and Research Department of Computer Science, Presidency College, Chennai. She has guided more than fifteen M.Phil and Ph.D research scholars and has been an active researcher in the areas of Neural Networks, Image Processing, and Data Mining.

Dr. Balakrishnan Doraisami, is an ENT surgeon with three decades of experience. He has authored several research papers. He has written a book titled Clinical ENT Made Easy and a booklet in Tamil, on Malignancies in the Head and Neck. He is currently working in SRM Medical College, SRM Institute of Science & Technology as Professor of ENT and is the Head of the Department of Audiology & Speech. His areas of research are in ENT, Audiology and Allied health sciences. His current project is on Hearing in diabetics and in new born.