# Comparison of Three Different Softwares for Evaluation of Voice Disorders

Register No.M9806

This Dissertation submitted as part fulfilment for the Final Year M.Sc, (Speech and Hearing), submitted to the University of Mysore, Mysore.

ALL INDIA INSTITUTE OF SPEECH AND HEARING

MYSORE 570006

MAY 2000

"With a head
bowed in
Reverence and a
heart
filled with love
and gratitude,
this
research work is

dedicated to
my dearest father
***Mr. Siva Subramanian***

MAY 2000

# CERTIFICATE

This is to certify that this Dissertation entitled : Comparison of Three Different Softwares for Evaluation of Voice Disorders is the bonafide work in part fulfilment for the degree of Master of science (Speech and Hearing) of the student with Register No.M9806

Mysore
May, 2000

Director
All India Institute of
Speech and Hearing
Mysore 570 006.

# CERTIFICATE

This is to certify that this Dissertation entitled : **COMPARISION OF THREE DIFFERENT SOFTWARES FOR EVALUATION OF VOICE DISORDERS** has been prepared under my supervision and guidance.

Mysore,

May, 2000


**Dr.Nataraja N.P.**

Prof. & HOD Sp.Sciences
All India Institute of
Speech and Hearing
Mysore 570 006.

# DECLARATION

This Dissertation entitled : *Comparison of Three Different Softwares for Evaluation of Voice Disorders* is the result of my own study under the guidance of Dr.N.P.Nataraja, Prof, and HOD, Department of Speech Sciences, All India Institute of Speech and Hearing, Mysore and has not been submitted earlier at any University for any other diploma or degree.

Mysore
May 2000

Reg. No.M9806

# ACKNOWLEDGEMENTS

I extend my heartful gratitude to my guide Dr.N.P.Nataraja,PROFESSOR AND HEAD OF DEPARTMENT OF SPEECH SCIENCE, AI.I.S.H, MYSORE for his valuable guidance,support and encouragement to enable me complete this work.

My thanks to Dr.Nikam-DIRECTOR,A././.S.HMysore for granting me the permission to carry out this study.

I thank Dr.MJayaram, DIRECTOR,A.I.I.S.H,Mysore.

I am grateful to the staff of the SPEECH SCIENCES Department of A.I.I.S.H, Mysore for their timely help and co-operation.

Thanks to the LIBRARY STAFF, A.I.I.S.H.Mysore for their co- operation in my library work related to this tudy.

I am obliged to all the subjects who lend their voice for this study.

Dearest A M M A......In silence does my heart weep__its hardly four monthsthat you have left for your HEAVENLY ABODE but it seems like years.I know not what the future has in hold but I know that it will be your blessings and memories which will continue to show the right path to tread upon.

Dearest Appa : You have been like a lonely mighty tree fighting against the storm in the outside world. You have set a wonderful example and enabled me to realisethat one can survive a storm of any kind provided one has the collage the will power and a positive attitude to do so. I am indebted to you.

Dearest T V M Ammai: You have been an epitome of love and sacrifice. My dearest Granny... I can never repay you for all that you have done.

Snarath, Atchyuth and Srikanth. My dearest trio brothers .You have made me realise that together we can fight against thetoughest times with a smile . I thank the ALMIGHTY for having given me such wonderful brothers.

Dear Athimber.You have provided help in innumerable ways which will forever be remembered.

Dearest NKM :As the years go by, our relationship continues to grow stronger with each passing day. With your support and encouragement the "bumpy roads of life seemall the more smoother ".

Dear Uncle and Aunty (Mrs.&. Mr.Nandur). You have been with my family during our worst times.I can never thank you enough for your constant support and encouragement.

Dearest LURIA(Anita) :Our friendship is like a "golden chain "with each link forged of memories that make youstill more dear, the thoughtful things you say and do become the links of gold that I will treasure through the years and in my heart will I hold.

Dear Vasanti: We have had more number of disagreements than agreements, but still the bond of friendship holds us both together.......forever!

Ruchi,Archdnd,Prarthand,Saritha: You have set an important landmark in my life___you have given me the chance to be the sole privilidged listener of your "first and last album.. the RAPS". You have always kept me smiling and you have been there when i needed you. Thank you for everything.

Sarah___I am glad i got to know you. Your small help made a big difference. Amirtha .you will always be remembered for all your "clownish acts ".

Kusum,Nivedithd,Beena,Sowjanya Chaitdnya. Your love and affection is immeasurable.! am glad I have got such wonderful friends.

I thank Akila ,Reddy and Sapna for constantly enquiring about the progress in my work.

I thank all my cldssmates for all their support and encouragement

I thank all my juniors(a special thanks to Bhanumdthi) for helping me with my research work when needed.

I thank Rdjdidkshmi akka without whose quick,patient and efficient typing this work would not have been submitted on time.

I thank all those whose names I would have missed but who have helped me with my work.

# TABLE OF CONTENTS

# LIST OF TABLES

# LIST OF GRAPHS

# INTRODUCTION

Voice which has been defined as "the laryngeal modulation of the pulmonary airstream, which is further modified by the configuration of the vocal tract" (Michael and Wendhal, 1971) depends on the synchrony or coordination between respiratory, phonatory and resonatory systems. Deviation in any of these systems may lead to voice problems.

There are various means of analysing voice, developed by different workers, to note the factors which are responsible for creating an impression of a particular voice (Hirano, 1981; Rashmi, 1985).

Development of technology has permitted the analysis and measurement of various aspects of vocal function. There have been many attempts over the years to find different voice parameters using objective methods that aid in early detection, diagnosis and treatment of dysphonics. The various objective approaches are high speed cinematography, stroboscopy, electroglottography, sound spectrography, photoglottography, echoglottography and inverse filtering. Even though, these techniques have been promising, there have been problems with instrumentation, methodology and analysis.

Presently, acoustic analysis of voice is gaining more importance. Hirano (1981) states that "____this may be one of the most attractive method of assessing the phonatory function or laryngeal pathology because it is non-invasive and provides objective and quantitative data . Acoustic analysis can be done by using methods such as spectrography, peak picking, inverse filtering, computer based methods and others.

In computer based techniques there are many softwares or programs which are designed to extract different parameters of voice. However, three softwares or programs that are available at All India Institute of Speech and Hearing, Mysore and are used in the present study i.e. (i) Multidimensional voice program-model 4305 developed and marketed by Kay Elemetrics Inc, New Jersey (ii) Vaghmi software developed by Voice Speech Systems, Bangalore, (iii) Dr.Speech software developed and marketed by Tiger DRS. These software acquires, analyses and displays the voice parameters from a single vocalisation.

There have been various Indian studies which have used the above mentioned softwares for evaluation of voice disorders. Studies conducted by Anitha (1994), Biswajith Das (1995), Aparna (1996) and Preethi (1998), David (1998), have used these softwares.

Thus eventhough these softwares have been used by various workers, it leads to confusion to a clinician regarding the selection and use of any particular software. Hence the present study was undertaken to compare three different software or programs in terms of these efficacy in the evaluation of voice disorder.

In the present study a total of 48 parameters were used out of which 29 parameters were extracted using MDVP software, 8 parameters were extracted using Vaghmi software and 11 parameters were extracted using Dr.Speech software.

The parameters extracted were :

1. Average Fundamental frequency
2. Average Pitch Period
3. Highest Fundamental Frequency
4. Lowest Fundamental Frequency

5. Standard Deviation of Fundamental Frequency
6. Phonatory Fundamental frequency Range in Semitones
7. Fundamental Frequency Tremor Frequency
8. Amplitude tremor frequency
9. Absolute Jitter
10. Jitter Percent
11. Relative Average Perturbation
12. Pitch Period Perturbation Quotient
13. Smoothed Pitch Period Quotient
14. Fundamental Frequency Variation
15. Shimmer in dB
16. Shimmer Percent
17. Amplitude Perturbation Quotient
18. Smoothed Amplitude Perturbation Quotient
19. Peak Amplitude Variation
20 Frequency Tremor Intensity Index
21. Amplitude Tremor Intensity Index
22. Degree of Voiceless
23 Noise to Harmonic Ratio
24. Voice Turbulence Index
25. Soft Phonation Index
26. Degree of Voice Breaks
27. Number of Voice Breaks
28. Number of Subharmonic Segments
29. Degree of Subharmonic Components
30. Fundamental Frequency in Phonation
31. Fundamental Frequency in Speech
32. Maximum Fundamental Frequency in Speech

33. Minimum Fundamental Frequency in Speech
34. Extent of Fluctuation in Fundamental Frequency
35. Speed of Fluctuation in Fundamental Frequency
36. Extent of Fluctuation in Intensity
37. Speed of Fluctuation of Intensity in Phonation
38. Maximum Fundamental Frequency
39. Minimum Fundamental Frequency
40. Habitual Fundamental Frequency in Phonation
41. Fundamental Frequency Tremor
42. Amplitude Tremor Frequency
43. Standard Deviation of Fundamental Frequency
44. Jitter %
45. Shimmer %
46. Harmonics to Noise Ratio
47. Normalised Noise Energy
48. Signal to Noise Ratio.

Parameters 1-29 were evaluated using MDVP software
Parameters 30-37 using Vaghmi software
Parameters 38-48 using Dr.Speech software.

A group of 30 normal subjects (15 males and 15 females) in the age range of 17 to 25 years and a second group of 30 dysphonic subjects which formed the experimental group (15 males and 15 females) in the age range of 18 to 65 years were considered for the study.

All the above mentioned parameters were measured for vowels /a/ /i/ and /u/ phonated by each subject and sentence /idu papu/ /idu koti/

/idu kempubanna/ spoken by each subject and ability to discriminate dysphonics from normal by each software based on the parameters of voice, measured by them has been determined.

It was hypothesized that -

This is no difference among the three softwares i.e. MDVP, Dr.Speech and Vaghmi in terms of their efficacy of evaluation of voice disorders.

The study is limited to (1 ) comparision of three different softwares for the evaluation of voice disorders. A combination of other objective methods can be used in future studies (2) Most of the analysis was carried out on phonation sample though speech samples were also used in the study. (3) Only 30 normals and 30 dysphonics have been used for the study.

*Implications of the study.* Comparison of the softwares and their usefulness in evaluating the voice.

# REVIEW OF LITERATURE

"Every human society, no matter how primitive, has developed the ability to communicate through speech and our ability to communicate through spoken and written language has been cited as the single most important characteristic that sets human apart from other animals" (Curtis, 1978).

Speech is a form of language that consists of the sounds produced utilising the flow of air from the lungs. Speech is the way of life for man and is its chief medium of social adaptation and control. According to Boone (1971), speaking and singing demand a combination or interaction of the mechanisms of respiratory phonation, resonation and speech articulation". The underlying basis of speech is voice/Voice plays the musical accompaniment to speech, rendering it tuneful, pleasing audible and coherent, and it is essential feature of efficient communication by the spoken word" (Greene, 1964). It is well known that voice has both linguistic and non-linguistic functions in any language. Voice is the carrier of speech. Variations in voice in terms of pitch and loudness, provide rhythm and also break the monotony. At the semantic level also voice plays an important role. The use of different pitches, high and low, with the same string of phonemes would mean different things. Speech prosody-the tone, the intonation and the stress or the rhythm of language is a function of vocal pitch and loudness as well as of phonetic duration.

Perkins (1971) has identified at least five non-linguistic functions of voice. Voice can reveal speaker identity i.e., voice can give information regarding sex, *age,* height and weight of the speaker. Lass et al. (1980) report several studies which have shown that it was possible to identify the speaker's age, sex, race, socio-economic status, social features, height and weight based on voice. Voice has also been considered to be reflecting the physiological state of the individual. For example, a very weak voice may indicate that the individual may not be keeping good health, or a denasal voice may indicate that the speaker has common cold. Apart from this, it is a well known fact that voice basically reflects the anatomical and physiological conditions of the respiratory, phonatory and resonatory systems, i.e. deviation in any of these systems may lead to voice disorders. The quality of voice may become important for certain professionals, for example, radio/TV announcers, actors and singers.

The term voice has been defined by Michel and Wendahl in 1971 as "The laryngeal modulations of the pulmonary airstream, which is then further modified by the configuration of the vocal tract. There have been controversies regarding what is normal voice? and who has an abnormal voice? At present it is difficult to find a comprehensive definition of normal voice". West, et al. (1957) offer the following criteria for normal voice "Adequate loudness, clearness of the tone, pitch appropriate to the age and sex, a slight vibrato and a graceful and constant inflection of pitch and force which follows the meaning of what is spoken". They state that the departure from these norms should be considered abnormal. On the other hand, Van Riper and Irwin (1963) state that "Voice can vary widely with respect to pitch, loudness and quality without appearing abnormal and the concept of normal voice may be related to cultural preferences, age and sex as well as to social and economic status". Both the above definitions are found to be vague and ambiguous. This ambiguity in the usage of the terms has percolated to the classification of voice disorders.

Many classifications of voice disorders have been putforth based on different points of view (Froschels, 1940; Broadnitz, 1959;1 Greene, 1964,-

Murphy, 1964; and Moore, 1971). Mysak (1966) has classified the voice disorders into the following categories :

1. Phonatory and resonatory disorders of infraglottal origin.
2. Phonatory and resonatory disorders of glottal origin.
3. Phonatory and resonatory disorders of supraglottal origin.

Under the first category, Mysak (1966) includes the problem of vocal weakness which according to him is caused by inadequate subglottic air pressure. Vocal cord paralysis, vocal nodules and laryngectomy are included in the second category while complexes associated with deficits of velopharyngeal closure finds a place in the last category. Sokoloff (1966) has given a classification of voice disorders which includes the following :

1. Phonatory problem due to hyperfunction
2. Phonatory problem due to hypofunction
3. Phonatory problem due to abnormal resonance (Supraglottal cavities)

The first category includes harsh and hoarse voice, pitch disorders, the second category consists of breathiness, hysterical aphonia etc., while the last category includes hypo and hypernasality. A similar classification of hypo and hyperfunction has been employed by Froschels (1940) and Broadnitz (1959). With reference to this classification (hyper and hypo function) if used excessively, over simplifies the complexities of laryngeal pathologies, placing excess emphasis on the degree of approximation of the vocal edges rather than one the multiple causes of such approximation deficits".

A classification based on etiology had been employed by Moore, 1971. The problems which are perceived as abnormal pitch, loudness or quality may be directly related to the mechanisms of the respiratory, phonatory and resonatory systems. When there is a voice disorder it would mean that one or more of the systems i.e. respiratory, phonatory and resonatory, is or are not functioning normal either because of structural or physiological conditions or due to faulty learning.

Boone (1977) classifies voice disorders based on changes in vocal fold mass-size and approximation. This is "based on the fact that normal phonation requires proper mass size adjustments and that the two vocal folds approximate one another optimally along their entire length". This classification is also considered as over simplification of the laryngeal function (Morris and Spriestersbach, 1978; Aronson, 1980). Pannbacker (1984) after reviewing these classification states that there is considerable overlap between these classifications, as the laryngeal structure and function and perceptual attributes are interrelated.

Fairbanks (1960) classifies voice quality disorders into (a) harshness (b) breathiness and (c) hoarseness. Though Jensen (1965) questioned the validity and reliability of this classification, still it is used. Hoarseness is a common symptom of many laryngeal disorders and many a times, it is the only and the first symptom to be noticed. Literature reveals that hoarseness is related to a large number of laryngeal disorders.

Sederholm et al. (1992) showed with the help of factor analysis that hyperfunction, breathiness and roughness are good predictors of hoarseness. Harshness and breathiness are two components of hoarseness. Harshness is perceived due to irregularity of vocal fold vibrations (Coleman, 1960,- Wendahl, 1963,- 1966; Moore, 1975) i.e. Variations or perturbations in both amplitude and time period from cycle to cycle give the impression in hoarseness. Breathiness is perceived by escape of air through partially closed glottis and the resultant turbulence noise reduces the harmonic to noise ratio (HNR). Excessive aperiodicity also generates noise and reduces the prominence of the harmonics, hence reducing the Harmonics to noise ratio. Thus, hoarseness is defined as a voice quality which clearly contains noise components and that can be labelled harsh and breathy (i.e. source noise elements plus friction noise); its perceived pitch tends to vary substantially,- common description of this quality are 'noisy', 'harsh', 'wet' (Anders et al. 1988).

The primary or common factor in hoarseness is noise of a relatively high frequency that is produced by transient vibrations. These sounds are combined with other phonatory sounds that are frequently at low pitch as the result of laryngeal disease or any other condition that would lower the frequency of vocal fold vibration. The transient disturbances seem to occur on the surface of the vocal folds, particularly along the glottis, but other laryngeal structures may also contribute to the total effect.

Sources of laryngeal transients can be grouped into four categories (1) Accumulation of sticky mucus secretion in the larynx. Excessive mucus tends to interfere with normal movements of vocal folds by weighing them unevenly and damp in their excursion through causing them to adhere to each other. (2) Relative flacidity of one or both vocal folds, the flacidity causes independent vibration, resulting transient disturbances. (3) Additions to the mass of the folds. Mass causes pitch changes, hoarseness by weighing stiffening and influencing vocal fold's compliance. (4) The destruction of all or part of the vocal folds cause random vibrations and transients resulting hoarseness (Moore, 1971).

Acoustic analysis of hoarseness has been extensively carried out in an attempt to understand the acoustic characteristics of hoarseness. The measures used in acoustic analysis of voice are convenient, non-invasive, objective, sensitive and quantitative method of studying laryngeal mechanism while producing speech. Studies have been conducted to identify measurable voice features that are correlated with hoarseness and thus effectively predict the degree of hoarseness perceived by listeners. Some of the advantages of these methods are that quantitative data from the correlated measurement could be easily stored or transmitted to those who need to see them. The measurement are repeatable from audio recorded voice samples and in due course might be obtained by standardized procedures in clinical or research contexts.

The purpose of research in acoustics is aimed at :

Speaker identification

Delineating the mechanophysiologic limitations of normal and pathological laryngeal performance (Scripture, 1906).

Detecting and discriminating the types of vocal pathology.

Monitoring and tracking response to therapy.

Searching for acoustic correlates of voice quality (Moore and Thompson, 1965) and check, their variations with voice production conditions caused by various pathologies (Ludlow, 1981).

In checking for the information regarding the magnitude of acoustic parameters that can be used in the field of speech synthesis helping in simulation of desired quality either normal or abnormal (Gill, 1961). This also helps in automatic analysis of fundamental frequency.

Correlating the movement for validating perception based on vocal hoarseness ratings as eg. those obtained on individual listeners or listener panels for either clinical research purposes.

The estimation of hoarseness and other vocal qualities by acoustic parameters varies with different methodological issues. Some of these are :

Contact mic/accelerometer - It is sensitive to body surface vibrations. When placed in intimate contact with the skin on surface of the neck, their output reflects vocal fold movement and the response of the body wall to the acoustic wave in trachea.

Fourcin (1 981 ) made simultaneous recordings of Electroglottograph and airflow velocity curves for different modes of phonations and described the method to interpret laryngeal wave forms. Electroglottograph reflects the vibratory cycle of the VF's with fairly high fidelity. According to Dejonckre and Lebacq (1995) "Electroglottograph reflects the glottal conditions more during the closed phase. As majority of laryngeal pathologies manifest abnormalities more during the closed phase, Electroglottograph has been considered as a better technique for studying VF movement in dysphonics .

Inverse filtering technique is an acoustic procedure with the inverse of the lip and the vocal tract effect radiations are done. The vocal tract spectral contributions are used to remove acoustic effects of the supra glottal vocal tract result in only with the glottal spectrum. One of the disadvantage is that, it is difficult to determine the parameters for the inverse filter model from the speech signals.

Cepstral analysis was first described by Noll (1 964). It relies on the fourier analysis of the speech signal. The speech signal is filtered by a low pass filter and then digitized in order to perform Fast Fourier Transform (FFT). Koike (1 986) applied cepstral analysis to study short term perturbation.

Hirano (1 989) had discussed the current international status of clinical voice evaluation on the basis of the answer to a questionnaire administered to laryngologists, phoniatricians, and speech pathologists working in various parts of the world, which is summarised below:

Measurements and procedures used :

1 . AIRFLOW : Subglottic pressure
Glottal resistance
Glottal efficiency
AC/DC ratio
Airflow/intensity ratio
Phonation Quotient (PQ)
Vocal Velocity Index (VVI)
Maximum Phonation Time (MPT)

2. Fo RANGE : SPL range
Habitual Fo
Habitual SPL
Voice Range Profile
Vocal Register Examination

3. STROBOSCOPY : Video Stroboscopy
Ultra High Speed Photography
Electroglottography
Photoelectric Glottography
Ultra Sound Glottography

4. TAPE RECORDING : Sound Spectrogram
Pitch Perturbation
Amplitude Perturbation
S/N Ratio
Noise Energy Measurement
Spectrum Envelope
LTAS

Inverse Filter Acoustic

Inverse Filter Aerodynamic

VOT

5. LARYNGEAL MIRROR : Telescopy of Larynx

Fibroscopy of Larynx

Microscopy of Larynx

6. X-RAY LARYNGOGRAPHY: Computed tomography

Magnetic Resonance Imaging

7. EMG NEEDLE SURFACE

8 VITAL CAPACITY : Pulmonary function test

Ribcage and Abdominal Movements.

9. AUDIOMETRY.

Read, et *al.* (1 990) in their survey of five microcomputer programs and two dedicated devices for recording, editing and analyzing speech have reviewed the capabilities, requirements and user interface of each system. The programs surveyed include :

Speech by Paul Milenlcovic, Madison, Wl, CSRE (Canadian Speech Research Environment) by Donald Jamieson, London, Ontario;

ILS-PC (Interactive Laboratory System) by Signal Technology Inc, Goleta, CA;

MSL (Microspeech Lab) by Software Research Corp,- Victoria, B.C. Mac Speech Lab. II G.W. Instruments, Cambridge, MA.

Dedicated Devices

Kay Model 5500 DSP Sona-Graph, by Kay Elemetrics, Pine Brook, **NJ.**

General capabilities

Table-1 shows the systems capabilities under the four headings of waveform acquisition and display, waveform operations,spectral and pitch analysis, and other functions. These are general capabilities : that vary in detail among the systems.

Waveform Acquisition and Display pertains to the initial capture of the signal. Listed under this heading are nine capabilities or specifications. The first two are the number of channels that can be simultaneously recorded (with at least 8 kHz sampling rate per channel) and simultaneously displayed. A range in the table means that the answer varies with the A/D board used. Some systems permit a rapid alternation, but not juxtaposition, of signal displays, we do not regard this as a simultaneous display. Capacity indicates whether the signal duration that can be recorded is limited by the amount of available memory or by the space available on a hard disk. Normally, the latter permits much longer signals to be recorded. Record/playback is the capability to sample and store a signal and to play the stored signal. Monitor display is the capability to display a signed on a monitor. Zoom is the ability to magnify a selected portion of a signal, so that the display represents a narrower segment of time or possibly frequency than in the original display, and

TABLE: 1

| | *CSpeech* | *CSRE* | *ILS-PC* | *Kay 7800* | *Kay 5500* | *MacSpeech* | *MSL* |
|---|---|---|---|---|---|---|---|
| **A. Waveform acquisition and display: Initial capture of signal** | | | | | | | |
| Recording: no. of channels | 1–4 | 1 | 1–3 | 2 | 2 | 1 | 1 |
| Display: max. no. of channels | 8 | 1 | 16 | 1 | 2 | 1 | 1 |
| Capacity (disk memory limit) | D | D | D | M | M | M | M |
| Record/playback | Y | Y | Y | N | Y | Y | Y |
| Monitor display | Y | Y | Y | N | Y | Y | Y |
| Zoom/scroll | Z | Z | Z | N | Z/S | Z | Z |
| Time readout | Y | Y | N | N | Y | Y | Y |
| Amplitude readout | Y | Y | N | N | Y | Y | Y |
| Amplitude resolution (max bits) | 16 | 16 | 16 | 12 | 16 | 12 | 10 |
| **B. Waveform operations (Subsequent to intial capture)** | | | | | | | |
| View & play selected segment | Y | Y | N | N | Y | Y | Y |
| Erase | Y | Y | Y | Y | Y | Y | N |
| Splice | Y | Y | Y | Y | Y | Y | N[1] |
| Replicate | Y | Y | Y | N | Y | Y | N |
| Taper ends | Y | N | Y | N | N | N | N |
| Save/retrieve | Y | Y | Y | N | N[2] | Y | Y |
| Label segment | Y | N | Y | N | N | N | N |
| **C. Spectral and pitch analysis** | | | | | | | |
| Spectrogram | N | Y | Y | Y | Y | Y | N |
| Formant tracing in s'gram | n/a | Y | Y | N | N | Y | n/a |
| Readout time, freq, ampl. | n/a | Y | N | N | Y | Y | n/a |
| Spectrum: FFT | Y | Y | Y | Y | Y | Y | Y |
| LPC | Y | Y | Y | N | Y | Y | N |
| Waterfall | Y | Y | Y | N | Y | N | N |
| Readout freq, ampl. | Y | Y | N | N | Y | Y | Y |
| Cepstrum | N | N | Y | N | N | N | N |
| Pitch extraction | Y | Y | Y | N | Y | Y | Y |
| Voice perturbation analysis | Y | Y | N | N | Y[3] | N | N |
| **D. Other functions** | | | | | | | |
| Speech synthesis | N | Y | Y | N | Y[4] | N | N |
| Data import/export | Y | N | N | N | Y | N | N |
| Stimulus presentation utils. | Y | Y | N | N | N | N | N |
| Simultaneous displays: | | | | | | | |
| Waveform & spectrum | N | N | Y | Y | Y | Y | Y |
| Waveform & spectrogram | n/a | N | Y | Y | Y | Y | n/a |
| Spectrum & spectrogram | n/a | Y | Y | Y | Y | Y | n/a |
| Spectra: FFT & LPC | Y | N | Y | n/a | N | Y | n/a |
| FO and waveform or s'gram | Y | Y | Y | n/a | Y | Y | Y |

[1] Can mark and save any continuous portion. In signal ABC, can mark B and play AC, but cannot save AC in a file.
[2] Can save & retrieve with optional computer interface.
[3] Optional analysis package.
[4] LPC editing and resynthesis with optional computer interface and LPC software.

scroll is the ability to move through and display *a* signal held in a buffer. Time readout is the numeric display of time values for selected points in the displayed signal. Similarly, amplitude readout is the numeric display of amplitude values for selected points in the signal. Amplitude resolution is the maximum signal resolution per channel in bits.

Waveform Operations are capabilities or features available subsequent to initial capture of the signal. View and play selected segment allows the user to select one portion of the captured signal for monitor display and playback. Erase is the capability to delete a selected part of the signal, and splice is the ability to join two waveforms that were originally noncontinuous. Replicate is the ability to splice a copy of any portion of a signal to the original. Taper ends refers to amplitude modulation of a waveform in which the ends of a selected sample can be shaped to initial and final values of zero (useful in splicing). Save/retrieve is the capability to store a waveform in *a* file that can be subsequently retrieved. Label segment is a feature by which the user can markand label portions of the speech signal, such as for phonetic segmentation.

Spectral and Pitch Analysis includes various forms of spectrograms, spectral analyses, and pitch extraction. Spectrogram refers to the traditional three-dimensional (frequency time-intensity) display of a running short-term spectrum. Formant tracing is the ability to display formant frequency traces superimposed on a spectrogram. This is achieved either by manual tracing from the displayed spectrogram (MacSpeech Lab II), or by superimposing LPC- derived formant frequencies on a spectrogram (ILS-PC). Readout of time, frequency, and amplitude is the numeric display of values in these domains for selected points on a spectrogram. Spectrum indicates analyses such as FFT (Fast Fourier Transform) or DFT (Discrete Fourier Transform), LPC (Linear Predictive Coding), and waterfall (time history of spectra). Readout of frequency and amplitude refers to the numeric display of frequency and amplitude for a selected point on the spectrum. Cepstrum is the Fourier transform of the logarithm of the amplitude spectrum (or, simply, the inverse transform). Pitch extraction is the determination of the fundamental frequency of voiced portions of the speech signal. Voice perturbation analysis is the capability to analyze jitter (cycle-to-cycle perturbations in the fundamental period of the waveform) and/or shimmer (cycle-to-cycle perturbations in the amplitude of the waveform).

Other Functions include a variety of additional features. Speech synthesis is the capability to generate speech signals either by parameter synthesis from tables data or by LPC resynthesis. Data import/export is the means to transfer data files from and to other programs (including header information). Stimulus presentation utilities refers to capabilities for preparing stimuli for listening studies. Simultaneous displays indicates the ability to display simultaneously various combinations of waveforms, spectra, spectrograms, and fundamental frequency contours. It is not necessary the case that the analyses are simultaneously generated, that is, one analysis may be displayed first and then the other added to the display.

ILS-PC has so many additional capabilities that have been listed in Table 2.

Table-2: ILS analysis capabilities

| | |
|---|---|
| **Spectral analyses** | |
| LPC, "Burg" analysis | Pitch synchronous analysis |
| Cepstrum | 3-D log-magnitude LPC spectra |
| Frequency spectrum | Real or complex cepstrum |
| LPC, formant synthesis | Spectral peak parameters |
| Spectrogram | 3-D FFT spectra |
| LPC spectrum | Pitch synchronous synthesis |
| Amplitude vs. frequency plots of smoothed spectra | Pitch asynchronous synthesis |
| AR modeling, cepstral pitch extraction | Spectral peak selection by root solving |
| **Fundamental frequency analyses** | |
| Peak locations for pitch synchronous analysis | AR analysis, SIFT pitch extraction |
| Auto correlations, Cross-correlations | AR modeling, cepstral pitch extraction |
| LPC pitch extraction. | |
| **Filtering** | |
| Elliptic, Butterworth, Chebychev I/II, linear phase, lo-pass, hi-pass, band reject, octave bands. | |
| **Waveform and signal processing** | |
| Add, substract, multiply, divide, average, convolve, multiplex, demultiplex, clip, random noise generation, many operations, interpolations, extrema locating. | |
| Hilbert transforms | Linear and dynamic time warping |
| Residue calculation (for inverse filtering) | |
| **Other** | |
| Test function generators | Vocal tract modeling |
| 3D area functions | Detection probabilities |
| Parameter amplitude plots, histograms, scatterplots | |
| Statistics : conventional, plus linear regression, curve fitting, compare and classify, discrimination analysis (cigenvectors), feature extraction. | |

## *LONG TERM AVERAGE SPECTRUM (LTAS)*

Recent research has shown that LTAS is a reasonable index of vocal quality (Carr and Trill (1964). The rational behind this technique is that vocal tract transfer function gets nullified after averaging out the various spectra over a prolonged period and the averaged spectrum is the true representative of only the glottal signal. LTAS often reveals pathological laryngeal conditions (Hecker and Kreul, 1971). On the other hand, it does not allow a definite classification of normal and pathological laryngeal conditions. This ambiguity may be caused by the influence of the vocal tract on the spectrum (Klingholtz, 1990) i.e. to say that articulatory behaviour masks laryngeal features in LTAS which questions the basic assumption of the method. The LTAS fails to detect all the fine temporal details of the speech pitch or amplitude (Schoentgen, 1989). But period to period measurements have established statistical measure of period and period perturbation distributions (Askenfelt and Hammarberg, 1986).

### *Speech Sample*

Most of the studies have employed sustained vowels rather than running speech (Horii, 1979). According to Horii (1979) there was a paucity of data on large quantities of speech because of the lack of efficient instrumentation and measurement procedures. There is a strong argument for sustained vowels because it gives only the random perturbations associated with physiological limitations of the glottal sound, source and controls supraglottal sources of variations. This allows measurement of only short term perturbations and checks long term systematic perturbation due to phonetic context, stress and intonation. It is to be stressed that most investigators, whatever their choice of speech material, considered the feasibility of separating normals and dysphonic subjects at a reasonable level of performance. However, looking into these factors, the use of the mid-portion of sustained vowel produced at a natural comfortable pitch and intensity level appears to be the most appropriate phonatory task when changes in perturbation caused by automatic physiologic conditions, of the larynx are in question (Koike, 1 969,- Iwata and Von Leden, 1 970),- Iwata, 1 9 7 2 ; Horii, 1 9 7 9 ) .

***Manual* Vs. *Automatic Analysis***

This include hand marking of analog oscillograms, semiautomatic methods using interactive digital wave form editors and both hardware and software automatic pitch tracers. Some of the earliest studies have involved the use of hand measurments (Lieberman, 1 9 6 1 ). This method is extremely tedious and time consuming because of their minute nature. More recent studies (Horii, 1979; 1 980; Wilcox, 1980) have used computerised instrumentation which is fast and precise, in between these two are the semi automatic instruments. A lot of subjective judgement is required in both manual and semiautomatic and hence automatic extraction gives much precision.

Many have applied computer techniques which use a formula or algorithm for the analysis of the acquired waveform. The majority of acoustic perturbation studies as well as the spectral noise studies have been limited to analysis by means of a single formula (Lieberman, 1961, 1963) and others two formulas (Horii, 1980). Regardless of the algorithm each investigation found that their measure provided some degree of discrimination between normals and pathologic subjects.

Qi and Shipp (1 992) devised a new method for tracking irregularities in the acoustic waveform of a sustained phonation using the adaptive Wiener filter. Irregularities were determined by the techniques of correlation cancellation. The alogrithm was evaluated using sustained vowels produced bya format synthesier and by subjects with and without phonatory disorders. Results indicate that the method is capable of differentiating between normal and abnormal voices.

***Temporal Resolution***

The number of times analog acoustic waveform is in a second during digitization process is termed as temporal resolution. This is also referred to as the sampling frequency or sampling rate and is commonly expressed in the unit of cycles per second (cps) or Hertz (Hz). Temporal resolution is a critical factor affecting all the acoustic mcasurments but especially the accuracy of jitter measurement is limited by the temporal resolution which becomes more important when peak to peak measures are the basis of acoustic analysis.

Cox, et al. (1989a) reported that increasing the sampling frequency from 10 kHz to 20 kHz had little effect on DFT based Harmonics to noise ratio estimates with all differences being 0.6 dB in perturbed data. However, the same in perturbation free data brought HNR from 21.9 dB to 41.2 dB for */i/* vowel and from 29.4 dB to 49.0 dB for/a/ vowel suggesting that over sampling brings a significant improvement in perturbation free datd.

***Amplitude Resolution***

This is commonly known as bit resolution which gives the resolution of a system along the ordinate where the amplitude of the acoustic wave is represented. This is usually expressed in terms of number of bits which can easily be converted into relatively simpler unit of amplitude resolution,- i.e. the number of samples per unit amplitude. Lower bit resolution produces the bit noise contaminating the original analog signal. A minimum of nine bits of resolution are needed to minimize the contamindting bit noise without intra polation (Titze, et al. 1987).

***Interpolation***

Interpolation is a mathematical process which calculates probability estimates of numbers between the actual numbers obtained from the digital sampling of the analog signal. Interpolation provides an obvious advantage for the estimation of filter, particularly if relatively low sample rate is used. The use of interpolation between samples in the extraction of normal vocal jitter was recommended by Titze et al. (1987), Deem et al. (1989) reported that the use of interpolation with peak picking extraction procedures had little effect on the jitter values. On the other hand, the extraction procedures using interpolation with zero crossing yielded the lowest jitter values.

***Waveform Marking***

After successful A-D conversion the data is stored in a file ready for analysis by the program. The various techniques are used to mark the points of interest in each period of the waveform. The user may choose whether to make the maximum peaks, the minimum peaks or the points where the waveform crosses zero line.

According to Titze et al. (1987) - overall peak picking techniques have yielded larger jitter values than the zero crossing techniques. Deem et al. (1989) reported that zero crossing procedures resulted in jitter values approximately 2 to 6 micro seconds lower than obtained with peak picking procedures.

***Sample Duration***

Sample duration in acoustic studies depends upon the optimum size of the window (token) and optimum number of tokens.

*Size of window* - Titze et al. (1987) suggested a window of 20-30 cycles within a given token of steady vowel phonation.

*Type of window* - A tapered window function is reported to be advantageous in HNR estimation for reducing sensitivity to errors in demarcation of data segments.

*Number of Tokens* - A single token of a steady vowel is insufficient to establish a reliable acoustic measure. Hence multiple tokens of an utterance are necessary to obtain a stable mean for perturbation measures (Titze et al. 1987).

## *Vowels*

Perturbation measures have been shown to be different among different vowels by Horii (1979). Normative data from Wilcox and Horii (1980) have shown that /u/ was associated with significantly smaller jitter (OI.55 96) than /a/ and /i/ for which the values were 0.6896 and 0.69% respectively.

Cox et al. (1989 c) reported that HNR varied as much as 25 dB at a given level of perturbation depending on whether/a/, /i/or/u/was being analyzed.

### ***Fundamental Frequency***

The Fo of speech also is an important factor for quantifying Hoarseness. Heiberger and Horii (1982) reported that jitter is systematically affected by the fundamental frequency of the voice i.e., jitter found to be large for low frequency phonation and small for high frequency phonation. Cox et al. (1989 c) reported that the HNR tend to increase with Fo. Increase of Fo from 103 Hz to 203 Hz led to variations of over 6 dB in HNR.

***Sex***

The two sexes differ in terms of their vocal Fo and hence sex itself becomes an important factor in acoustic parameters (Emanuel, et al. 1973).

***Age***

Wilcox (1978) and Wilcox and Horii (1980) reported that a greater magnitude of jitter occurs with advancing age and this, they attribute to the reduced sensory contributions from the laryngeal mechano receptors.

The parameters considered in the present study were :

1) Frequency Parameters

1. Habitual fundamental frequency
2. Fundamental frequency in phonation
3. Fundamental frequency in speech
4. Average fundamental frequency
5. Average pitch period
6. Highest fundamental frequency
7. Lowest fundamental frequency
8. Standard deviation of fundamental frequency
9. Phonatory fundamental frequency range in semitones
10. Fo tremor frequency
11. Absolute jitter
12. Jitter percent
13. Relative average perturbation

14. Pitch period perturbation quotient
15. Smoothed pitch period quotient
16. Co-efficient of Fo variation

Intensity Parameters

1. Shimmer in dB
2. Shimmer percent
3. Amplitude perturbation quotient
4. Smoothed amplitude perturbation quotient
5. Co-efficient of amplitude variation

Frequency and Intensity Parameters

1. Extent of fluctuation in fundamental frequency
2. Speed of fluctuation in fundamental frequency
3. Extent of fluctuation in intensity
4. Speed of fluctuation of intensity

Other parameters

1. Signal to noise ratio
2. Normalised noise energy
3. Noise to harmonic ratio
4. Length of analysed sample.
5. Voice turbulence index
6. Soft phonation index
7. Frequency tremor intensity index

8. Amplitude tremor intensity index

9- Degree of voice breaks

10. Degree of sub-harmonic components

11. Degree of voiceless

12. Number of voice breaks

13. Number of subharmonic segments

***Fundamental Frequency***

The perception of the pitch of a signal is common. The term pitch refers to the human psychophysical response to the acoustic signal and is difficult to quantify. On the other hand, the physical basis of pitch i.e. the fundamental frequency (Fo) of a periodic tone is relatively easy to quantify and measure (Hollien, 1981).

The human listener perceives the lowest frequency in the spectra, the fundamental frequency, as the speakers pitch. Plomp (1967) states that even in a complex tone, where the fundamental frequency is absent or weak, the ear is capable of perceiving the fundamental frequency based on periodicity of pitch, Erickson (1959) is of the opinion that the vocal cords are the ultimate determiners of the pitch and that the same general structure of the cords seem to determine, the range of frequencies that one can produce. The perception of pitch and measurement of fundamental frequency are based on the systematic opening and closing of the vocal folds during the production of voiced speech signals. Hence when fundamental frequency is measured acoustically, the process is actually to count these openings and closings by some method.

Various objective methods have been used to measure the fundamental frequency of vocal cords. For example stroboscopic procedures, pitch meter, high speed cinematography, electroglottography, ultrasonic recordings, cepstrum pitch detection, the 3M plastiform magnetic tape viewer spectrography, digipitch, pitch computer, high resolution signal analyzer, visipitch.

The fundamental frequency of a voiced sound is a function of the mass, elasticity, compliance and length of the vocal folds. Davis (1981) sates that fundamental frequency depends to some extent on subglottal pressure and configuration of vocal and configuration of vocal tract. "More massive folds (longer and thicker) vibrate at naturally lower frequencies than shorter and thinner folds. More elastic folds vibrate at high frequencies because they bounce back faster. Vocal folds vibrate faster when they are tense than when they are slack. The primary way to make a given set of vocal folds more tense is to stretch them. Thus longer folds contribute to increased mass and lower Fo in one condition and to increased tension and higher Fo in another condition. This is because a longer pair of vocal folds (compared to other speakers) will be more massive and produce a low frequency voice, men s voices are lower than children's voices. Vet the lengthening of vocal folds (within the same (speaker) will stretch out and thin the effective vibrating portion of the vocal folds, adding tension and thereby producing a higher fundamental frequency" (Borden and Harris, 1980).

Studies have been carried out by various investigators to provide data about changes fundamental frequency of voice as a function of age (Fairbanks, 1940; 1949/ Curry, 1940; Snidecor, 1943; Hanky, 1949; Mysak, 1959,- Samuel, 1973; Usha Abraham, 1978; Gopal, 1980,- KushalRaj, 1983 and Rashmi, 1985). The voice of a new born has been found to be around 400 Hz

(Indira, 1982). Upto puberty there is very little difference between the voice of boys and girls. The voice change is prominent at puberty.

Broadnitz (1959) states that the voice changes in puberty should be interpreted as the intensification of a process that has begun ready at a much earlier period. There is gradual decrease in fundamental frequency since infancy but this change is marked at puberty.

The fundamental frequency drops slightly during the first three weeks or so, but then increases until about the fourth month of life, after which it stabilizes for a period of approximately five months. Beginning with the first year, fundamental frequency decreases sharply until about three years of age, when it makes a more gradual decline, reaching to the onset of puberty at 11 or 12 years of age. A sex difference is apparent by the age of thirteen years, which marks the beginning of a substantial drop in male voices, the well known adolescent voice change in case of males. The decrement in fundamental frequency from infancy to adulthood among females is somewhat in excess of an octave, whereas males exhibit an overall decrease approaching two octaves (Kent, 1976).

Eguchi and Hirsh (1969) report that children have a fundamental frequency of about 300 Hz even upto the age of 8 and 10 years. They have also stated that the fundamental frequencies of children and adult females are higher than those of the adult male. There is no significant difference in fundamental frequency between 7-8 year old boys and girls (Fairbanks, et al. 1949). Kent (1976) has reported that the fundamental frequency values are distinguished by sex, only after the age of 11 years.

Studies on Indian population have shown that, in males the lowering in the fundamental frequency is gradual till the age of 10-12 years, after which, there is a sudden marked lowering in the fundamental frequency, which is attributable to the changes in the vocal apparatus at puberty. In the case of females, a gradual lowering of fundamental frequency is seen (George, 1973,- Usha, 1979; Gopal, 1980; Kushal Raj, 1983; Rashmi, 1985).

Kent (1976) has cautioned against considering the findings from these cross-sectional studies as the representative of the actual developmental course of voice fundamental frequency at various age levels. Further then he has stated that the longitudinal studies are required. One such study was carried out by Loebell and Karger (1976). The voice of twenty-five children were recorded during puberty for two years, every month. The results showed a significant descent of fundamental frequency, for all subjects during the lapse studied.

The aging trend for males with respect to the mean fundamental frequency is one of a progressive lowering of pitch level from infancy through middle age followed by a progressive raise in the old age (Mysalc, 1966). However, among females, the mean fundamental frequency levels of the 7 and 8 year olds were the highest. A progressive lowering of fundamental frequency level is seen till the age of a young adult female. No significant change is seen from young adulthood to the aged group which is in contrast to the male population (Mysak, 1966).

It is generally believed that the voice becomes weak and tremulous and high-pitched in old age, and it is obvious that the singing voice deteriorates much earlier than the speaking voice. There are great individual variations in age of onset and degree of vocal deterioration in old age. Much depends upon the quality of phonation earlier. A fine voice and especially a trained voice need not deteriorate at all in speech (Greene, 1 9 7 2 ). The factors that might be responsible for voice changes in old age include calcification of laryngeal cartilages, loss of elasticity, and atrophy of laryngeal muscles. In some, auditory changes might be responsible for characteristic voice changes.

The study of fundamental frequency has important clinical implications.

Jayaram (1 97 5) has found a significant difference in habitual frequency measures between normals and dysphonics. Cooper (1 9 7 4 ) has used spectrographic analysis, as a clinical tool to describe and compare the fundamental frequency and hoarseness in dysphonic patients before and after vocal rehabilitation. Study carried out by Asthana (1 9 7 7 ) has shown that the cleft palate speakers had significantly less nasality at higher pitch levels than habitual pitch. But degree of perceived nasality did not change significantly when habitual pitch was lowered.

Most of the therapies aim to alter the habitual pitch level of the patients or make the patient to use his optimum pitch (Cowan, 1 9 3 6 ; West et al. 1957,- Thurman, 1 958,- Anderson, 1 9 6 1 ; Greene, 1 964,- Murphy, 1 9 6 4 ). Nataraja and Jayaram (1 9 8 2 ) have reported that most of the therapies of voice disorders are based on the assumption that each individual has an optimum pitch at which the voice will be of a good quality and will have maximum intensity with least.expense of energy and goal of therapy is to alter the habitual pitch level so that the patient uses optimum pitch.

Anitha (1994), David (1998), Preethi (1998) found significant differences in fundamental frequency between males and females in both the normal and dysphonic groups. Biswajit (1 9 9 5 ) and David (1 9 9 8 ) found no significant differences between normals and dysphonics in terms of fundamental frequency.

However a study done by Anitha (1994) indicated significant differences between normals and dysphonics in terms of fundamental frequency.

Thus measurement of fundamental frequency is useful both in diagnosis and therapy.

***Fundamental Frequency in Phonation***

Fundamental frequency is the lowest frequency that occurs in the spectrum of a complex tone, invoice also, the fundamental frequency is considered the lowest frequency in the voice spectrum. This keeps varying depending upon several factors.

Arnold (1961) suggested that "both quality and loudness of voice are mainly dependant upon the frequency of vibration. Hence, it seems apparent that frequency is an important parameter of voice. There are various objective methods to measure the fundamental frequency of the vocal cords. Cooper (1974) uses spectrographic analysis, as a clinical tool to determine and compare the fundamental frequency in dysphonic before and after vocal rehabilitation. Jayaram (1975) found a significant difference in habitual frequency measures between normals and dysphonics. Preethi (1998) stated that the fundamental frequency in phonation in normal and dysphonic males is lower than the fundamental frequency in phonation in normal and dysphonic females. These findings are consistent with the findings of Nataraja (1986) and Jayaram (1975).

***Fundamental Frequency in Speech***

Many investigators have studied fundamental frequency as a function of age and in various pathological conditions. Different types of speech samples, i.e., phonation, reading, spontaneous speech and singing have been used in different studies. Clinical experience has shown that the subjects use different fundamental frequencies under different conditions. Nataraja and Jagadish (1 984) conducted an experiment to verify this clinical impression. They measured fundamental frequency in phonation, reading, speaking and singing and also the optimum frequency in thirty normal males and thirty normal females. They observed that the fundamental frequency increased from phonation to singing with speaking and reading in between.

The age dependent variations of mean SFF reported by Bohme and Hecker (1 970) indicate that the mean SFF decreases with age upto the end of adolescence. A marked lowering take place during adolescence in men. In advanced age, the mean SFF becomes higher in men but is slightly lower in women. Hudson and Holbrook (1981) investigating mean model frequency, in reading in hundred young black adults whose ages ranged from 1 8 to 29 years and found to be 110.15 Hz in males and 193.10 Hz in females.

Shipp and Huntington (1 965) reported that no significant differences have been noticed in the mean and median SFF between laryngitic and nonlaryngitic voices. Murray (1 978) studying the SFF characteristics of four groups of subjects, namely vocal fold paralysis, benign mass lesion, cancer of the larynx and normals, noted that the parameters of mean SFF failed to separate the normals from the three groups of pathological subjects.

Hammerberg (1980) studied the pitch and quality characteristics of mutational voice disorders before and after therapy. This study included 1 3 young men with mutational voice disorder, age ranging from 1 3 to 1 8 years (1 0 subjects) while 3 subjects were between the age of 26 and the results of this study showed a difference of approximately 1 0 octaves between the pitch levels of 1 3 and 1 8 years old group. The mean value speaking Fo in pretherapy lowered to a mean value of 1 1 9 Hz after therapy. Production measures such as directional and magnitudinal perturbation the SFF improved the discriminant function between normal voice and voice of the patients with malignancy of the larynx. It is considered that the FF in voice disorders would act as a diagnostic and prognostic indicator.

Preethi (1 9 9 8) found no significant difference between dysphonic males and normal males and significant difference between dysphonic females and normal females in terms of fundamental frequency for speech.

***Average Fundamental Frequency (Fo/Hz)***

Average value of all extracted period-to-period fundamental frequency values. Voice break areas are excluded.

Fo is computed from the extracted period-to-period pitch data as :

1 . *Absolute jitter/seq/or/jita:*

$$Fo = \frac{1}{N-1} \sum_{i=1}^{N-1} Fo^{(i)}$$

Where

$$Fo^{(i)} = \frac{(1)}{To^{(i)}} \quad \text{Period-to-period fundamental frequency}$$

$To^{(i)}$, i = 1 , 2, ... N extracted pitch period data

N = PER, Number of extracted pitch periods.

Biswajit (1995) found no significant difference between normals and dysphonics in terms of average fundamental frequency.

*Highest Fundamental Frequency (HFo/Hz)*

The greatest of all extracted period-to-period fundamental frequency values. Voice break areas are excluded. It is computed as

$$Fhi = Max. [Fo^{(i)}), i = 1, 2, ..., N$$

***Lowest Fundamental Frequency (LFo/Hz)***

The lowest of all extracted period-to-period. It i computed as :

$$F1o = Min(Fo^{(i)}), i = 1, 2, ...., N$$

The lowest fundamental within the defined period is extracted and displayed as Flo. However, the pitch extracted range is defined to either search for periods from 70-625 Hz or 200-1 000 Hz. Therefore, the 'high' range will not determine a fundamental under 200 Hz.

Biswajit (1995) found no significant difference between male normals and dysphonics in terms of lowest fundamental frequency for vowels/a//i/and /u/, but a significant difference in sentences was noted. Anitha (1 994) found that the mean of the 'lowest-Fo' of/a/ was lower when compared to/i/ and /u/ and the mean of the 'lowest Fo' was lower when compared to /a/, /i/ and /u/ is the case of normal males. In the case of normal females, the order of increase in the mean 'lowest Fo' were/a/ (232.69), N (242.73), /u/ (249.51) and sentence (173.27). However in the case of dysphonic males and females the mean value of 'lowest Fo' of sentence is lowest when compared to /a//i/ and /u/.

***Frequency Range in Phonation and Speech***

Humans are capable of producing a wide variety of acoustic signals. The patterned variations of pitch over linguistic units of differing length (syllables, words phrases) yield in critical prosodic features namely intonation (Freeman, 1982).

Variations in fundamental frequency and the extent of range used also relate to the intent of the speakers (Fairbanks and Pronovost, 1939). More specifically, the spread of frequency range used corresponds to the mood of the speaker, that is, as Skinner (1935) reports, cheerful animated speech exhibits greater range than serious, thoughtful speech.

Jayaram (1975) reported that a significant difference in the frequency range was obtained for males and females in the normal group at both the levels of significance, while the males and females in the dysphonic group differed only at 0.05 level of significance.

Hudson and Holbrook (1981) studied the fundamental vocal frequency range in reading, in a group of young black adults, age range from 18-29 years. Their results showed a mean range from 81.95-158.50 Hz in males and from 139.05 Hz to 266.10 Hz in females.

Nataraja (1986) found that the frequency range did not change much with age i.e. in the age range 16-45 years. He also found that females showed a greater frequency range than males in both phonation and speech. Gopal (1986) from a study of normal males from 16-65 years, reported slightly lower frequency range in speech.

Hanson, et al. (1983), suggested that majority of phonatory dysfunctions are associated with abnormal and irregular vibrations lead to the generation of random acoustic energy, i.e. noise, fundamental frequency and intensity variations. This random energy and a periodicity of Fo is perceived by human ears as hoarseness. Hence, the spectral, intensity and Fo parameters are more appropriate in quantifying phonatory dysfunctions. The frequency related parameters are the most rugged and sensitive in detecting anatomical and sensitive in detecting anatomical physiological changes in the larynx (Hanson, et al. 1983).

Preethi (1998) found significant differences between normals and dysphonics in terms of frequency range in phonation and speech. No significant difference between males and females both in normal and dysphonic groups in terms of frequency range in phonation and speech was also reported.

***Standard Deviation of Fundamental Frequency (STD)/Hz)***

Standard deviation of all extracted period-to-period fundamental frequency values. Voice break areas are excluded.

$$STD = \sqrt{\frac{1}{N} \sum_{i=1}^{N-1} (Fo-Fo^{(i)})^2}$$

where,

$$Fo = \frac{1}{N} \sum_{i=1}^{N-1} Fo^{(i)}$$

$$Fo = \frac{1}{To^{(i)}}$$ - Period to period of values

$To^{(i)}$, i = 1, 2, ...., N extracted pitch period data

N = Number of extracted pitch periods.

Anitha (1 994) found that in the case of dysphonic males and females, the mean value of STD was higher for sentence when compared to /a//i/ and /u/. Biswajit (1 995) found no significant difference between male normals and dysphonics in terms of standard fundamental frequency.

***Phonatory Fundamental Frequency range (PFR)/Semitones)***

The range between Fhi and Flo expressed in number of semitones. The ratio of two consecutive semi-tones is equal to 1 2th root of 2.

First all frequencies of semitones Fst - F1, lc=1 , 2, ... are computed within the frequency range 55 Hz to 1055 Hz.

where $a = \sqrt[12]{2}$

$f1 = 55$ Hz, $f2 = 1055$ Hz and $f1 < Fst^{k} < f2$

Anitha (1 994) found that in the case of dysphonic males and females, the mean PFR value of sentence was higher than /a/ /i/ and /u/. But in the dysphonic males and females the PFR values of phonation of vowels /a/, /i/ and /u/ were higher than that of normal males and females respectively.

*Fo - Tremor Frequency (FFTR)/Hz)*

The frequency of the most intensive low frequency Fo- modulating component in the specified Fo-tremor analysis range. If the corresponding FTRI value is below the specified threshold, the Fftr value is zero.

The method for frequency tremor analysis consists of the following :

A. Division of the fundamental frequency period-to-period (Fo) data into 2 sec windows at 1 sec step between. For every window, the following procedures apply.

1. Low-pass filtering of the Fo data at 30 Hz and down sampling at 400 Hz.
2. Calculation of the total energy of the resulting signal.
3. Subtraction of the DC component.
4. Calculation of an auto correlation function on the residue signal.
5. Division by the total energy and conversion to (%)
6. Extraction to the period of variation.
7. Calculation of Fftr corresponding to the period of variation found.

B. Computation of the average autocorrelator curve and average Fftr for all processed window.

Biswajit (1995) found significant difference between male normals and dysphonics in terms of fundamental tremor frequency for vowels */a/* and */i/* and a significant difference for /u/ and sentence.

*Absolute Jitter (Jita/usec)*

An evaluation of the period to period variability of the pitch period within the analysed voice sample. Voice break areas are excluded. Jita is computed as :

$$Jita = \frac{1}{N-1} \sum_{i-1}^{N-1} \left| To^{(i)} - To^{(i+1)} \right|$$

where, $To^{(i)}$, i = 1, 2, ... N extracted pitch period data
N = Number of extracted pitch periods.

Absolute jitter measures the very short term (cycle-to-cycle) irregularity of the pitch periods in the voice sample This measure is widely used in the research literature on voice perturbation (Iwata and Vonleden, 1970). It is very sensitive to the pitch variations occurring between consecutive pitch periods. However, pitch extraction errors may affect absolute jitter significantly.

The pitch of the voice can vary for a number of reasons, cycle-to-cycle irregularity can be associated with the inability of the vocal cords to support a periodic vibration for a defined period. Usually this type of variation is random. They are typically associated with hoarse voices.

Both Jita and Jitt represent evaluations of the same type of pitch perturbation. Jitfa is an absolute measure and shows the result in micro seconds which makes it dependent on the average fundamental frequency of voice. For this reason, the normative values on Jitta for men and women differ significantly. Higher pitch results into lower Jita. That's why, the Jitta value of two subjects with different pitch are difficult to compare.

Anitha (1 994) found mean absolute jitter values to be higher in normal males than in normal females and in case of dysphonic males and females there was significant difference with males having higher values for all the vowels and the sentence. Biswajit (1 995) stated that a significant difference was noticed between male normals and dysphonics in terms of absolute jitter for vowels /a/ /i/ /u/ and sentence. David (1998) found a significant difference between normals and dysphonics in terms of jitter mean fundamental frequency. A significant difference was also found between males and females with reference to normals and dysphonics.

*Jitter Percent (Jitt) (%)*

Relative evaluation of the period-to-period (very short term) variability of the pitch within the analysed voice sample. Voice break areas are excluded. It is computed

$$Jitt = \frac{\frac{1}{N-1} \sum_{i-1}^{N-1} \left| To^{(i)} - To^{(i)+1} \right|}{\frac{1}{N} \sum_{i=1}^{N} To^{(i)}}$$

where, $To^{(i)}$, $i = 1, 2, \ldots, N$ extracted pitch period data
$N = PER$, Number of extracted pitch periods.

Jitter percent measures the very short term cycle-to-cycle irregularity of the pitch period of the voice. Jitt is a relative measure and the influence of the average fundamental frequency of the subject is significantly reduced.

Anitha (1994) found *a* significant difference in terms of Jitter for phonation of vowels /a//V and sentence incase of dysphonic males and females but for normals no significant difference between males and females for vowels was noticed. Biswajit (1995) found no significant difference sentence male normals and dysphonics in terms of jitter percentage for vowels /a//i//u/ and sentence. David (1998) found a significant difference between normals and dysphonics in terms of percent jitter and a significant difference between males and females in terms of normals and dysphonics.

***Relative Average Perturbation (RAP) (%)***

Relative evaluation of the period-to-period variability of the pitch within the analysed voice sample with smoothing factor of three periods. Voice breaks areas are excluded. It is computed as :

$$RAP = \frac{\frac{1}{N-1} \sum_{i=2}^{N-1} \left| To^{(i-1)} + To^{(i)} + To^{(i-1)} \right|}{\frac{1}{N} \sum_{i=1}^{N} To^{(i)}}$$

where, $To^{(i)}$, $i = 1, 2, \ldots, N$ extracted pitch period data
$N = PER$, Number of extracted pitch periods.

Relative average perturbation measures the short term (cycle-to-cycle with smoothing factor of three periods) irregularity of the pitch period of the voice. The smoothing reduces the sensitivity of RAP to pitch extraction errors. However, it is less sensitive to the very short term period-to-period variations, but describes the short-term pitch perturbation of the voice very well.

The pitch of the voice can vary for a number oF reasons, cycle-cycle irregularity can be associated with the inability oF the vocal cords to support a periodic vibration with a defined period. Hoarse and/or breathy voices may have an increased RAP.

Anitha (1994) and Biswajit (1995) Found no significant difference between male normals and dysphonics in terms of relative average perturbation for vowels /a/ /i/ /u/ and sentence. David (1998) found no significant difference between normal and dysphonic in terms of jitter relative average perturbation and no significant difference between males and females.

***Pitch Period Perturbation Quotient (PPO %)***

Relative evaluation of the period-to-period variability of the pitch within the analysed voice sample with a smoothing factor of five periods. Voice break areas are excluded. PPQ is computed as,

$$APQ = \frac{\frac{1}{N-4} \sum_{r=1}^{N-4} \left| \frac{1}{5} \sum_{i=0}^{4} To^{(i+r)} - To^{(i+2)} \right|}{\frac{1}{N} \sum_{i=1}^{N} To^{(i)}}$$

where $To^{(i)}$, $i = 1, 2, \ldots, N$ extracted peak-to-peak amplitude data

$N$ = Number of extracted impulses.

PPQ measures the short-term (cycle-to-cycle with a smoothing factor o five periods) irregularity of the pitch period of the voice. The smoothing reduces the sensitivity of PPQ to pitch-extraction errors while it is less sensitive to period-to-period variations, it describes the short-term pitch purturbation of the voice very well. Hoarse and/or breathy voices may have an increased PPQ.

Anitha (1994) and Biswajit (1995) found no significant difference between male normals and dysphonics in terms of pitch perturbation quotient for vowels /a/ /i/ /u/ and sentence.

***Smoothed Pitch Period Perturbation Quotient (SPPQ %)***

Relative evaluation of the short or long term variability of the pitch period within the analysed voice sample at smoothing factor defined by the user. The factory setup for the smoothing factor is 55 periods. Voice break areas are excluded.

$$SPPQ = \frac{\frac{1}{N-Sf+i} \sum_{i=1}^{N-Sf+1} \left| \frac{1}{Sf} \sum_{r=0}^{Sf-1} To^{(i-r)} + To^{(i-m)} \right|}{\frac{1}{N} \sum_{i=1}^{N} To^{(i)}}$$

where, $To^{(i)}$, $i = 1, 2, \ldots, N$ extracted peak to peak amplitude data
$N$ = PER, Number of extracted impulses.
SF = Smoothing factor.

SPPQ allows the experiments define his own pitch perturbation measure by changing the smoothing factor from 1 to 99 periods. This is desirable because in the scientific literature researchers use pitch perturbation measures with different smoothing factor or without smoothing.

With a small smoothing factor, SPPQ is sensitive mostly to the short-term pitch variation of the voice impulses. With a smoothing factor of 1 (no smoothing), SPPQ is identical to Jitter percent (Jitt). It is very sensitive to the pitch variations occurring between consecutive pitch periods. Usually this type of variation is random. It is typical for hoarse voices. However, pitch extraction errors may affect jitter percent significantly.

With a smoothing factor of 3, SPPQ is identical to the relative average perturbation introduced by Koike (1973). With a smoothing factor of 5, SPPQ is identical to the pitch perturbation quotient introduced by Koike and Calcatera (1977). At high smoothing factors SPPQ correlates with the intensity of the long-term pitch period variations. The studies of patients with spasmodic dysphonia (Deliyski, et al. 1991) show that SPPQ with smoothing factor set in the range 45-65 period has increased values in case of regular long-term pitch variations (frequency voice tremors).

The SPPQ smoothing factory set-up is 55 periods. This set up allows using SPPQ as an additional evaluation of the frequency tremors in the voice. The intensity and the regularity of the frequency tremors can be assessed using SPPQ (55) in combination with VFo. The difference between VFo and SPPQ (55) is that VFo represents a general evaluation of the fundamental frequency (pitch) variation of the voice signal. The VFo value increases regardless of the type of pitch variation. Either random or regular short-term or long-term variations increase the value of VFo. However, SPPQ (55) is more sensitive to regular long term variations with a period near and above 55 pitch periods. If both SPPQ (55) and VFo are low, the intensity of pitch variations in the voice signal is very low. If VFo is high but SPPQ (55) is low, there are pitch variations but not a long-term periodic one. If both SPPQ (55) and VFo are high, there is a long-term periodic pitch variation (most likely a frequency tremor).

Anitha (1994) and Biswajit (1995) found no significant difference between male normals and dysphonics in terms of smoothed pitch perturbation quotient for sentences but *a* significant difference for vowels /a//i//u/.

***Coefficient of Fo Variation VFo /%/***

Relative standard deviation of the fundamental frequency. It reflects, in general, the variation of Fo (short to long- term), within the analysed voice sample. Voice break areas are excluded.

$$VFo = \frac{\frac{1}{N} \sum_{i-1}^{N} \sqrt{\frac{1}{N} \sum_{j=1}^{N} \left[ \frac{1}{N} Fo^{(i)} - Fo^{(i)} \right]^{(2)}}}{\frac{1}{N} \sum_{i=1}^{N} Fo^{(i)}}$$

Where, $Fo^{(i)}$, **i = 1, 2, .... N extracted peak-to-peak amplitude data**

N = PER, number of extracted pitch periods

VFo reveals the variations in the fundamental frequency. The VFo value increases regardless of the type of pitch variation. Either random or regular short-term or long-term variations increases the value of VFo. Because the sustained phonation normative thresholds assume that the Fo should not change, any variations in the fundamental frequency are reflected in VFo. These changes could be frequency tremors or non-periodic changes, very high jitter or simply rising a falling pitch over the analysis length. Anitha (1994) and Biswajit (1995) found no significant difference between male normals and dysphonics in terms of coefficient of fundamental frequency variation for vowels /a//i//u/ and sentences.

## INTENSITY PARAMETERS

### *Shimmer in dB (ShdB) /dB/*

Evaluation in dB of the period-to-period (very short-term) variability of the peak-to-peak amplitude within the analysed voice sample. Voice break areas are excluded. ShdB is computed as,

$$ShdB = \frac{1}{n-1} \sum_{i=1}^{N-1} \left| 20 \log \left( \frac{A^{(i+1)}}{A^{(i)}} \right) \right|$$

where, $A^{(i)}$, $i = 1, 2, \ldots N$ extracted peak-to-peak amplitude
$N$ = Number of extracted impulses.

Shimmer in dB measure the very short term (cycle-to-cycle) irregularity of peak-peak amplitude of the voice. This measure is widely used in the research literature on voice perturbation (Iwata and Von Leden, 1970). It is very sensitive to the amplitude variation occurring between consecutive pitch periods. However, pitch extraction errors may affect shimmer percent significantly.

The amplitude of the voice varies due to several factors. Cycle-to-cycle irregularity of amplitude can be associated with the inability of the vocal folds to support a periodic vibration for a defined period and with the presence of turbulent noise in the voice signal usually, this type of variation is random. It is typically associated with hoarse and breathy voices. APQ is the preferred measurement for Shimmer because it is less sensitive to pitch extraction errors while still providing a reliable indication of short-term amplitude variability in the voice.

Both shim and ShdB are relative evaluations of the same type of amplitude perturbation but they use different measures for the result-percent an dB.

Anitha (1994) and Biswajit (1995) found no significant difference between male normals and dysphonics in terms of shimmer in dB for vowels /a//i/ /u/ and sentence. David (1998) found significant difference between normal and dysphonics in terms of shimmer in dB and a significant difference between males and females in shimmer in dB.

***Shimmer Percent (%)***

Relative evaluation of the period-to-period (very short term) variation of the peak-to-peak amplitude within the analysed voice sample. Voice break means are excluded.

$$Shim = \frac{\frac{1}{n-1}\sum_{i=1}^{N-1}\left|A^{(i)} - A^{(i+1)}\right|}{\frac{1}{N}\sum_{i=1}^{N}A^{(i)}}$$

where, $A^{(i)}$, $i = 1, 2, \ldots N$ extracted peak-to-peak amplitude
$N$ = Number of extracted impulses.

Shimmer percent measure the very short term (cycle-to-cycle) irregularity of the peak-to-peak amplitude of the voice. Anitha (1994) reported that the mean value of shimmer in 96 for /u/ was lower when compared to /a/ /i/ and the sentence. The mean value for sentence was highest when compared to /a/ /]/ and /u/ for both normal males and females. For dysphonic males and females the mean value of sentence was highest when compared to /a/ /i/ and /u/. It was also found that females had lower values of shimmer in 96 when compared to males for both the groups. Biwajit (1995) found no significant difference between male normals and dysphonics in terms of shimmer in percent for vowel /a/ and sentence but a significant difference for vowels /i/ and /u/.

### *Amplitude Perturbation Quotient (APQ) (%)*

Relative evaluation of the period-to-period variation, variability of the peak-to-peak amplitude within the analysed voice sample at smoothing of 11 periods. Voice break areas are excluded.

$$APQ = \frac{\frac{1}{N-4} \sum_{i=1}^{N-4} \left| \frac{1}{5} \sum_{r=0}^{4} A^{(i+r)} - A^{(i+2)} \right|}{\frac{1}{N} \sum_{i=1}^{N} A^{(i)}}$$

where, $A^{(i)}$, $i = 1, 2, \ldots N$ extracted peak-to-peak amplitude

$N$ = Number of extracted impulses.

APQ measures the short-term (cycle-to-cycle with smoothing factor of 11 periods) irregularity of the peak-to-peak amplitude of the voice. While it is less sensitive to the period to period amplitude variations it still describes the short-term amplitude perturbation of the voice very well breathy and hoarse voice usually have an increased APQ. APQ should be regarded as the preferred measurement for shimmer in MDVP.

Anitha (1 994) found a significant difference in the mean APQ values for vowels /a//i/ /u/ and sentence between the normal males and females and between the dysphonic males and females. Biswajit (1 995) found no significant difference between male normals and dysphonics in terms of amplitude perturbation quotient for sentence in a significant difference between male normals and dysphonics for vowels /a//i/ and /u/. David (1 998) found no significant difference between normal and dysphonic females in terms of shimmer amplitude perturbation quotient. A significant difference between normal and dysphonic males in terms of APQ values and between males and females in both normal and dysphonics was also reported.

***Smoothed Amplitude Perturbation Quotient (SAPQ (%)***

Relative evaluation of the short or long term variability of the peak-to-peak amplitude within the analysed voice sample at smoothing factor defined by the user. The factory set up for the smoothing factor is 5 5 periods (providing relatively long-term variability, the user can change this value as desired). Voice break areas are excluded.

$$SAPQ = \frac{\frac{1}{N-Sf+1} \sum_{i=1}^{N-Sf+1} \left| \frac{1}{Sf} \sum_{r=0}^{Sf-1} A^{(i+r)} - A^{(i+m)} \right|}{\frac{1}{N} \sum_{i=1}^{N} A^{(i)}}$$

where, $A^{(i)}$, $i = 1, 2, \ldots N$ extracted peak-to-peak amplitude

N = Number of extracted impulses.

SF = Smoothing factor.

SAPQ allows and Biswajit (1995) found no significant difference between male normals and dysphonics in terms of smoothed amplitude perturbation quotient for sentence and a significant difference for vowels/a//i/ and /u/ between male normals and dysphonics.

***Coefficient of Amplitude Variation (VAm) /%/***

Relative standard deviation of peak-to-peak amplitude. It reflects in general to peak-to-peak amplitude variations (short to long term) within the analysed voice sample, voice break areas are excluded. VAm is computed as ratio of the standard deviation to the average value of the extracted peak-to-peak amplitude data as

$$VAm = \frac{\sqrt{\frac{1}{N}\sum_{i=1}^{N}\left[\frac{1}{N}\sum_{j=1}^{N}A^{(j)} - A^{(i)}\right]^2}}{\frac{1}{N}\sum_{i=1}^{N}A^{(i)}}$$

where, $A^{(i)}$, $i = 1, 2, \ldots N$ extracted peak-to-peak amplitude

$N$ = Number of extracted impulses.

VAm reveals the variations in the cycle-to-cycle amplitude of the voice. The VAm value increases regardless of the type of amplitude variation. Either random or regular short-term or long-term variation increases the value of VAm.

Anitha (1994) found that in normal males and females the mean values of VAm of /i/ and /u/ were lower than /a/ and sentence and the mean value of sentence was highest when compared to /a//i/ and /u/. In case of dysphonic males the mean value for sentence was higher when compared to /a//i/ and /u/. In case of dysphonic females the mean value for sentence was highest when compared to /a/ /i/ and /u/ and the mean value for /u/ was lower when compared to /a/ and sentences. Biswajit (1995) found no significant difference between normals and dysphonics in terms of coefficient of amplitude variation for sentences but a significant difference for vowels /a/ /i/ and /u/.

***Frequency and Intensity Parameters***

***Extent and Speed of Fluctuation in Frequency and Intensity***

The extent and speed of fluctuation in frequency and intensity are also measures of fundamental frequency and intensity variation measurements. The fluctuations in frequency and intensity in phonation sample may include the physiological (Neuromucular) or pathological changes in the vocal mechanism.

i) *Extent of Fluctuation in Fundamental Frequency*

The extent of fluctuation as defined as the percent score of the ratio of the peak to peak value of fluctuation (Fo) to the mean fundamental frequency (Fo).

Preethi (1998) reported no significant difference between males and females both in normals and dysphonics and significant differences between normals and dysphonics in terms of fluctuation in fundamental frequency.

ii) *Speed of Fluctuation in Fundamental Frequency*

This has been defined as the peak to peak value in decibels measured on an average amplitude display.

iii) *Extent of Fluctuation in Intensity*

This has been defined as the peak to peak value in decibels measured on an average amplitude display.

Preethi (1998) found no significant difference between males and females in both normal and dysphonic group in terms of extent of fluctuation in intensity and speed of fluctuation in intensity.

iv) *Speed of Fluctuation of Intensity*

This was defined as the number of positive peaks on an amplitude display with in one sec. Peaks of 3 dB or greater from adjacent trough have been counted.

The results of Kim, et al. (1982)'s study was indicated that among the above mentioned acoustic parameters significant differences were found between the control and the diseased groups in terms of Fluctuation of fundamental frequency. Vanaja (1986), Thamar (1991) and Suresh (1991) have reported that as the age increases there was increase in fluctuations in frequency and intensity of phonation and this difference was more marked in females.

Nataraja (1986) found that speed of fluctuation in fundamental frequency and extent of fluctuation in intensity parameters were sufficient to differentiate the dysphonics from the normal. Preethi (1908) reported that both males and females showed significant differences in both normals and dysphonic groups in terms of speed of fluctuation in intensity.

## OTHER PARAMETERS

### . *Signal to Noise Ratio (SNR)*

Fourier expansion to separate the noise from the periodic components was used by Jojima, et al. (1980) to compute the signal to noise ratio as an objective estimate of hoarseness. The resolution of voice into signal and noise components may not be satisfactory, since only three pitch periods used in the fourier transform i.e. one third of the Fourier components was counted as the signal (Hiraoka et al. 1984). This method has theoretical limitations also with regard to the accuracy of estimated noise levels, since the fourier co- efficients derived from a signal with duration T provides estimates of the noise components only at multiple frequencies of 1/T whereas the noise has a continuous frequency spectrum. This method is too complex and time consuming to apply to clinical use.

### *Normalized Noise Energy*

Kasuya et al. (1 986) proposed normalized noise energy (NNE) which was considered to be superior to other measures of spectral noise. The NNE, was automatically computed from the voice signals using an adaptive comb filtering method performed in the frequency domain. Experiments with the voice samples have show that NNE is especially effective for detecting the glottic cancers. Since the NNE measures primarily the turbulence noise caused by the closing in sufficiency of the glottis during the phonation. It is very useful in the detection of these diseases. But NNE is not sensitive to the noise caused by irregular vibration motion of the vocal folds. Hence NNE is not an effective measure for those laryngeal conditions which produce hoarseness because of a periodicity of vocal fold movements.

***Noise to Harmonic Ratio (NHR)***

Average ratio of the in harmonic spectral energy in the frequency range 1 500-4500 Hz to the harmonic spectral energy in the frequency range 70-4500 Hz. This is general evaluation of noise present in the analysed signal. NHR is computed using a pitch synchronous frequency domain method. In general terms, the algorithm functions as follows:

A. Divides the analysed single into windows of 81 .92 ms (4096 points at 50 kHz sampling rate or 2048 at 25 kHz). For every windows the following steps apply.

1 . Low pass filtering 6 kHz (order 22) with Hamming window, down sampling of the single data down to 1 25 kHz and conversion of the real signal into an analytical one using the Hilbert transform.
2. 1024 points complex fast Fourier Transform (FFT) on the analytical signal corresponding to a 2048 - points FFT on real data.
3. Calculation of the power spectrum from the FFT.
4. Calculation of the average fundamental frequency within the window synchronously with the pitch extraction results.
5. Harmonic/inharmonic separation of the current spectrum synchronously with the current window fundamental frequency.
6. Computation of the noise-to-harmonic ratio of the current window. NHR is the ratio of the inharmonic (1 500-4500 Hz) to the harmonic spectral energy (70-4500 Hz).

B. Computes the average values of NHR for all previously processed windows.

Increased values of NHR are interpreted as increased spectral noise which can be due to amplitude and frequency variations (i.e. Shimmer and Jitter) Turbulent noise, subharmonic components and/or breaks which affects NHR globally measures the noise in the signal (includes contributions of jigger, shimmer and turbulent noise).

Anitha (1 994) found that the mean value of NHR of/u/ was greater when compared to the mean values of */a/,* /"/ and sentence in males and females of both the groups. It was also reported that the value of NHR increased in both dysphonic males and females when compared to normal males and females. Biswajit (1995) stated that there is no significant difference between male normals and dysphonics in terms of noise to harmonic ratio for vowels /a//i//u/ and sentence.

***Voice Turbulence Index (VTI)***

Average ratio of the spectral in harmonic high frequency energy in the range 2800-5800 Hz to the spectral harmonic energy in the range 70-4500 Hz in the areas of the signal where the influence of the frequency and amplitude variations, voice breaks and subharmonic components are minimal. VTI measures the relative energy level of high frequency noise. VTI is computed using a pitch synchronous frequency domain method. The algorithm consists of the following steps :

A. Selects upto four but atleast two 81 .92 msec, windows where the frequency and amplitude perturbations are lowest for the signal. These windows are located in different areas of signal and don't include voice breaks and subharmonic components.

For every window, the following steps apply :

1. Low-pass filtering at 6 kHz.
2. Down sampling at 12.5 kHz.
3. Conversion of the real signal to analytical one.
4. Computation of a 1024 points complex fast fourier transform on the analytical signal.
5. Computation of power spectrum from the FFT.
6. Calculation of the average fundamental frequency within the window.
7. Harmonic/inharmonic separation of the current spectrum synchronously with the current window Fo.
8. Computation of the TI for every window, VTI is the ratio of the spectral in harmonic high frequency energy (2800-5800 Hz) to the spectral harmonic energy (70-4500 Hz).

B. Calculate the average VTI values for all processed windows. VTI measures the relative energy level of high-frequency noise. Biswajit (1995) reported of no significant difference between male normals and dysphonics in terms of voice turbulence index for vowels /a//i//u/ and sentence.

VTI mostly correlates with the turbulence caused by incomplete or loose adduction of the vocal folds. VTI, unlike NHR, analyses high frequency components to extract an acoustic correlate to "breathiness". However, it is unlikely that users will find a one-to-one correspondence between their perceptual impression of a voice and this acoustic analysis. However, VTI is a new attempt to compute a parameter which correlates with breathiness. Because VTI is a new parameter, normative values cannot be found in the professional literature.

***Soft Phonation Index (SPI)***

Average ratio of the lower-frequency harmonic energy in the range of 70-1 600 Hz to the higher frequency harmonic energy in the range 1 600-4500 Hz.

SPI is computed using a pitch synchronous frequency domain method. The algorithm does the following procedures.

A. *Divides the analysed signal into windows of 81 -92 ms.*

For every one of these windows, the following steps apply :

1 . Low-pass filtering at 6 kHz order 22 with Hamming window, down sampling of the signal data down to 12.5 Hz and conversion of the real signal ratio analytical one using Hilbert transform.
2. 1 024 points complex fast fourier transform on the analytic signal.
3. Computation of the power spectrum from the FFT.
4. Calculation of the average Fo within the window synchronously with the pitch extraction results.
5. Harmonic/inhdrmonic separdtion of the current spectrum synchronously with the current window Fo.
6. Computdtion of SPI of the current window. SPI is a ratio of the lower-frequency' (70-1 600 Hz) to the higher frequency (1 600-4500 Hz) hdrmonic energy.

B. *Computes the average of SPI for all previously processed windows.*

SPI can be thought of as an indicator of how completely or tightly the vocal folds adduct during phonation. Increased value of SPI is generally an indication of loosely or incompletely adducted vocal folds during phonation. However, it is not necessarily an indication of a voice disorder. Similarly, patients with "pressed" phonation may likely have a "normal" SPI though their pressed voice characteristic may not be desirable. Therefore, a high SPI value is not necessarily bad, nor a low SPI value necessarily good. Subjects with glottal chinks (determined stroboscopically) or with high phonatory air flow rates often exhibit an increased SPI. Spectral analysis will show a well defined higher formants when SPI is low, and less well defined when SPI is high. SPI is very sensitive to the vowel formant structure because vowels with lower high frequency energy will result in higher SPI, only values computed for the same vowel can be compared.

Increased SPI values may be due to a number of factors. The subject may have a "soft" phonation because of a voice or speech disorder and may not be able to strongly adduct his vocal folds. However, the subject may naturally speak with a softer "attack" and hence have an elevated SPI. Psychological stress could also be a factor that may increase SPI. Another important factor is the amplitude of the sustained vowel, if the subject phonates softly, SPI may be high. Anitha (1994) and Biswajit (1995) reported no significant difference between male normals and dysphonics in terms of soft phonation index for vowels /a//i//u/ and sentences. Arun Biran (1995) reported of no significant difference in SPI between males and females of different age groups in terms of phonation in both normal and dysphonic groups.

. ***Frequency Tremor Intensity Index (FTRI) /%/***

Average ratio of the frequency magnitude of the most intensive low-frequency modulating component (Fo-tremor) to the total frequency magnitude of the analysed voice signal.

The method for frequency tremor analysis consists of the following steps:

A. Division of the fundamental frequency period-to-period (Fo) data into 2 secs windows. For every window, the following procedures apply.

1. Low-pass filtering of the Fo data at 30 Hz and down sampling at 400 Hz.
2. Calculation of the total energy of the resulting signal.
3. Subtraction of the DC component.
4. Calculation of an autocorrelation function on the residue signal.
5. Division by total energy and conversion to percent.
6. Extraction of the period of variation.
7. Calculation of Fftr and Ftri corresponding to the period of variation found.

B. Computation of the average autocorrelation curve and average FTRI for all processed windows.

The algorithm for tremor analysis determines the strongest periodic frequency and amplitude modulation of voice. Tremor has both frequency and amplitude components (i.e. the Fo may vary and/or the amplitude of the signal may vary in a periodic manner). Tremor frequency provides the rate of change with Fftr providing the rate of periodic tremor of the frequency and Fatr providing the rate of change of the amplitude. The program will determine the Fftr and Fatr of any signal if the magnitude of these tremors is above a low threshold of detection. Therefore, the magnitude of the frequency tremor and the magnitude of the amplitude tremor are more significant than the respective frequencies of the tremor.

***Amplitude Tremor Intensity Index (ATRI) (%)***

Average ratio of the amplitude of the most intense low-frequency amplitude modulating component to the total amplitude of the analysed voice signal.

The method for computation is same as FTRI except that here the peak-to-peak amplitude data has been taken into consideration instead of Fo data.

Biswajit (1995) found no significant difference between male normals and dysphonics in terms of amplitude tremor intensity index for vowel /u/ and sentence a significant difference for vowels /a/ and /i/. Anitha (1994) reported that in normal males the mean value of ATRI of /u/ was lesser when compared to /a//i/and sentence. The man value of sentence was highest when compared to /a//i/ and /u/. In case of normal females the mean values of/i/ and /u/ were lesser when compared to /a/ and sentence. The mean value of sentence was highest when compared to /a//i/ and /u/.

In case of dysphonic males and females the mean value of sentence was highest when compared to /a//i/ and /u/.

***. Degree of Voice Breaks (DVB)( %)***

Ratio of the total length of areas representing voice breaks to the time of the complete voice sample.

$$DVB = \frac{t1 + t2 \ldots + tn}{Tsam}$$

where, t1 , t2 ......tn lengths of the 1st, 2nd ... voice break

Tsam - length of analysed voice data samples.

DVB does not reflect the pauses before the first and after the last voiced areas of the recording. It measures the ability of the voice to sustain uninterrupted voicing. The normative threshold is o' because the normal voice, during the task of sustaining voice, should not have any voice break areas. In case of phonation with pauses (such as running speech, voice breaks, delayed start or earlier end of sustained phonation), DVB evaluates only the pauses between the voiced areas.

***Degree of Sub-harmonic Components (DSH) /%/***

Relative evaluation of sub-harmonic to Fo components in the voice sample. DSH is computed as a ratio of the number of autocorrelation segments where the pitch was found to be sub-harmonic of the real pitch (NSH) to the total number of autocorrelation segments. The degree of sub-harmonic components in normal voices should be equal to zero. It is expected to increase in voices where double or triple pitch periods replace the fundamental in certain segments over the analysis length. These effects are typical for diplophonic voices and voices with glottal fry. The experimental observation of patients with functional dysphonia or neurogenic voice disorders may show increased values of DSH.

Biswajit (1 995) reported of significant difference between male normals and dysphonics in terms of degree of subharmonic components for vowels/a//i/ and

/u/ and sentence. The results of this study was in accordance with the study done by Anitha (1994)

***Degree of Voiceless (DOV)/%/***

Estimated relative evaluation of non-harmonic areas (where Fo cannot be detected) in the voice samples. DOV is computed as a ratio of the number of autocorrelation segments where an unvoiced decision was made to the total number of auto-correlation segment. DOV measures the ability of the voice to sustain uninterrupted voicing. The normative threshold is o' because of normal voice, in the defined task of sustaining voicing, should not have any voiceless segments. In case of phonation with pauses (such as running speech, voice breaks, delayed start or earlier end of sustained phonation), DOV also evaluates the pauses before, after and/or between the voiced areas.

Biswajit (1995) reported significant difference between male normals and dysphonics in terms of degree of voiceless for vowels/a//i//u/ and sentence. Anitha (1 994) reported that the degree of voiceless was higher for sentence than in the phonation of vowels /a/ /i /u/ in all the four groups - normal males, normal females, dysphonic males and dysphonic females.

***Number of Voice Breaks (NVB)***

Number of times the fundamental period was interrupted during the voice sample (measured from the first detected period to the last period). NVB does not reflect the pauses before the first and after the last voiced areas of the recording. However, like NVB, it measures the ability of the voice to sustain uninterrupted voicing. The normative threshold is o' because of normal voice, during the task of sustaining voice, should not have any voice breaks. In cases of phonation with pauses (such as running speech, voice breaks, delayed start or earlier end of sustained phonation) NVB evaluates only the pauses between the voiced areas.

Biswajit (1995) and Anitha (1994) reported of the presence of voice break areas in phonation and sentence in dysphonics and in case of normal subjects the number of voice break areas in phonation of vowels were new, but in sentence it was present.

***Number of Sub-Harmonic Segments (NSH)***

Number of autocorrelation segments where the pitch was found to be a sub-harmonic of Fo. The number of sub-harmonic components in normal voices should be equal to zero. It is expected to increase in voices where double or triple pitch period replaces the fundamental in certain segments over the analysis length. These effects are typical for diplophonic voices and voices with glottal fry.

Thus to conclude a number of parameters can be used for evaluation of voice quality and these parameters can be evaluated using software programs.

The three software programs routinely being used in the "Speech Science" Department of All India Institute of Speech and Hearing, Mysore for objective evaluation of voice quality for voice disordered patients are :

1. Multidimensional Voice Profile (developed and marketed by Kay Elemetrics Inc. NJ).

2. Vaghmi Software (developed by Voice and Speech Systems, Bangalore).
3. Dr.Speech Software (developed by Tiger DRS).

1. *Multidimensional Voice Profile*

This is a computer based program which extracts several parameters of voice. This program options acquires, analyses and displays upto 33 voice parameters from a single vocalisation. The 33 extracted parameters are available as a numerical file or they can be displayed graphically in a data base. The 33 parameters can be grouped into 8 groups for analysis.

1. Fundamental frequency related measurements.
2. Long and short term frequency perturbation measurements.
3. Long and short term amplitude perturbation measurements.
4. Noise related measurements.
5. Tremor related measurements.
6. Subharmonic components related measurements.
7. Voice break related measurements.
8. Voice irregularity based measurements.

This software program has been used by various researchers for evaluation of voice disorders.

Biswajit (1995), Anitha (1994) had used this software program for analysis of voice disorders. Aparna (1996) used MDVP for multidimensional analysis of voice in hearing-impaired population. Multidimensional voice analysis in children has been conducted by Arun Biran (1995).

2. *Vaghmi Software*

This is a software program developed by the Voice Speech Systems, Bangalore and provides for both frequency and intensity related measurements. The parameters which were extracted in the present study with the help of this software were :

1. Mean fundamental frequency
2. Maximum fundamental frequency
3. Minimum fundamental frequency
4. Extent of fluctuation in frequency
5. Speed of fluctuation in frequency
6. Extent of fluctuation in intensity
7. Speed of fluctuation in intensity

Various studies have been conducted using this software program.

Preethi (1998) has evaluated the "Factors in normal and abnormal voice' using Vaghmi software.

Temporal and acoustic analysis of speech of Kannada speaking hearing-impaired children using Vaghmi software was done by Jayaprakash (1998). Divya (1996) analysed the voice before and after treatment with Vaghmi software.

3. *Dr. Speech*

This software program has been developed by Tiger DRS and in the present study has been used to extract the following parameters.

1. Habitual fundamental frequency
2. Maximum fundamental frequency
3. Minimum fundamental frequency
4. Fundamental frequency tremor
5. Standard deviation of fundamental frequency
6. Jitter percent
7. Shimmer percent
8. Normalised noise energy
9. Harmonics to noise ratio
10. Signal to noise ratio
11. Amplitude tremor

Thus as can be noted there have been various studies using these sofware programs independently for evaluation of voice quality. However, there has been no study which has compared all the three software programs for evaluation of the voice quality.

Thus, the purpose of the present study is to compare three different software programs (MDVP, Vaghmi and Dr.Speech) in terms of their efficacy of evaluation of voice disorders .

# METHODOLOGY

Over the years, the development of technology has permitted the analysis and measurement of various aspects of vocal function. There have been various objective approaches developed such as stroboscopy, electroglottography, inverse filtering, computer based methods etc. for analysis of vocal function. In computer based methods, there are many software programs which are designed to extract different parameters of voice. As a result, the speech-language pathologist often has difficulty, in deciding as to which the software is effective in evaluation of voice disorders among the various software. The present study thus attempts to evaluate three softwares in terms of their ability to evaluate the voice disorders. Thus the purpose of the present study was to compare three different software programs :

1 ) Multidimensional voice program :

(developed and marketed by Kay Elemetrics Inc., NJ).

2) Vaghmi software :

(developed by Voice and Speech Systems, Bangalore).

3) Dr.Speech software

(developed by Tiger DRS) in terms of their evaluation of voice quality).

The following table indicates the parameters which were evaluated by the three different softwares.

( + ) indicates that the parameter is evaluated by the software and

(-) indicates that the parameter is not evaluated by the software.

Table-3 : Shows the parameters which were evaluated by the three different softwares.

| Parameters | MDVP | Vaghmi | Dr.Speech |
|---|---|---|---|
| **FREQUENCY PARAMETERS** | | | |
| 1 . Habitual Fundamental Frequency | - - | | + |
| 2. Average Fundamental frequency | + | + | - |
| 3. Average Pitch Period | + - | | - |
| 4. Highest Fundamental Frequency | + | + | + |
| 5. Lowest Fundamental Frequency | + | + | + |
| 6. Standard Deviation of Fundamental Frequency | + - | | + |
| 7. Phonatory Fundamental frequency Range in Semitones | + - | | - |
| 8. Fundamental Frequency Tremor Frequency | + - | | + |
| 9. Amplitude Tremor Frequency | + - | | - |
| 10. Absolute Jitter | + - | | - |
| 1 1 . Jitter Percent | + - | | + |
| 1 2. Relative Average Perturbation | + - | | - |
| 1 3. Pitch Period Perturbation Quotient | + - | | - |
| 1 4. Smoothed Pitch Period Quotient | + - | | - |
| 1 5. Co-efficient of Fo Variation | + - | | - |
| **INTENSITY PARAMETERS** | | | |
| 1 . Shimmer in dB | + - | | _ |
| 2. Shimmer Percent | + - | | + |
| 3. Amplitude Perturbation Quotient | + - | | - |
| 4. Smoothed Amplitude Perturbation Quotient | + - | | - |
| 5. Coefficient of Amplitude Variation | + - | | - |
| 6. Amplitude Tremor | - - | | + |

| FREQUENCY AND INTENSITY PARAMETERS | | | |
|---|---|---|---|
| 1 . Extent of Fluctuation in Fundamental Frequency | - | + | - |
| 2. Speed of Fluctuation in Fundamental Frequency | - | + | - |
| 3. Extent of Fluctuation in Intensity | - | + | - |
| 4. Speed of Fluctuation of Intensity | -- | + | - |
| **OTHER PARAMETERS** | | | |
| 1 . Signal to Noise Ratio | - | - | + |
| 2. Normalised Noise Energy | - | - | + |
| 3. Noise to Harmonic Ratio | + | - | - |
| 4. Voice Turbulance Index | + | - | - |
| 5. Soft Phonation Index | + | - | - |
| 6. Frequency Tremor Intensity Index | + | - | - |
| 7. Amplitude Tremor Intensity Index | + | - | - |
| 8. Degree of Voice Breaks | + | - | - |
| 9. Degree of Subharmonic Components | + | - | - |
| 1 0. Degree of Voiceless | + | - | - |
| 1 1 . Number of Voice Breaks | + | - | - |
| 1 2. Number of Subharmonic Segments | + | - | - |
| 1 3. Harmonics to Noise Ratio | - | - | + |

**PROCEDURE**

Instrumentation : The following instruments were used in the tudy

i) Dynamic microphone

ii) Preamplifier

iii) Sony tape deck(TCFXi 70)

iv) Speech interface unit

v) Meltrack CR-X90 audio cassette

vi) Philiamp 60

vii) Headphones (Sukawa)

viii) PC-AT (486DX Vaghmi software)

ix) PC 486-SX MDVP software

x) PC II with 350 MHz speed Dr.Speech software.

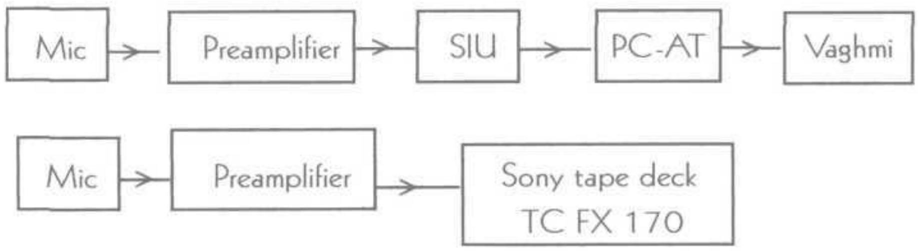
***Subjects***

30 normals (15 males, 15 females) in the age range of 18 to 25 years and 30 dysphonics (15 males, 15 females) in the age range of 18 to 65 years were studied. The group of dysphonics were chosen from among the patients who visited the All India Institute of Speech and Hearing with a complaint of voice problem. These cases were diagnosed as having voice disorder after the routine speech science, speech pathology, otolaryngological and psychological examinations. The subjects of the normal group had no apparent speech, hearing or ENT problem and were judged as normal by qualified speech pathologists.

***Recording the Data***

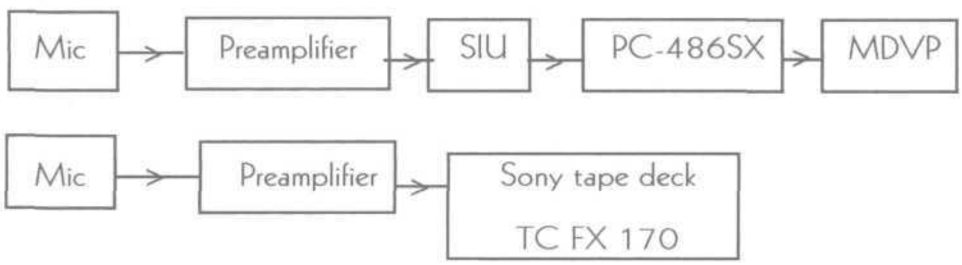
I. Phonation of /a/, */i/,* /u/ for about 5 to 6 seconds and three Kannada sentences /idu papu/ /idu koti/ /idu kempu banna/ spoken by each subject were recorded on an audio cassette and also on the computer using MDVP software.

Block diagram showing instrumentation for Vaghmi Software



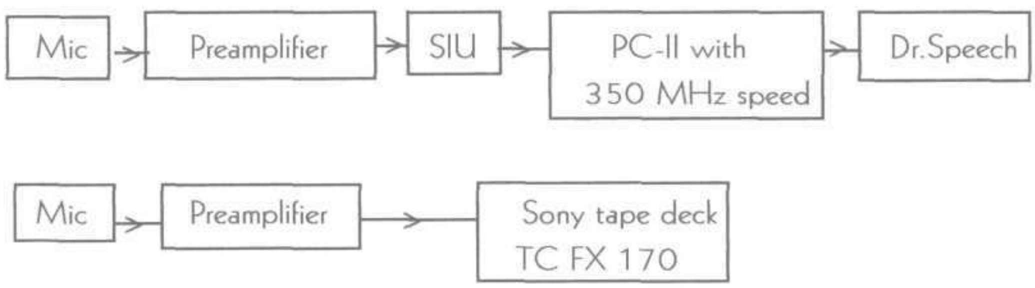

The following instructions were given to each subject prior to recording : "Say/a/ three times with comfortable loudness until I signal you to stop". Similar instructions were given for recording of vowels / i / and / u / and for speech. The microphone was kept 4-6 inches from the subjects mouth. In this set-up, the recorded samples ie, phonation of/a//i//u/ and speech samples were individually fed through the speech interface unit (SIU) (12 bit A/D converter) and digitized at a sampling rate of 1600 kHz and stored at the hard disc of the computer. Before digitizing, each sample was passed through filter at 7.5 KHz with a roll off of 78 dB/octave. The level indicators of the speech interface unit was used to monitor the intensity level to avoid any distortion while digitizing the signal. Each sample was analysed using Vaghmi (Voice-Inton) software. A total of eight parameters were obtained using Vaghmi software (Refer Table 3)

Block diagram showing instrumentation for MDVP Software



The samples of phonation and speech produced by each subject which were recorded and stored on the cassette was used for this part of the experiment. In this set-up, the recorded sample i.e. phonation of /a/, /i/, /u/ and speech samples were fed through the speech interface unit and after digitization were fed into MDVP software. A total of 29 parameters were evaluated using MDVP software (refer to table 3).

Block diagram showing instrumentation for Dr.Speech Software




The samples of phonation and speech produced by each subject which were recorded and stored on the cassette was used for this part of the experiment. In this set-up, the recorded sample i.e. phonation of /a/, i/, /u/ and speech samples were fed through the speech interface unit and after digitization were fed into Dr.Speech software. A total of 1 1 parameters were evaluated using Dr.Speech software (refer to table 3).

***Statistical Analysis***

SPSS (Statistical Package of Social Sciences) program was used for descriptive and canonical discriminant analysis. Descriptive analysis was done to calculate the mean, standard deviation and range of all the parameters.

Further the data was treated with canonical discriminant analysis for classification of parameters within and across groups for all the three different softwares.

## RESULTS AND DISCUSSION

The main purpose of the present study was to compare three different software programs in terms of their efficacy of evaluation of voice disorders and to differentiate hoarse voice from normal voice using the parameters evaluated by the three software systems. The parameters evaluated using MDVP software were :

1. Average Fundamental Frequency (Fo)
2. Average Pitch Period (To)
3. Highest fundamental Frequency (Fhi)
4. Lowest Fundamental Frequency (Fhi)
5. Standard Deviation of Frequency (STD)
6. Phonatory Frequency Range (PFR)
7. Fundamental Frequency Tremor Frequency (FFTR)
8. Amplitude Tremor Frequency (FATR)
9. Absolute Jitter (Jitt)
10. Jitter Percent (Jitt %)
11. Relative Average Perturbation (RAP)
12. Pitch Period Perturbation Quotient (PPQ)
13. Smoothed Pitch Period Perturbation Quotient (SPPQ)
14. Variation in Fundamental Frequency (VFo)
15. Shimmer in dB (ShdB)
16. Shimmer Percentage (Sh%)
17. Amplitude Perturbation Quotient (APQ)
18. Smoothed Amplitude Perturbation Quotient (SAPQ)
19. Variation in Amplitude (VAM)
20. Noise to Harmonic Ratio (NHR)
21. Voice Turbulence Index (VTI)

22. Soft Phonation Index (SPI)
23. Frequency Tremor Intensity Index (FTRI)
24. Amplitude Tremor Intensity Index (ATRI)
25. Degree of Voice Breaks (DVB)
26. Number of Subharmonic Segments (NSH)
27. Degree of Subharmonic Components (DSH)
28. Degree of Voiceless (DUV)
29. Number of Voice Breaks (NVB)

The parameters evaluated using Vaghmi software are as follows:

1. Fundamental Frequency in Phonation.
2. Fundamental Frequency in Speech
3. Maximum Fundamental Frequency in Phonation.
4. Minimum Fundamental Frequency in Phonation.
5. Extent of Fuctuations in Fundamental Frequency in Phonation.
6. Speed of Fluctuations in Fundamental Frequency in Phonation.
7. Extent of Fluctuations in Intensity in Phonation.
8. Speed of Fluctuations in Intensity in Phonation.

The parameters evaluated using Dr.Speech software were :

1. Habitual Fundamental Frequency in Phonation.
2. Maximum Fundamental Frequency.
3. Minimum Fundamental Frequency.
4. Fundamental Frequency Tremor.
5. Amplitude Tremor

6. Standard Deviation of Fundamental Frequency.
7. Jitter %
8. Shimmer 96.
9. Harmonics to Noise Ratio.
10. Normalized Noise Energy.
11. Signal to Noise Ratio.

Thus the three systems used measured a total of 48parameters out of which 3 parameters were common in all three systems and 4 parameters were common among by two system (MDVP and Dr.Speech).

The three systems differed from each other in terms of technique or procedure. Since the main purpose of the study was to find out the efficacy of the systems in differentiating dysphonics from normal, their ability to differentiate was examined i.e. the minimum number of parameters required to carry out this differentiation between the dysphonics and the normal was examined.

The results and discussion of the parameters evaluated using MDVP software are presented.

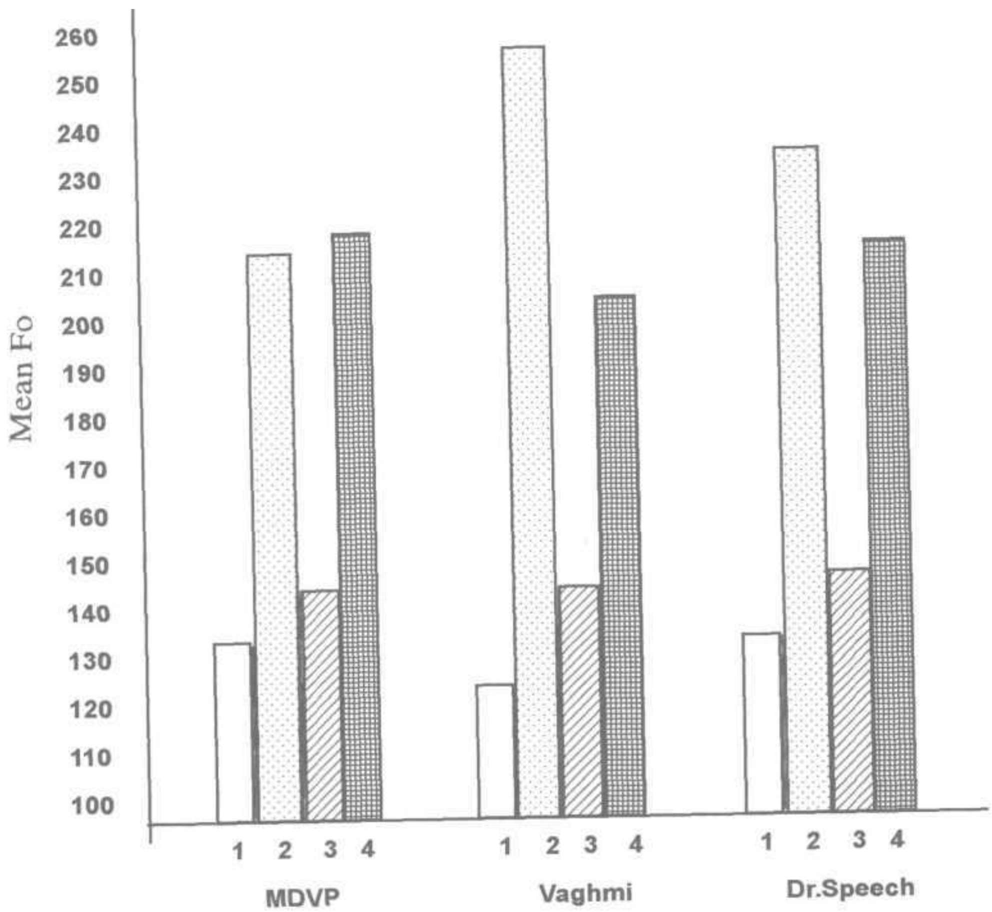
1. *Average Fundamental Frequency (Fo/Hz)* : Defined as the average value of all extracted period to period fundamental frequency values. Voice break areas are excluded. Table 4: Indicates the mean, SD and range of average Fo in phonation and speech in normal and dysphonic groups as evaluated by MDVP, Vaghmi and Dr.Speech softwares.

| Vowels | Subjects | MDVP | | | VAGHMI | | | Dr.Speech | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | Mean | Range | SD | Mean | Range | SD | Mean | Range | SD |
| /a/ | NM | 126.62 | 63.96 | 16.69 | 119.61 | 47.12 | 13.66 | 126.64 | 14.12 | 107.71-146.07 |
| | NF | 206.76 | 136.21 | 64.28 | 246.61 | 164.14 | 66.00 | 223.68 | 43.87 | 96.32-229.67 |
| | DM | 134.49 | 129.88 | 36.12 | 133.80 | 122.82 | 34.61 | 143.20 | 18.39 | 196.26-269.30 |
| | DF | 223.06 | 166.78 | 38.41 | 198.43 | 197.79 | 38.46 | 208.21 | 26.94 | 171.99-247.67 |
| /i/ | NM | 130.86 | 63.128 | 16.67 | 123.68 | 29.36 | 9.94 | 134.16 | 14.91 | 12.46-116.03 |
| | NF | 224.66 | 74.32 | 22.36 | 269.178 | 168.06 | 91.82 | 238.62 | 41.71 | 98.49-260.67 |
| | DM | 142.66 | 169.124 | 39.06 | 143.47 | 128.06 | 33.19 | 161.84 | 17.37 | 210.27-268.70 |
| | DF | 216.99 | 268.28 | 42.63 | 210.01 | 189.06 | 38.10 | 219.91 | 28.12 | 168.73-266.78 |
| /u/ | NM | 134.11 | 49.36 | 16.71 | 127.13 | 39.978 | 36.49 | 134.70 | 13.21 | 113.69-168.20 |
| | NF | 212.27 | 140.62 | 39.92 | 263.66 | 169.16 | 62.82 | 242.11 | 42.68 | 62.11-246.96 |
| | DM | 162.97 | 139.41 | 38.86 | 163.46 | 63.06 | 32.80 | 147.70 | 16.93 | 216.08-268.76 |
| | DF | 217.76 | 194.62 | 194.62 | 204.69 | 186.62 | 34.23 | 216.64 | 37.12 | 164.14-270.13 |
| /sp/ | NM | 130.87 | 67.218 | 19.66 | 136.82 | 46.76 | 40.37 | | | |
| | NF | 216.29 | 73.81 | 23.41 | 276.64 | 162.32 | 67.36 | - | - | - |
| | DM | 166.26 | 141.23 | 39.06 | 162.72 | 168.73 | 36.86 | - | - | - |
| | DF | 226.12 | 186.66 | 40.69 | 220.67 | 192.67 | 69.72 | - | - | - |

**Table 4: Mean, range, SD values of Fo in normal groups (males and females) and dysphonic groups (males and females) in phonation and speech as evaluated by MDVP, Vaghmi and Dr.Speech.**

| | MDVP | Vaghmi | Dr. Speech |
|---|---|---|---|
| NM | 130.16 | 123.44 | 131.43 |
| NF | 214.19 | 256.48 | 234.73 |
| DM | 143.37 | 143.57 | 147.58 |
| DF | 219.26 | 204.37 | 214.58 |

Table 5 : Mean values of Fo in phonation of /a/ /i/ /u/ in normal and dysphonic groups as evaluated by the three software systems.




**1 - Normal male; 2 - Normal female; 3 - Dysphonic male; 4 - Dysphonic female**

**Graph-1:** Mean values of Fo in Phonation of /a/ /i/ /u/ in normal and dysphonic groups as evaluated by the three software systems.

As seen in the table 4, for all the three softwares, the average fundamental frequency in phonation were found to be lower in normal males when compared to normal females. Similarly, the dysphonic males had a lower average fundamental frequency than the dysphonic females. The results of this study are consistent with the earlier studies done by Sheela (1 974), Vanaja (1 986), Suresh (1 991 ).

| Investigators | Male | Female |
|---|---|---|
| Sheela (1974) | 126 | 217 |
| Jayaram (197 5) | 123 | 225 |
| Nataraja and Jagadish (1 984) | 141 | 237 |
| Vanaja (1986) | 127 | 234 |
| Nataraja (1986) | 119 | 223 |
| Sreedevi (1987) | 119 | 218 |
| Tharmar (1991) | 124 | 233 |
| Suresh (1991) | 123 | 219 |
| Sanjay(1991) | 131 | 220 |
| Rajashekar (1991) | 148 | - |
| Krishnan (1992) | 122 | 231 |
| Pathak (1997) | 126 | 231 |
| Prabha (1997) | 125 | 214 |
| Pradeep(1997) | 136 | 240 |
| Rajkumar (1998) | 140 | 240 |
| David (1998) | 127 | 232 |
| Present Study | 128 | 206 |

Table 6: Values of mean fundamental frequency (in Hz) on normal Indian population as reported by various investigators.

| Investigators | Male | Female |
|---|---|---|
| Jayaram (197 5) | 174 | 202 |
| Nataraja (1986) | 152 | 200 |
| Sanjay (1991) | 157 | 233 |
| Pathak (1997) | 141 | 234 |
| Prabha (1997) | 159 | 217 |
| Rajkumar(1998) | 172 | 229 |
| David (1998) | 156 | 209 |
| Present study | 143 | 204 |

Table 7 : Values of mean fundamental frequency (in Hz) for phonation on dysphonic Indian population as reported by various investigators.

The males and females of the dysphonic group showed greater variation than the males and females of the normal groups. The average fundamental frequency of phonation of vowels / i / and / u / were greater than that for vowel /a/ in both the normal group (males and females) and the dysphonic groups (males and females). The mean values of Fo was higher in sentences compared to phonation in both the normal and dysphonic groups.

The results of the above parameters can be discussed as follows:

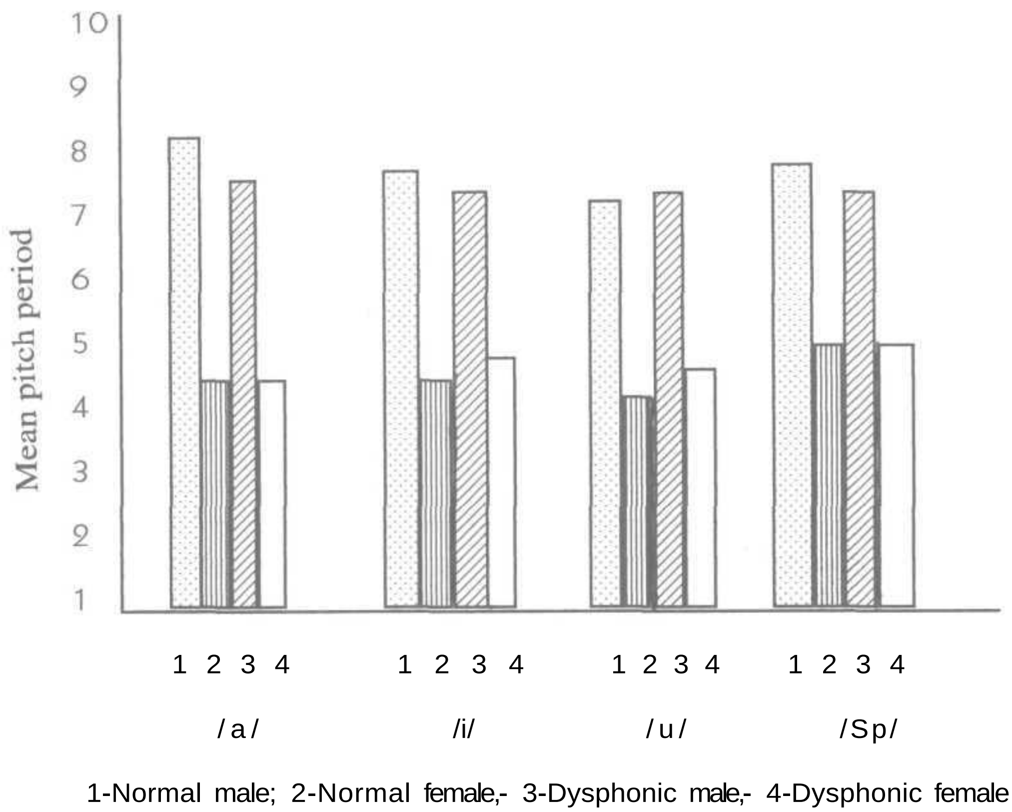
Vowels */i/* and /u/ are high vowels and the larynx is pushed upward resulting in the increase in the distance between the thyroid angle and arytenoid cartilage thereby stretching the vocal cords. This results in increase in tension (as mass is constant) leading to increased frequency of vibration of vocal folds and therefore Fo increases.

Vowel /a/ is a low mid vowel and the level of the larynx is lower when compared to the level of larynx during the phonation of /i/ and /u/ which results in decreased distance between the thyroid angle and arytenoid cartilage thereby relaxing the vocal folds, leading to a reduced frequency of vibration of vocal folds, thus decreasing the fundamental frequency. The higher value of mean Fo in sentences may be due to the variation in fundamental frequency during speech.

2. *Average Pitch Period (To)* : The mean, SD and range values of average pitch period are in phonation and speech for both normal groups (males and females) and dysphonic groups (males and females) are presented in table-8.

| Vowels | Subjects | Mean | Range | SD |
|---|---|---|---|---|
| /a/ | **NM** | 8.07 | 3.14 | 0.93 |
| | **NF** | 4.34 | 4.31 | 1.01 |
| | **DM** | 7.51 | 5.22 | 1.47 |
| | **DF** | 4.81 | 4.79 | .790 |
| ***/i/*** | **NM** | 7.80 | 2.90 | .970 |
| | **NF** | 4.30 | 4.22 | .94 |
| | **DM** | 7.22 | 7.90 | 2.02 |
| | **DF** | 4.90 | 6.11 | 1.55 |
| */u/* | **NM** | 7.02 | 8.37 | 2.05 |
| | **NF** | 4.09 | 7.25 | 1.76 |
| | **DM** | 7.09 | 3.86 | 1.34 |
| | **DF** | 4.61 | 8.30 | 1.12 |
| ***/SP/*** | **NM** | 7.92 | 3.88 | 1.11 |
| | **NF** | 4.96 | 1.85 | .549 |
| | **DM** | 7.19 | 5.88 | 1.36 |
| | **DF** | 4.98 | 3.65 | 1.25 |

Table 8: Mean, SD, range values in phonation and speech of average pitch period in both normal groups (Males and females) dysphonic groups (Males and females) as evaluated by MDVP software.




Graph-2 : Mean of average pitch period in both normal groups (males and females) and dysphonic groups (males and females) in phonation and speech as evaluated by MDVP software.

As can be seen from the above, table 8 the To for /a/ and sentence is higher than for /i/ and/u/ in case of normal males, as has already been discussed previously. In case of normal females - the mean values of To were 4.34 for /a/ 4.30 For/i/ 4.09 for/u/ and 4.96 for sentence. In case of dysphonic males, the average pitch period of /a/ was more than /i/ and for dysphonic females, the mean values obtained were/a/ 4.81 , /i/ 4.90, /u/ 4.61 /sp/ 4.98. The dysphonic group showed greater variability compared to the normal group for both phonation and sentence.

The above parameter can be discussed as follows:

The average pitch period (To) is more for low vowels /a/ and less for high vowels / i / and /u/. This is because : when the frequency of vibration of the vocal folds increases, as in the case of high vowels / i / and /u/, the To decreases and for low vowels /a/, the frequency of vibration is reduced and hence an increase in To is seen (The reason for the increase and decrease in the frequency of vibration of vocal folds for high and low vowels has been discussed earlier).

3 . *Highest Fundamental Frequency (HFO):* This is defined as the greatest of all extracted period to period fundamental frequency values.

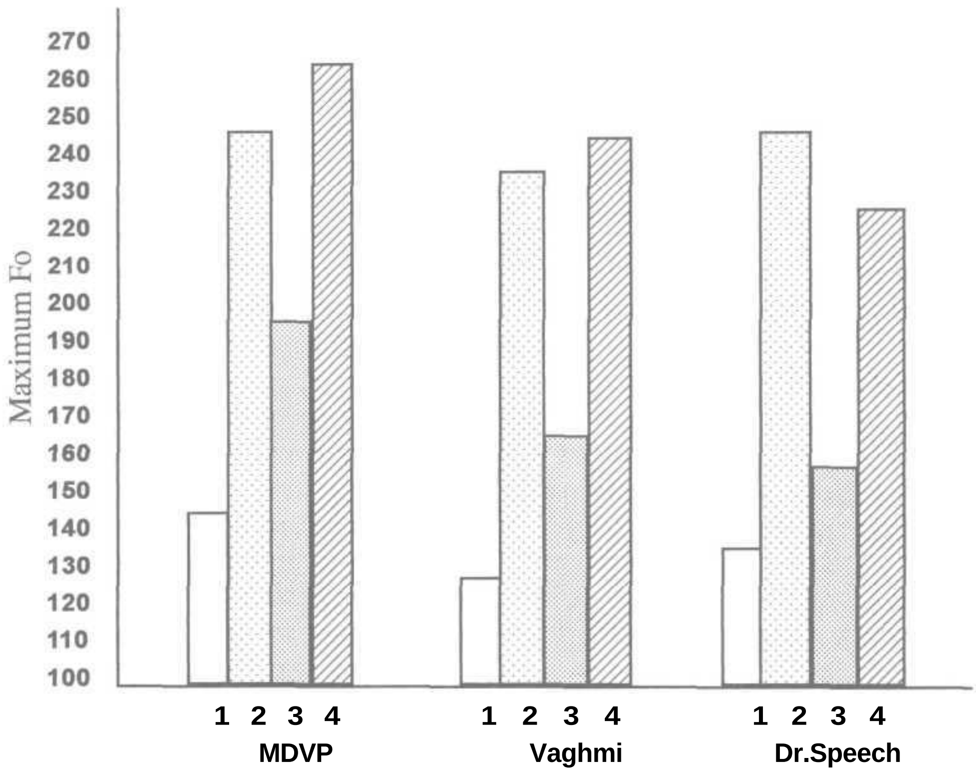
The highest fundamental frequency during phonation and sentence production for normal male and females groups and dysphonic male and female groups are presented in table 9.

| Vowels | Subjects | MDVP | | | Vaghmi | | | Dr.Speech | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | Mean | Range | SD | Mean | Range | SD | Mean | Range | SD |
| /a/ | NM | 134.88 | 82.34 | 20.31 | 124.364 | 47.29 | 13.60 | 128.72 | 13.92 | 112.21-149.49 |
| | NF | 231.76 | 64.83 | 20.22 | 224.26 | 68.03 | 20.93 | 228.78 | 43.91 | 98.44-236.83 |
| | DM | 151.09 | 160.92 | 40.09 | 149.79 | 61.47 | 21.70 | 148.46 | 18.88 | 197.76-262.60 |
| | DF | 262.80 | 160.32 | 46.67 | 232.26 | 226.13 | 80.66 | 216.36 | 27.87 | 177.11-264.91 |
| /i/ | NM | 141.46 | 89.61 | 26.61 | 130.63 | 63.40 | 16.28 | 137.97 | 16.64 | 113.66-160.96 |
| | NF | 238.63 | 87.87 | 28.10 | 231.42 | 60.72 | 20.86 | 246.12 | 61.67 | 102.32-308.39 |
| | DM | 201.80 | 308.01 | 103.68 | 169.48 | 200.76 | 61.67 | 161.07 | 19.19 | 217.24-286.38 |
| | DF | 261.31 | 236.437 | 67.66 | 246.62 | 230.13 | 79.63 | 229.69 | 33.16 | 175-304.14 |
| /u/ | NM | 149.57 | 71.076 | 20.60 | 133.72 | 33.36 | 9.31 | 138.06 | 13.94 | 116.69-164.66 |
| | NF | 271.44 | 166.62 | 46.96 | 241.76 | 61.65 | 21.72 | 261.69 | 44.04 | 103.28-280.89 |
| | DM | 226.68 | 284.51 | 100.67 | 170.66 | 210.66 | 66.66 | 162.44 | 18.22 | 220.60-282.69 |
| | DF | 270.62 | 396.38 | 86.66 | 246.66 | 231.73 | 67.92 | 228.82 | 42.89 | 168.97-304.14 |
| /sp/ | NM | 169.97 | 194.36 | 39.22 | 146.67 | 46.66 | 14.52 | - | - | - |
| | NF | 248.43 | 160.40 | 70.16 | 260.72 | 60.62 | 21.36 | - | - | - |
| | DM | 262.32 | 292.72 | 66.12 | 180.66 | 216.62 | 69.26 | - | - | - |
| | DF | 326.21 | 306.60 | 60.76 | 261.66 | 236.46 | 66.72 | - | - | - |

Table 9: Mean, Range, SD values of maximum fundamental frequency in phonation and speech in both normal group (males and females) and dysphonic groups (males and females) as evaluated by MDVP, Vaghmi andDr. Speech softwares.

| Subjects | MDVP | Vaghmi | Dr. Speech |
|---|---|---|---|
| NM | 141.97 | 129.55 | 134.91 |
| NF | 247.24 | 234.24 | 245.08 |
| DM | 193.19 | 163.33 | 157.43 |
| DF | 261.57 | 241 .50 | 224.95 |

Table-1 0 Mean values of maximum Fo in phonation in normal and dysphonic groups as evaluated by all the 3 softwares.




**1= Normal male; 2- Normal female 3- Dysphonic male; 4-Dysphonic female**

**Graph-3:** Shows Mean values of maximum fundamental frequency in phonation and speech in both normal group (males and females) and dysphonic groups (males and females) as evaluated by MDVP, Vaghmi andDr. Speech softwares.

As seen in the table 9 for all the softwares the mean value of highest Fo for /a/ is less than / i / and /u/ and the sentence was the highest when compared to /a/, /i/ and u/ in case of both normal males and females. In case of dysphonic males and females, the mean values of "Highest Fo" for the sentence was the highest when compared to the mean values of/a//i/ and /u/. As in sentence the speech sample consists of both high and low vowels the resultant HFO may be higher. The mean values of this study are in correlation with the study done by Anita (1 994) who reports of the following mean values for vowels and sentences as indicated in table 1 1 .

| Vowels and sentence | Mean | |
|---|---|---|
| | Normals | Dysphonics |
| /a/ | 133.17 | 177.12 |
| /i/ | 146.55 | 199.03 |
| /u/ | 146.74 | 188.50 |
| Sentence | 1 16.14 | 213.70 |

Table 1 1 Indicates mean values of normal and dysphonics of HFo in phonation and speech as reported by Anitha (1 994).

The variability of "Highest Fo" is greater in the dysphonic groups (both males and females) when compared to the normal groups (both males and females). This can be attributed to the inability of dysphonics to maintain a constant pitch level.

4. *Lowest Fundamental Frequency (LFO):* This is defined as the lowest of all extracted period to period fundamental frequency values.

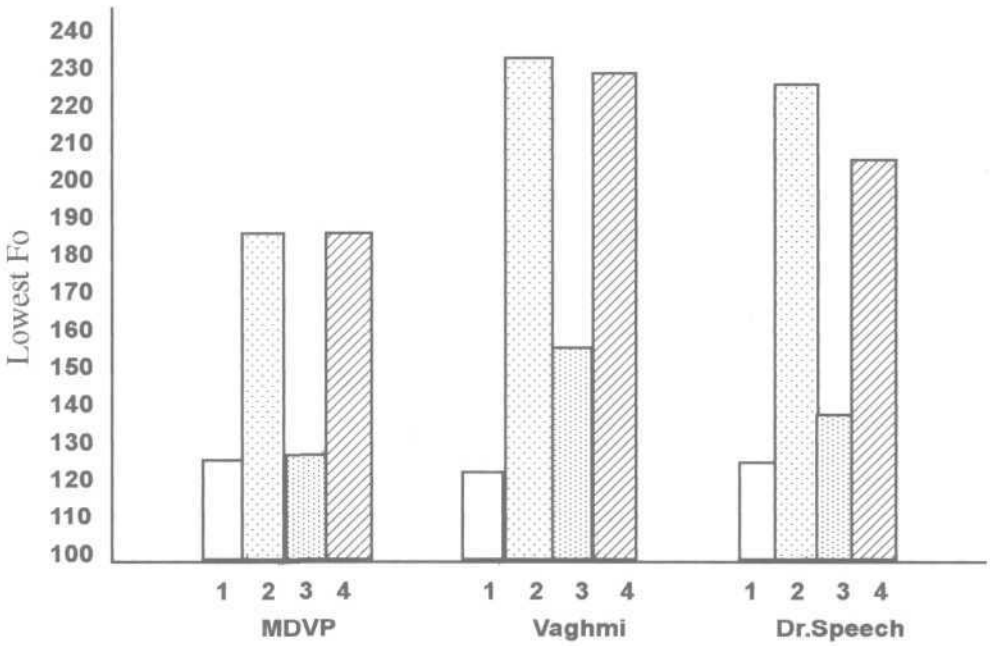
Table 1 2 indicates the mean, SD and range of the lowest fundamental frequency for phonation and speech in both normal groups (males and females) and dysphonic groups (males and females).

| | | MDVP | | | Vaghmi | | | Dr.Speech | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Vowels | Subjects | Mean | Range | SD | Mean | Range | SD | Mean | Range | SD |
| /a/ | NM | 119.82 | 62.63 | 17.09 | 119.02 | 46.32 | 13.25 | 122.16 | 13.17 | 105-140 |
| | NF | 171.29 | 70.89 | 18.71 | 219.32 | 56.05 | 19.65 | 219.67 | 42.26 | 60.66-219.40 |
| | DM | 120.71 | 166.68 | 38.16 | 143.72 | 60.49 | 20.76 | 132.61 | 18.83 | 190.09-266.40 |
| | DF | 202.27 | 262.73 | 64.17 | 224.62 | 223.03 | 79.36 | 198.77 | 27.93 | 162.73-242.31 |
| /i/ | NM | 124.29 | 70.35 | 18.28 | 120.65 | 52.60 | 14.37 | 124.30 | 25.88 | 113.65-148.99 |
| | NF | 213.53 | 60.36 | 18.40 | 225.62 | 58.65 | 19.76 | 233.07 | 30.66 | 95.66-124.77 |
| | DM | 142.65 | 159.12 | 39.05 | 154.72 | 199.35 | 60.23 | 142.16 | 17.88 | 206.12-264.07 |
| | DF | 190.25 | 152.63 | 41.64 | 223.75 | 226.13 | 78.56 | 208.68 | 30.35 | 165.79-260.59 |
| /u/ | NM | 127.96 | 83.93 | 22.20 | 130.65 | 32.72 | 8.27 | 131.30 | 12.47 | 110.80-153.66 |
| | NF | 175.22 | 166.54 | 57.04 | 247.75 | 60.65 | 20.86 | 234.07 | 38.99 | 60.33-225 |
| | DM | 111.41 | 74.58 | 28.23 | 168.76 | 208.06 | 64.65 | 140.33 | 16.93 | 207.17-264.07 |
| | DF | 152.29 | 266.02 | 74.49 | 232.96 | 228.76 | 66.86 | 202.98 | 36.13 | 149.49-247.75 |
| /sp/ | NM | 83.63 | 84.62 | 21.96 | 142.37 | 43.62 | 13.23 | | | |
| | NF | 138.68 | 126.63 | 37.88 | 256.45 | 59.56 | 20.30 | | | |
| | DM | 100.80 | 78.86 | 22.85 | 175.32 | 213.56 | 58.16 | | | |
| | DF | 122.39 | 161.47 | 49.58 | 238.76 | 232.32 | 64.86 | | | |

Table 12 Mean, range, SD values of lowest Fo in phonation and speech in both normals and dysphonics as evaluated by MDVP, Vaghmi and Dr. Speech softwares.

| Subjects | MDVP | Vaghmi | Dr. Speech |
|---|---|---|---|
| NM | 124.02 | 123.44 | 125.92 |
| NF | 186.34 | 230.89 | 228.93 |
| DM | 124.92 | 155.73 | 138.36 |
| DF | 181.60 | 229.11 | 203.47 |

Table-13 Mean values of lowest Fo in phonation (/a/ /i/ /u/)as evaluated by MDVP, Vaghmi and Dr.Speech Software.




**1= Normal male; 2- Normal female 3- Dysphonic male; 4-Dysphonic female**

Graph 4 :Mean, Range, SD values of lowest fundamental frequency in phonation in both normal group (males and females) and dysphonic groups (males and females) as evaluated by MDVP, Vaghmi andDr. Speech softwares.

The mean values of LFo in this study are in accordance with the study done by Anitha (1994). The results of her study of mean values of normals and dysphonics are presented in table 14.

| Vowel/Sentence | Mean | |
|---|---|---|
| | Normals | Dysphonics |
| /a/ | 123.42 | 29.01 |
| /i/ | 134.31 | 148.21 |
| /u/ | 134.17 | 148.91 |
| /sent/ | 101.94 | 123.38 |

Table-1 4: Mean values of normals and dysphonics of lowest Fo according to Anitha (1994).

As can be seen from the table 1 2, for MDVP software the mean of the "lowest Fo" of/a/ is lower when compared to /i/ and /u/ and the mean values of the "lowest Fo" was lower in sentences when compared to/a//i/ and/u/ in case of normal males. In case of normal females, the order of increase in the mean "lowest Fo" are/a/171 .29, /i/ 21 3 .53, /u/ 1 7 5.22 and speech 138.68. As indicated by all the thre softwares. In case of dysphonic males and females the mean value of the : lowest Fo of sentence is lowest when compared to /a//i/ and /u/ for all the 3 softwares. For Vaghmi software, the mean values of minimum fundamental frequency of /a/ was lower than /i/ and /u/ and sentence for both normal and dysphonic groups. Dr.Speech software evaluation of Flo indicated that in case of males the mean values of dysphonic group were higher than normal group. While in females, the normal females had higher mean values than normal males. The variability was greater in the dysphonic group when compared to the normal group and this can be attributed to the inability of the dysphonics to maintain constant pitch level.

The results of this parameters can be discussed as follows :

Since in sentence due to inflections used during the production of sentence, use of different speech sound having different vocal tract configuration which would indirectly affect the fundamental frequency of the voice had led to an increase in *L.FO* in normals. The dysphonic group due to various vocal pathologies, their ability to control the vocal system decrease and hence low fundamental frequency compared to normals.

5. *Standard Deviation of Fo (STD) :* Defined as the standard deviation of all extracted period to period fundamental frequency values.

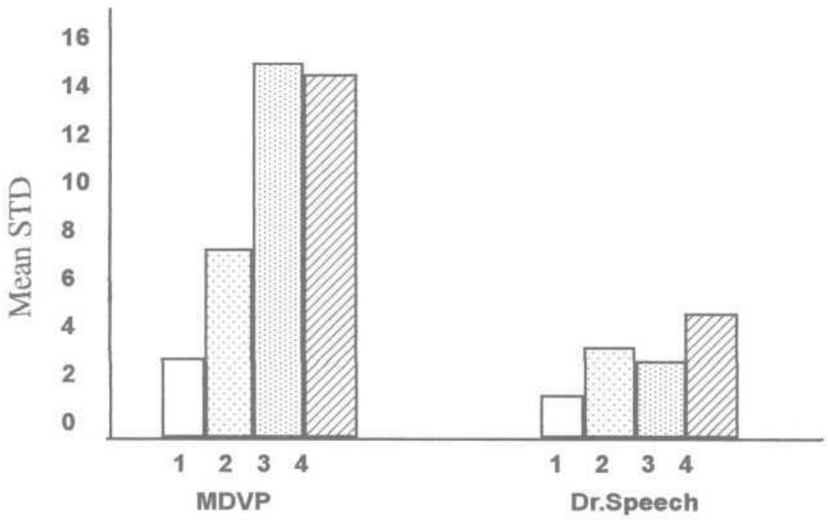
Table 1 5 indicates the mean, SD and range of "standard deviation of Fo" for phonation and sentences in both normal groups (males and females) and dysphonic groups (males and females) as evaluated by MDVP and Dr.Speech software..

| Vowels | Subjects | MDVP | | | Dr.Speech | | |
|---|---|---|---|---|---|---|---|
| | | **Mean** | **Range** | **SD** | **Mean** | **Range** | **SD** |
| ***/a/*** | **NM** | **1.68** | **2.34** | **.66** | **1.19** | **.42** | **.63-2.08** |
| | **NF** | **2.87** | **6.06** | **1.29** | **3.21** | **9.66** | **.68-46.30** |
| | **DM** | **4.77** | **23.84** | **6.06** | **1.72** | **.68** | **1.12-3.73** |
| | **DF** | **12.74** | **74.19** | **18.84** | **3.62** | **1.42** | **.78-6.93** |
| ***/i/*** | **NM** | **1.73** | **2.24** | **.78** | **1.28** | **.46** | **.46-2.03** |
| | **NF** | **3.12** | **4.74** | **1.36** | **3.66** | **6.72** | **.86-24.24** |
| | **DM** | **10.47** | **40.16** | **13.90** | **2.10** | **1.33** | **1.08-6.00** |
| | **DF** | **7.58** | **36.71** | **8.83** | **3.99** | **1.79** | **1.09-6.96** |
| ***lul*** | **NM** | **3.22** | **10.46** | **2.92** | **2.66** | **6.38** | **.67-22** |
| | **NF** | **16.23** | **83.68** | **21.11** | **3.80** | **9.66** | **.68-46.3** |
| | **DM** | **32.13** | **166.62** | **42.03** | **3.66** | **1.73** | **1.12-7.03** |
| | **DF** | **23.17** | **82.28** | **26.84** | **4.62** | **1.42** | **.78-6.93** |
| ***Ispl*** | **NM** | **14.93** | **23.13** | **6.40** | | | |
| | **NF** | **26.74** | **43.38** | **12.66** | - | | - |
| | **DM** | **36.34** | **36.16** | **7.60** | | | - |
| | **DF** | **38.47** | **74.78** | **21.74** | - | - | - |

**Table 16: Mean, range and SD values of STD in phonation and speech for both normal groups (males and females) and dysphonic groups as evaluated by MDVP, and Dr. Speech softwares.**

| Subjects | MDVP | Dr. Speech |
|---|---|---|
| NM | 2.21 | 1 .70 |
| NF | 7.42 | 3.55 |
| DM | 15.79 | 2.46 |
| DF | 14.49 | 4.07 |

Table-1 6 :Mean values of STD in phonation (of /a/ /i/ /u/) for both normal and dysphonic groups as evaluated by MDVP and Dr.Speech.




**1= Normal male; 2- Normal female 3- Dysphonic male; 4-Dysphonic female**

Graph-5: Indicates Mean values ofstandard deviation of frequency in phonation in both normal group (males and females) and dysphonic groups (males and females) as evaluated by MDVP and Dr. Speech softwares.

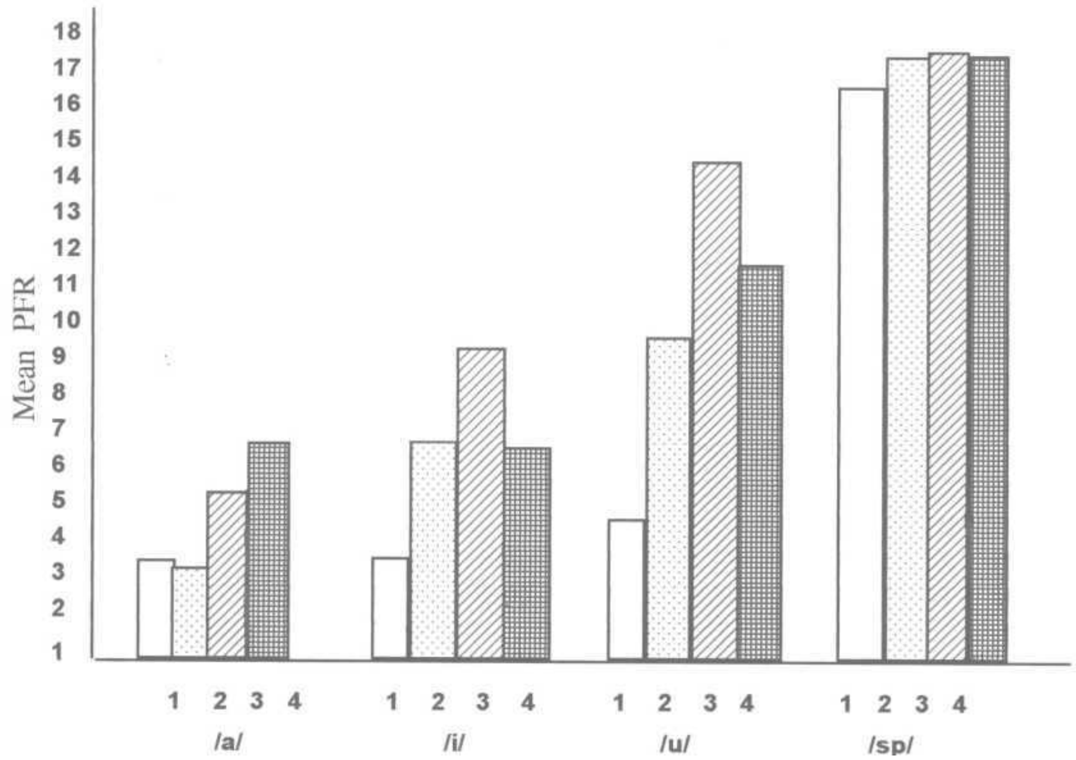
As can be seen from table 1 5 when measured with MDVP and Dr.Speech the mean values of "standard deviation of Fo" in case of normal males and normal females was less for/a/ when compared to / i / and /u/. In case of dysphonic males and females, the mean value of STD when measured with MDVP software was higher for sentence when compared to /a/, / i / and /u/. This was seen in case of normal males and females also. The variability of SD Fo was higher in the dysphonic group compared to the normal group. The increase in STD in dysphonics may be attributd to the inability to maintain a constant pitch and intensity during phonation and sentence due to various vocal pathology.

6. *Phonatory Fo Range (PFR)* : It is defined as the range between Fhi and Flo expressed in number of semitones.

Table 1 7 indicates the mean, SD and range of "Phonatory Fo Range" for phonation and sentences in both normal groups (males and females) and dysphonic groups (males and females) as evaluated by MDVP software.

| Vowels | Subjects | Mean | Range | SD |
|---|---|---|---|---|
| /a/ | **NM** | 3.13 | 5.00 | 1.50 |
| | **NF** | 3.06 | 3.00 | 0.96 |
| | **DM** | 5.00 | 15.00 | 4.35 |
| | **DF** | 6.93 | 19.00 | 5.86 |
| /i/ | **NM** | 3.26 | 8.00 | 2.43 |
| | **NF** | 6.93 | 4.00 | 1.16 |
| | **DM** | 9.00 | 29.0 | 9.91 |
| | **DF** | 6.40 | 14.00 | 4.35 |
| | **NM** | 4.60 | 15.00 | 4.06 |
| | **NF** | 9.60 | 22.00 | 7.61 |
| | **DM** | 14.13 | 31.00 | 9.17 |
| | **DF** | 11.76 | 28.00 | 9.37 |
| /SP/ | **NM** | 16.33 | 10.00 | 3.06 |
| | **NF** | 17.06 | 14.00 | 4.14 |
| | **DM** | 17.56 | 11.00 | 4.12 |
| | **DF** | 17.26 | 27.00 | 7.11 |

Table 1 7 :Mean, range, SD values of phonatory frequency range in phonation and speech for both normal groups (males and females) and dysphonic groups (males and females) as measured by MDVP software.




**= Normal male; 2- Normal female 3- Dysphonic male; 4-Dysphonic female**

Graph-6: Mean values of phonatory frequency range in phonation and speech in both normal group (males and females) and dysphonic groups (males and females) as evaluated by MDVP software.

Taking into consideration the mean values of phonatory frequency range for the phonation of different vowels and sentence, it was found that the mean PFR value for sentence was highest when compared to /a/, /u/ and /i/ in all the four groups (normal males, normal females, dysphonic males and dysphonic females).

In case of normal males, normal females and dysphonic males, the values of mean PFR for /a/ and /i/ was lower than that of /u/. However in case of dysphonic females the value of mean PFR of /i/ was less than /a/ and /u/.

The dysphonic group (males and females) had the mean PFR values of phonation of /a/ /i/ and /u/ higher than that of normal males and females respectively. The results of this parameters are in accordance with the study done by Anitha (1994).

The above results can be discussed as follows:

The mean PFR value for sentence was higher for phonation of vowels /a/ /i/ and /u/ in all the four groups (normal males, normal females, dysphonic males, dysphonic females). This could be due to the inflections used during the production of the sentence, use of different speech sounds having different vocal tract configuration which would in directly affect the fundamental frequency of the voice and hence the range of Fo higher for sentence than for phonation of vowels /a/, /'\j and /u/.

The mean values of PFR for vowels in dysphonic males and females were higher than in normal males and females which could be attributed to the inability of the dysphonics to maintain a constant pitch during sustained phonation.

7. *Fo Tremor Frequency (FFTR):* This is the frequency of the most intensive low frequency Fo modulating component in the specified Fo tremor analysis range.

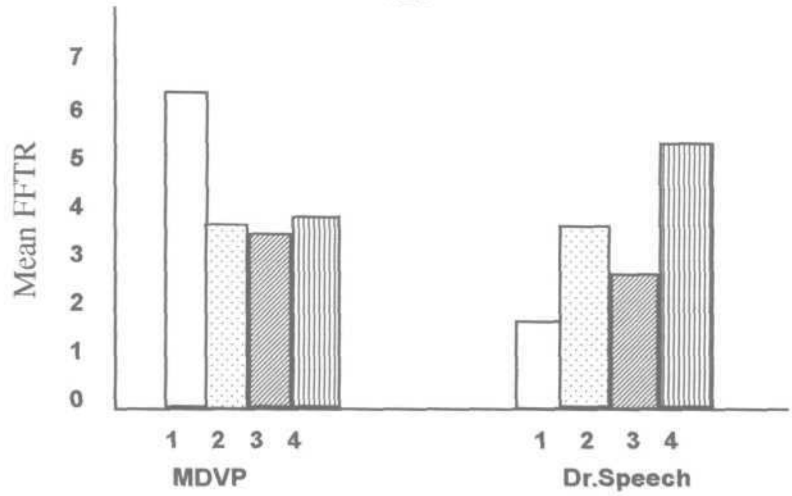
Table 1 8 indicates the mean, SD, and range of "fundamental frequency tremor frequency for phonation in both normal groups (normal males and females) and dysphonic groups (dysphonic males and females) as evaluated by MDVP and Dr.Speech software.

| Vowels | Subjects | MDVP | | | Dr.Speech | | |
|---|---|---|---|---|---|---|---|
| | | Mean | Range | SD | Mean | Range | SD |
| /a/ | NM | 6.32 | 5.29 | 19.96 | 1.99 | 1.33 | 1-4.72 |
| | NF | 3.92 | 3.31 | 12.67 | 2.28 | 1.33 | 1-4.87 |
| | DM | 2.18 | 1.26 | 4.23 | 2.12 | 1.93 | 1-7.2 |
| | DF | 2.78 | 3.83 | 16.00 | 3.56 | 3.36 | 1.34-14.64 |
| /i/ | NM | 7.53 | 7.26 | 22.22 | 1.44 | 0.56 | 1-3.10 |
| | NF | 3.4 | 2.86 | 9.75 | 3.12 | 3.43 | 1.02-12.86 |
| | DM | 4.97 | 4.79 | 14.81 | 1.24 | 0.26 | 1-2.01 |
| | DF | 5.23 | 6.48 | 20.77 | 4.32 | 4.36 | 1-13.72 |
| /u/ | NM | 5.79 | 6.89 | 21.17 | 2.06 | 1.10 | 1.01-4.07 |
| | NF | 3.93 | 3.86 | 11.76 | 5.23 | 4.88 | 1-14.35 |
| | DM | 5.62 | 6.39 | 17.39 | 4.92 | 4.90 | 1-14.88 |
| | DF | 3.81 | 3.12 | 10.20 | 7.68 | 5.12 | 1.06-14.92 |

Table 1 8: Mean, range and SD values of normal males and females and dysphonic males and females of FFTR during phonation as evaluated by MDVP, and Dr. Speech softwares.

| Subjects | MDVP | Dr.Speech |
|---|---|---|
| NM | 6.21 | 1 .83 |
| NF | 3.75 | 3.54 |
| DM | 3.69 | 2.76 |
| DF | 3.94 | 5.18 |

Table-1 9: Mean values of FFTR in phonation in normal and dysphonic groups as evaluated by MDVP and Dr.Speech




**1= Normal male; 2- Normal female 3- Dysphonic male; 4-Dysphonic female**

Graph - 7: Mean values of Fo Tremor fundamental frequency in phonation in both normal group (males and females) and dysphonic groups (males and females) as evaluated by MDVP and Dr. Speech softwares.

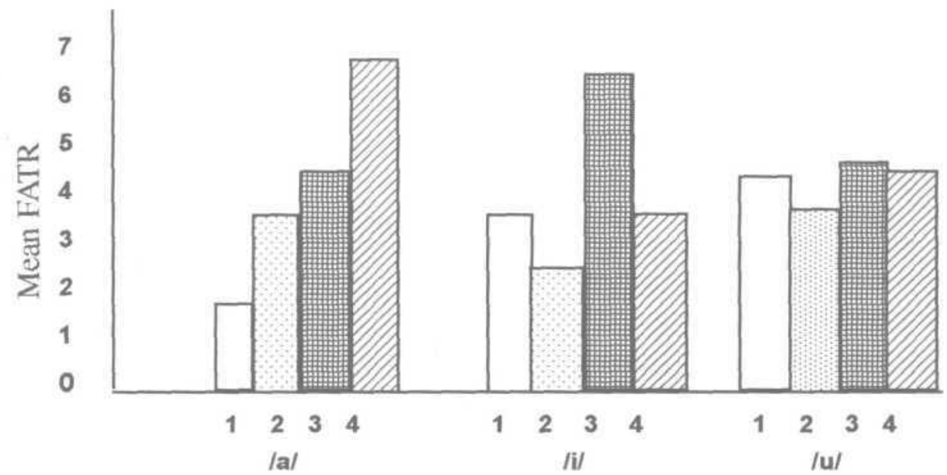
As seen in table 1 8, as measured by MDVP software the mean value of FFTR for normal males was more for/i/ compared to /a/ and /u/, while in normal females this value was less than /a/ and /u/. In case of dysphonic males, the mean value of FFTR was highest for/u/, while in dysphonic females it was highest for/i/ compared to /a/ and /u/. As measured by Dr.Speech, the mean values of FFTR for normal males and females was highest for /u/ compared to and /i/. Similar results were noted for the dysphonic group also.

8. *Amplitude Tremor Frequency (FATR)* : It is defined as the frequency of the most intensive low frequency amplitude modulating component in the specified amplitude tremor analysis range.

| Vowels | Subjects | Mean | Range | SD |
|---|---|---|---|---|
| /a/ | NM | 1 .90 | 4.34 | 16.66 |
| | NF | 3.46 | 4.70 | 16.00 |
| | DM | 4.40 | 5.90 | 21.06 |
| | DF | 6.93 | 2.99 | 10.81 |
| /i/ | NM | 3.49 | 6.15 | 21 .05 |
| | NF | 2.26 | 2.69 | 16.05 |
| | DM | 6.31 | 6.15 | 22.22 |
| | DF | 3.66 | 5.63 | 19.04 |
| /u/ | NM | 4.09 | 6.81 | 21.05 |
| | NF | 3.71 | 5.61 | 16.67 |
| | DM | 4.76 | 4.48 | 17.39 |
| | DF | 4.34 | 4.77 | 14.28 |

Table 20 - Indicates the mean, range and SD values for normal males, normal females, dysphonic males, dysphonic females of the 'Amplitude Tremor Frequency" in phonation of/a//i/ and/u/ vowels as evaluated by MDVP software.

As can be noted from table 20, the dysphonic group (males and females) showed greater mean values of FATR than the normal group (males and females). This can be attributed to the inability of the dysphonics to maintain a constant pitch and intensity.




**1= Normal male; 2- Normal female 3- Dysphonic male; 4-Dysphonic female**

Graph-8: Mean of amplitude tremor frequency in phonation in both normal group (males and females) and dysphonic groups (males and females) as evaluated by MDV software.

Observation of table-20 indicates that in case of normal males and females, the mean value of FATR is largest for /u/ compared to /a/ and /i/. The dysphonic male group indicated larger mean value of FATR for/i/ compared to /a/ and /u/, while the dysphonic female group indicated larger mean value of FATR for /a/ compared to */i/* and /u/. The results of this parameter are in accordance to the mean values of this parameter reported by Anitha (1994).

| Vowels | Mean | |
|---|---|---|
| | Normal | Dysphonics |
| /a/ | 2.30 | 3.46 |
| /i/ | 2.33 | 4.91 |
| /u/ | 2.83 | 3.08 |

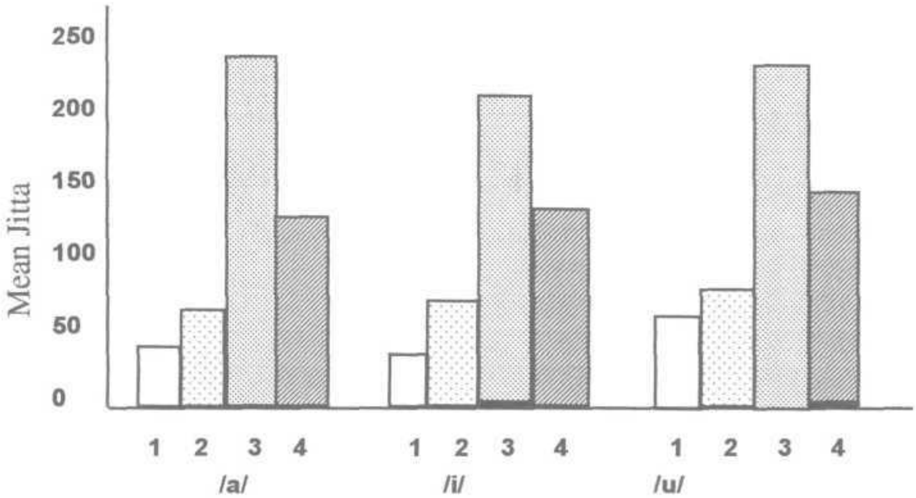
Table 21 : Mean values of FATR for phonation of normal (males and females) dysphonic (males and females) groups as reported by Anitha (1994).

9. *Absolute Jitter (Jitta)* : It is an evaluation of the period to period variability of the pitch period within the analyzed voice sample.

Table 22 indicates the mean, SD and range of absolute jitter for normal males, normal females, dysphonic males and females for phonation of vowels /a/, /i/ and /u/ as measured by MDVP software.

| Vowels | Subjects | Mean | Range | SD |
|---|---|---|---|---|
| /a/ | NM | 46.16 | 239.84 | 72.53 |
| | NF | 56.58 | 149.942 | 34.05 |
| | DM | 240.84 | 562.04 | 225.60 |
| | DF | 129.62 | 425.65 | 140.72 |
| /i/ | NM | 42.08 | 167.261 | 41.55 |
| | NF | 69.64 | 111.66 | 35.84 |
| | DM | 201.32 | 425.25 | 150.66 |
| | DF | 135.76 | 356.55 | 91.42 |
| /u/ | NM | 50.76 | 484.76 | 136.72 |
| | NF | 70.12 | 496.76 | 145.72 |
| | DM | 236.75 | 962.56 | 192.36 |
| | DF | 146.45 | 346.75 | 245.44 |

Table 22: Mean, SD and range values of Absolute jitter for normal and dysphonic groups for phonation of /a/, / i / and / u / as measured by MDVP software.




**1= Normal male; 2- Normal female 3- Dysphonic male; 4-Dysphonic female**

Graph-9: Indicates Absolute Jitter in phonation in both normal group (males and females) & dysphonic groups (males and females) as evaluated by MDVP software.

Comparision of mean values of absolute jitter for the phonation of different vowels indicated that the mean absolute jitter value of /a/ and /i/ were lower than /u/ in case of normal males. In case of normal females, the mean value of /a/ was lesser than /i/ and /u/. In case of dysphonic males, /i/ had lower mean value when compared to /a/ and /u/.

As seen from the definition the following parameters: absolute jitter, jitter percent, relative average perturbation, pitch perturbation quotient and smoothed pitch perturbation quotient are interrelated. Hence, the results of all these parameters are discussed together. They all measure the short or long term variation of the pitch period within the analysed voice sample but they are different in terms of the smoothing factors used. In RAP, a smoothing factor of 3 is used, PPQ uses 5 whereas SPPQ uses 55 as a smoothing factor.

The mean absolute jitter values were higher in normal females than in normal males. This results is in contradiction with the results of Higgins and Saxman (1 989), but in agreement with that of Sorenson and Horii (1 983).

In case of dysphonic males and females and males had higher values for vowels than dysphonic females. The dysphonic group exhibited greater variability than the normal group. This may be due to inability of the dysphonics to maintain a constant pitch in phonation. The results of this study are in agreement with the results of the study done by Chandrashelcar (1 987), \XanLeden et al. (1 966) and Anitha (1994).

*10. Jitter Percent (Jitt 96) :* Defined as the relative evaluation of the period to period (very short term) variability of the pitch within the analyzed voice sample.

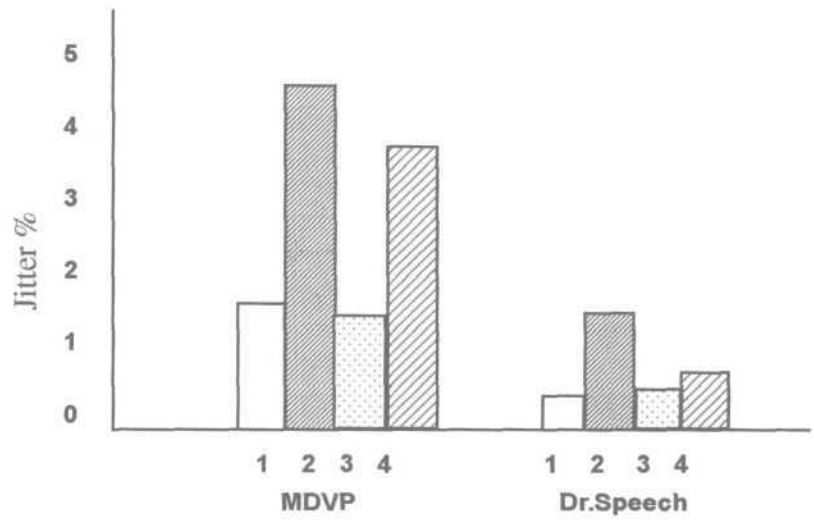
The mean, SD and range values for both normals and dysphonics for the phonation of /a/, /i/ and /u/ are given as evaluated by MDVP and Dr.Speech softwares.:

| Vowels | Subjects | MDVP | | | Dr.Speech | | |
|---|---|---|---|---|---|---|---|
| | | Mean | Range | SD | Mean | Range/ | SD |
| /a/ | NM | 1.25 | 2.63 | 0.81 | 0.28 | 0.12 | 0.11-.55 |
| | NF | 2.38 | 13.15 | 3.27 | 0.44 | 0.15 | 0.11-2.25 |
| | DM | 1.42 | 2.93 | 0.60 | 0.30 | 0.10 | 0.13-.33 |
| | DF | 3.01 | 5.10 | 1.82 | 0.48 | 0.18 | 0.03-.76 |
| /i/ | NM | 0.89 | 1.26 | 0.35 | 0.21 | 0.10 | 0 .12-.48 |
| | NF | 4.60 | 21.41 | 5.86 | 2.71 | 0.15 | 0.10-.53 |
| | DM | 1.46 | 2.78 | 0.79 | 0.23 | 1.75 | 0.10-0.45 |
| | DF | 2.69 | 3.81 | 1.23 | 0.6 | 0.58 | 0.13-2.21 |
| /u/ | NM | 2.54 | 6.99 | 2.22 | 0.25 | 1.6 | 0.12-.75 |
| | NF | 6.78 | 12.99 | 3.64 | 0.83 | 0.62 | 0.14-2.32 |
| | DM | 1.52 | 2.56 | 0.82 | 0.62 | 0.45 | 0.14-1.74 |
| | DF | 6.27 | 11.64 | 3.42 | 1.14 | 0.72 | 0.29-2.67 |

**Table 23: Mean, range and SD values for normal groups (males and females) and dysphonic groups (males and females) of jitter % for vowels.**

| Subjects | MDVP | Dr.Speech |
|---|---|---|
| NM | 1.56 | 0.24 |
| NF | 4.58 | 1.32 |
| DM | 1.46 | 0.38 |
| DF | 3.99 | 0.74 |

Table-24 Average mean values of Jitt % in phonation (/a//i//u/) in normal and dysphonics as evaluated by MDVP and Dr.Speech.




**1= Normal male; 2- Normal female 3- Dysphonic male; 4-Dysphonic female**

Graph - 10: Mean of Jitter percent in phonation in both normal group (males and females) and dysphonic groups (males and females) as evaluated by MDVP and Dr. Speech softwares.

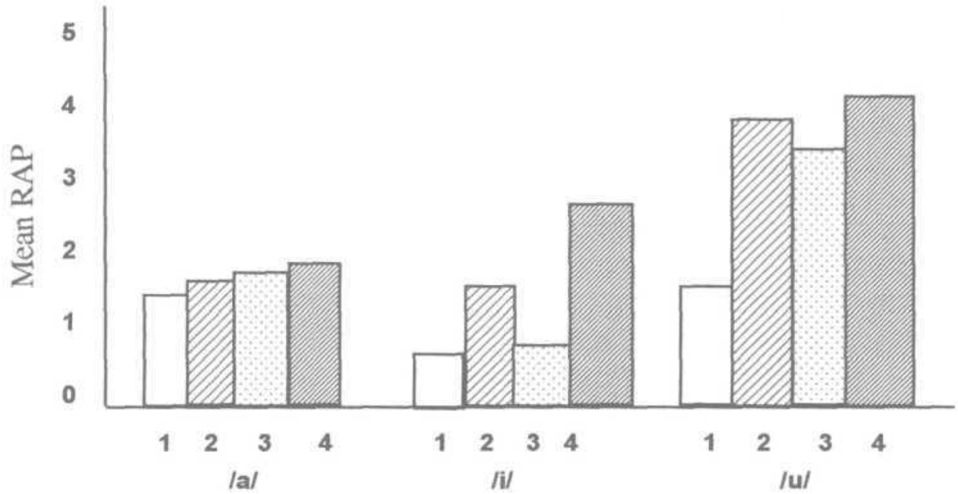
As measured by Dr.Speech in case of dysphonic males the mean value of jitter percent was highest for vowel /u/ followed by /a/ and /i/ while for normal females the mean value of/i/was highest followed by/u/ and/a/. As measured by Dr.Speech and MDVP software. In case of normal males, the mean values of "jitter percent" was highest for vowel /u/ followed by vowel /a/ and /i/. As measured by MDVP software in case of dysphonic males and normal female the mean value of jitter percent was highest for /u/ followed by /i/ and /a/. The dysphonic females indicated highest mean value of jitter percent for vowel /u/ followed by vowel /a/ and /i/.

In the present study while comparing the mean values of jitter percent for phonation of different vowels, the dysphonic group had higher values than the normal group. This is in agreement with the results of study done by Chandrashekar (1987) and VonLeden et al. (1966). On the whole, the dysphonic groups exhibited greater variability than the normal group. This could be attributed to the inability of the dysphonics to maintain a constant pitch in phonation.

*1 1. Relative Average Perturbation (RAP) :* It is defined as relative evaluation of the period to period variability of the pitch of the analysed voice sample with smoothing factor of three periods. Table 25 indicates the mean, range and SD values in phonation in RAP in both normal groups (males and females) and dysphonic groups (males and females) as evaluated by MDVP software.

| Vowels | Subjects | Mean | Range | SD |
|---|---|---|---|---|
| /a/ | NM | 1.93 | 1.50 | 0.49 |
| | NF | 1.74 | 2.76 | 1.74 |
| | DM | 1.79 | 3.26 | 1.89 |
| | DF | 1.94 | 6.94 | 1.01 |
| /i/ | NM | 0.59 | 1.03 | 0.29 |
| | NF | 1.60 | 2.10 | 3.39 |
| | DM | 0.88 | 1.72 | 0.49 |
| | DF | 2.46 | 2.15 | 0.73 |
| **/u/** | NM | 1.54 | 4.39 | 1.38 |
| | NF | 3.94 | 7.08 | 2.00 |
| | DM | 3.31 | 6.29 | 2.31 |
| | DF | 3.98 | 7.50 | 1.99 |

Table 25: Mean, SD and Range values of relative average perturbation of normal and dysphonic groups for/a///and /u/ as evaluated by MDVP software.




**1= Normal male; 2- Normal female 3- Dysphonic male; 4-Dysphonic female**

Graph - 11 : Mean values of relative average perturbation in phonation in both normal group (males and females) and dysphonic groups (males and females) as evaluated by MDVP software.

In case of normal males, the mean value of /a/ was highest when compared to /i/ and /u/. In case of dysphonic males, the mean value of RAP was highest for /u/ compared to /a/ and /i/. In case of normal females, the mean value of RAP was highest for /a/ when compared to /i/ and /u/. In case of dysphonic females, the mean value of RAP was highest for /u/ compared to /a/ and /i/. The dysphonic group exhibited greater variability than the normal group for all the vowels.

It was also seen that the mean values of this parameter for dysphonic males and females for all the vowels were higher than for normal males and females. The increase in the means of dysphonics can be attributed to their inability in maintaining a constant pitch while phonating. The results of the present study are in accordance with those of Anitha (1994).

| Vowels | Mean | | Present study | |
|---|---|---|---|---|
| | Normal | Dysphonics | Normal | Dysphonics |
| /a/ | 0.38 | 1 .36 | 1.53 | 1.86 |
| /i/ | 0.58 | 1.23 | 1.07 | 1.82 |
| /u/ | 0.49 | 1 .46 | 2.74 | 3.64 |

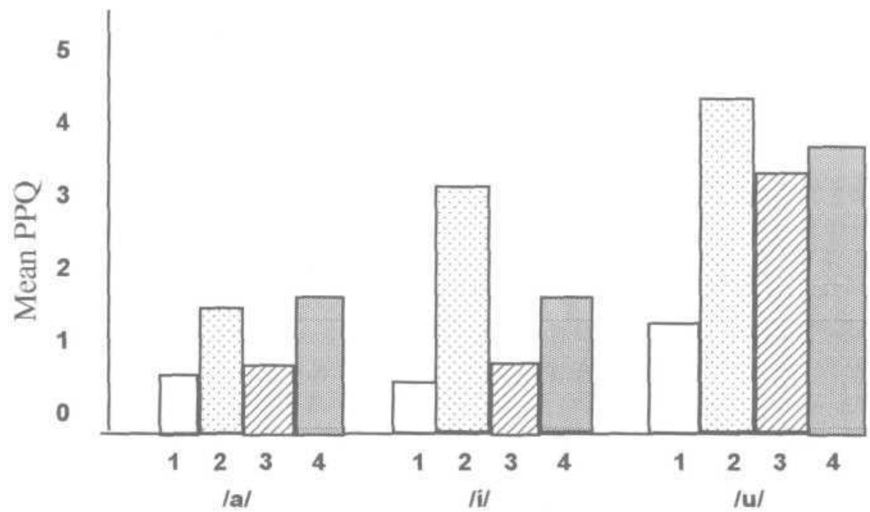
Table 26: The mean values of RAP in phonation of/a//i//u/ as reported by Anitha (1994).

*12. Pitch Period Perturbation Quotient (PPQ) :* Defined as the relative evaluation of the period to period variability of pitch within the analyzed voice sample with a smoothing factor of 5 periods.

The mean, SD and range of PPQ for the four groups are presented in table 27.

| Vowels | Subjects | Mean | Range | SD |
|---|---|---|---|---|
| /a/ | NM | 0 .73 | 1.49 | 0.46 |
| | NF | 1 .51 | 3.03 | 1.99 |
| | DM | 0.81 | 1.74 | 0.43 |
| | DF | 1 .85 | 3.36 | 1 .12 |
| /i/ | NM | 0.57 | 0.97 | 0.27 |
| | NF | 3.00 | 14.21 | 3.93 |
| | DM | 0.86 | 1.53 | 0.44 |
| | DF | 1 .65 | 2.65 | 0.81 |
| /u/ | NM | 1.48 | 3.78 | 1.23 |
| | NF | 4.32 | 9.52 | 2.50 |
| | DM | 3.31 | 7.74 | 2.8 |
| | DF | 3.75 | 8.72 | 2.12 |

Table 27: Mean, range and SD values for normal males, females, dysphonic males and females for PPQ during phonation of/a/, /i and/u/ as evaluated by MDVP software.




**1= Normal male; 2- Normal female 3- Dysphonic male; 4-Dysphonic female**

Graph -12: Mean of PPQ in phonation in both normal group (males and females) and dysphonic groups (males and females) as evaluated by MDVP software.

Comparision of the mean values of pitch perturbation quotient for the phonation of different vowels indicated that the mean pitch period perturbation quotient of /u/ was the highest followed by /a/ and /i/ in case of normal females. The dysphonic males and normal females indicated greater mean values of PPQ for /u/ when compared to /a/ and /i/. The mean PPQ value of /u/ was highest for dysphonic females followed by /i/ and /a/. The mean value of this parameter are in accordance with those reported by Anitha (1994). The values reported by Anitha (1994) are as follows:

| Vowels | Mean | |
|---|---|---|
| | Normal | Dysphonic |
| */a/* | 0.38 | 1 .24 |
| */i/* | 0.57 | 1 .71 |
| */u/* | 0.48 | 2.43 |

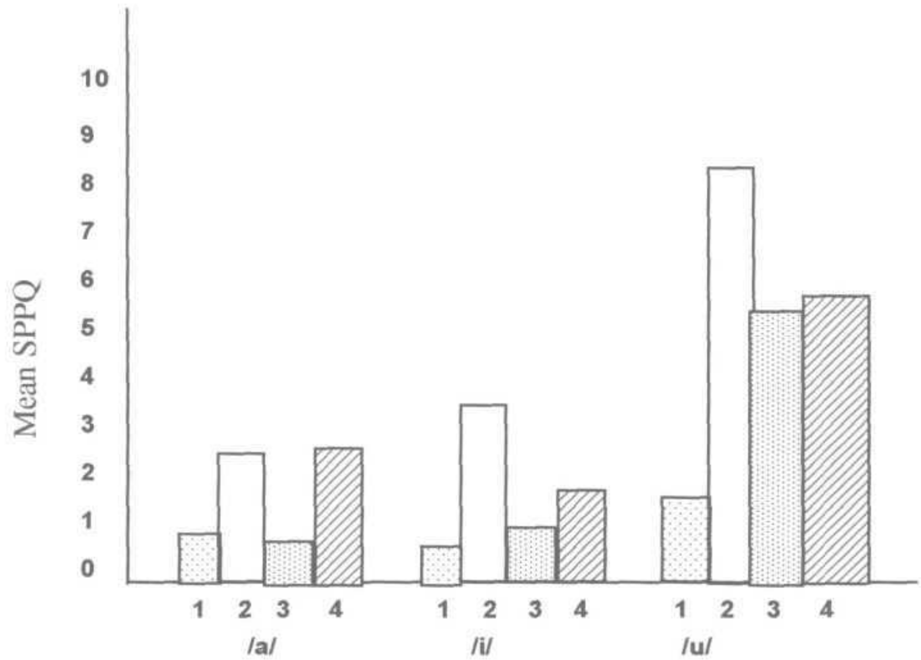
Table-28: Mean values of PPQ information in normal (males and females)and dysphonic group(males and females) as reported by Anitha (1994).

The variability was higher in the dysphonic group compared to the normal group which can be attributed to the inability of the dysphonics to maintain a constant pitch during phonation. The mean value of this parameter for dysphonic males and females for all the vowels were higher than for normal males and females.

*13. Smoothed Pitch Perturbation Quotient (SPPQ)* : This is the relative evaluation of the short/long term variability of the pitch period within the analysed voice sample at smoothing factor defined by the user. The mean, SD and range of SPPQ are presented in table 29 for normal males, normal females, dysphonic males and dysphonic females respectively.

| Vowels | Subjects | Mean | Range | SD |
|---|---|---|---|---|
| ***/a/*** | NM | 0.88 | 1.82 | 0.53 |
| | NF | 2.23 | 10.62 | 02.74 |
| | DM | 0.85 | 1.73 | 0.42 |
| | DF | 2.30 | 8.23 | 2.07 |
| ***/i/*** | NM | 0.65 | 0.94 | 0.26 |
| | NF | 3.28 | 13.59 | 4.60 |
| | DM | 0.92 | 1 .42 | 0.42 |
| | DF | 1.69 | 2.99 | 0.87 |
| ***/u/*** | NM | 1.61 | 3.78 | 1.30 |
| | NF | 8.17 | 28.36 | 7.65 |
| | DM | 5.28 | 20.87 | 6.76 |
| | DF | 5.82 | 20.95 | 6.01 |

Table 29: Mean, range and SD values of SPPQ for normal and dysphonic groups for phonation of/a/ /i/ and/u/ as evaluated by MDVP software




**1= Normal male; 2- Normal female 3- Dysphonic male; 4-Dysphonic female**

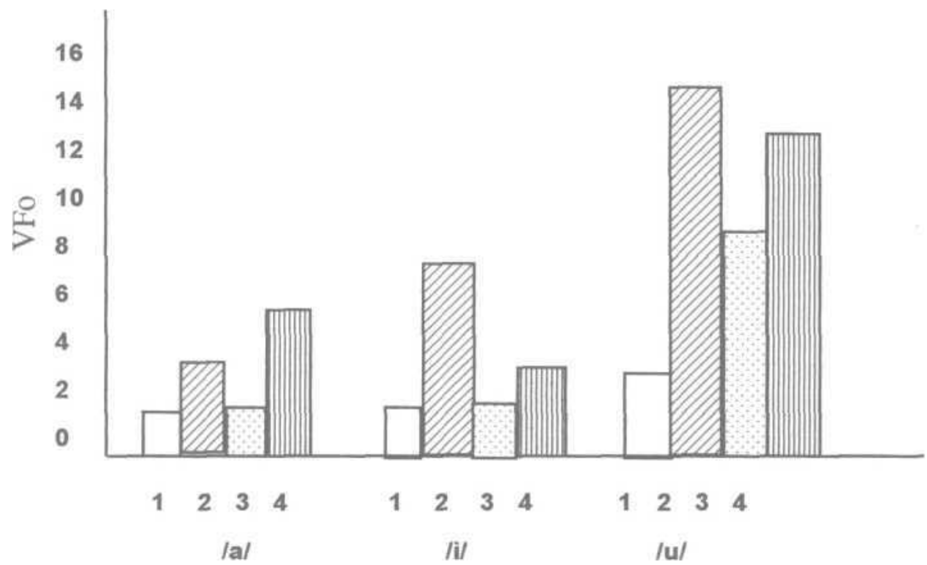
Graph - 1 3: Mean of SPPQ in phonation in both normal group (males and females) and dysphonic groups (males and females) as evaluated by MDVP software..

Comparision of the mean values of SPPQ for phonation of different vowels indicated that in case of normal males and dysphonic females, the mean values of SPPQ of /u/ was highest followed by /a/ and /i/. In case of dysphonic males and normal females the mean values of SPPQ and /u/ was highest followed by /(/ and /a/. The variability was more in the dysphonic group compared to the normal group which can be attributed to the inability of the dysphonics to maintain a constant pitch during phonation. The mean value of this parameter for dysphonic males and females for all the vowels were higher than normal males and females.

1 4. *Variation in Fundamental Frequency (VFo) :* This is defined as the relative standard deviation of the Fo and it reflects in general the variation of Fo. The mean, SD and range of this coefficient of Fo variation of normal groups (males and females) and dysphonic groups (males and females) are presented in table 30.

| Vowels | Subjects | Mean | Range | SD |
|---|---|---|---|---|
| **/a/** | NM | 1.22 | 2.92 | 0.68 |
| | NF | 3.77 | 21.57 | 5.41 |
| | DM | 1 .24 | 2.24 | 0.54 |
| | DF | 5.91 | 31.65 | 8.28 |
| **/i/** | NM | 1.32 | 2.00 | 0. 63 |
| | NF | 7.63 | 29.35 | 10.56 |
| | DM | 1.36 | 1.88 | 0.53 |
| | DF | 3.43 | 13.30 | 3.29 |
| **/u/** | NM | 2.80 | 8.48 | 2.36 |
| | NF | 14.25 | 65.26 | 18.47 |
| | DM | 8.58 | 36.06 | 10.82 |
| | DF | 12.52 | 42.12 | 14.46 |

Table 30: Indicating the mean, range and SD values of normal males, normal females, dysphonic males and dysphonic females on "coefficient of Fo variation for phonation"measured by MDVP software..




**1= Normal male; 2- Normal female 3- Dysphonic male; 4-Dysphonic female**

Graph - 14: Mean of variation of fundamental frequency in phonation in both normal group (males and females) and dysphonic groups (males and females) as evaluated by MDVP software.

While comparing the mean values of the coefficient of Fo variation for the phonation of different vowels it was found that in the case of all the groups (normal males, normal females, dysphonic males and dysphonic females), the order of increase in the mean co-efficient of Fo was /a/ followed by / i / and / u /, with / u / showing the highest values of mean Fo.

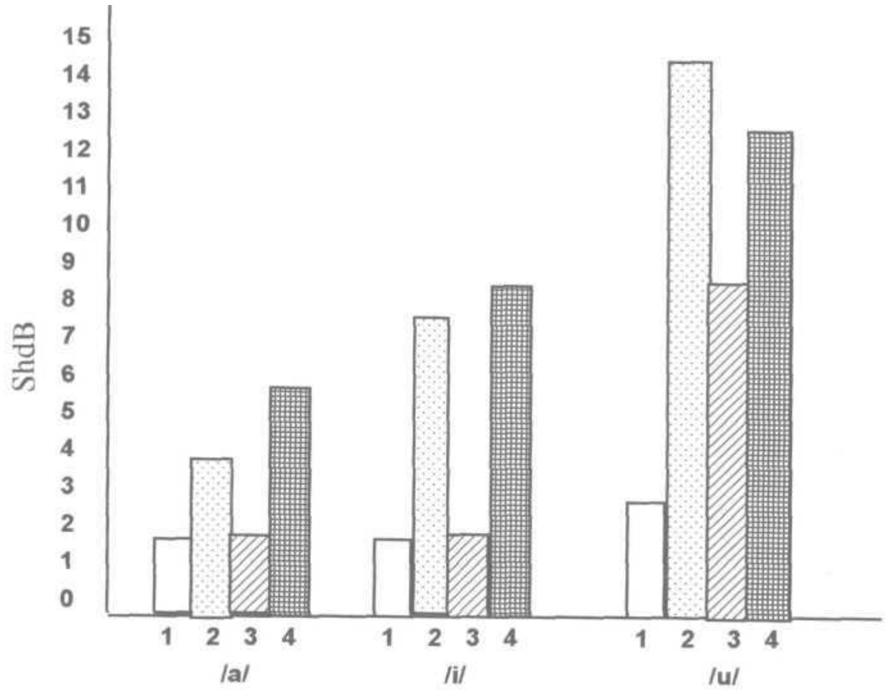
The variability was more in the dysphonic group compared to the normal group for all the vowels and also the mean value of the dysphonics was higher than that of normals for all the groups and for phonation of all the vowels.

The results of this parameter has been discussed as follows: VFo incrases in dysphonics because of the inability of the dysphonics to maintain a constant pitch while phonation

*15. Shimmer in dB (ShdB):* This measures the very short term (cycle to cycle) irregularity of the peak to peak amplitude of voice. The mean, SD and range for this measure are presented in table 3 1 for normal males, normal females, dysphonic males and dysphonic females as measured by MDVP software.

| Vowels | Subjects | Mean | Range | SD |
|---|---|---|---|---|
| **/a/** | **NM** | 1.22 | 2.92 | 0.68 |
| | **NF** | 3.77 | 21.57 | 05.41 |
| | **DM** | 1.24 | 2.24 | 0.549 |
| | **DF** | 5.91 | 31.65 | 8.28 |
| **/i/** | **NM** | 1.32 | 2.00 | 0.63 |
| | **NF** | 3.77 | 21.57 | 5.41 |
| | **DM** | 1.24 | 2.24 | 0.54 |
| | **DF** | 5.91 | 31 .65 | 8.28 |
| **/u/** | **NM** | 2.80 | 8.48 | 2.36 |
| | **NF** | 14.25 | 65.26 | 18.47 |
| | **DM** | 8.58 | 36.06 | 10.82 |
| | **DF** | 12.52 | 42.12 | 14.46 |

Table 3 1 : Indicates mean, range, SD of ShdB in phonation in normal and dysphonic groups as measured using MDVP software




**1= Normal male; 2- Normal female 3- Dysphonic male; 4-Dysphonic female**

Graph 15 : Mean values of Shimmer in dB in phonation and in both normal group (males and females) and dysphonic groups (males and females) as evaluated by MDVP software.

While comparing the mean values of ShdB for phonation of different vowels it was found that the mean value of /u/ was the highest followed by /]/ and /a/ for all the groups (normal males, normal females, dysphonic males and dysphonic females). Also, the mean values of this parameter for dysphonic males and females are higher than for normal males and females which is in agreement with the study done by Von Leden et al (1968), Venkatesh et al. (1992) and Kitajima and Gould (1976). This could be attributed to the inability of the dysphonics to maintain a constant intensity in phonation. The results of this study are in accordance with the study of Anitha (1994) and Biswajith (1995).

As it could be noted from the definition, the parameters shimmer in dB (ShdB) Shimmer 96, Amplitude Perturbation Quotient (APQ), Smoothed

Amplitude Perturbation Quotient (SAPQ), Coefficient of Peak Amplitude variation (VAM) are interrelated and hence the results of all these parameters are discussed together at the end of presentation of the results of the parameter coefficient of peak amplitude variation (VAM).

*16. Shimmer 96 :* defined as the relative evaluation of the period to period very short-term variation of the peak to peak amplitude within the analyze d voice sample.

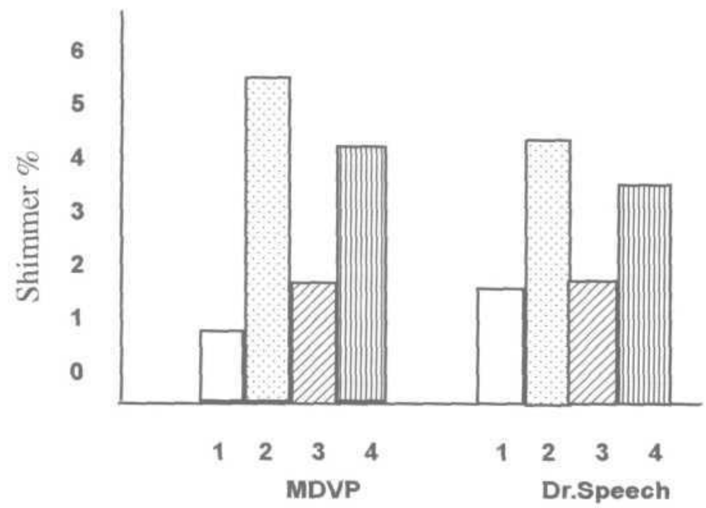
The Mean, SD and range of shimmer in percent are presented in table 31 for normal males, females and dysphonic males and females as measured by MDVP and Dr.Speech software.

| Vowels | Sub-jects | MDVP | | | Dr.Speech | | |
|---|---|---|---|---|---|---|---|
| | | Mean | Range | SD | Mean | Range | SD |
| /a/ | NM | 0.73 | 0.99 | 0.25 | 1.85 | 0.73 | 1.10-3.45 |
| | NF | 3.02 | 15.69 | 4.47 | 2.89 | 2.00 | 0.96-9.06 |
| | DM | 0.71 | 1.75 | 0.45 | 1.67 | 0.46 | 0 .97-2.46 |
| | DF | 3.98 | 10.37 | 3.62 | 2.63 | 1.70 | 0.44-7.62 |
| /iI | NM | 0.77 | 0.81 | 0.86 | 1.46 | 0.62 | 0.68-2.86 |
| | NF | 4.49 | 31.67 | 8.18 | 2.86 | 2.08 | 0.87-8.61 |
| | DM | 0.66 | 3.60 | 0.23 | 1.49 | 0.41 | 0.82-4.04 |
| | DF | 1.72 | 5.77 | 1.74 | 1.88 | 0.92 | 0.82-4.04 |
| lul | NM | 1.47 | 4.31 | 1.31 | 1.46 | 0.49 | 0.73-2.28 |
| | NF | 9.36 | 27.33 | 8.80 | 6.61 | 3.10 | 1.10-11.91 |
| | DM | 4.15 | 19.95 | 6.14 | 2.19 | 2.03 | 0.82-8.08 |
| | DF | 6.47 | 20.77 | 5.56 | 6.23 | 2.77 | 1.21-10 |

Table 32: The Mean, SD and range of normal groups and dysphonic groups for shimmer96 for phonation of/a//i/and/u/ as evaluated by MDVP and Dr.Speech software.

| Subjects | MDVP | Dr.Speech |
|---|---|---|
| NM | 0.99 | 1.58 |
| NF | 5.62 | 4.12 |
| DM | 1.83 | 1.78 |
| DF | 4.05 | 3.58 |

Table-33: Average of mean values of shim 96 in phonation of /a//i//u/ in both normal and dysphonic groups as evaluated by MDVP and Dr.Speech.




**1= Normal male; 2- Normal female 3- Dysphonic male; 4-Dysphonic female**

Graph 16 : Mean of Shimmer 96 in phonation (/a//i/ /u/) in both normal group (males and females) and dysphonic groups (males and females) as evaluated by MDVP and Dr.Speech softwares.

In the present study, while comparing the mean values of shimmer in percent for the phonation of different vowels, it was found that with both MDVP and Dr.Speech the mean value of /u/ was highest compared to /a/ and /i/ for both the normal groups (males and females) and the dysphonic groups (males and females).

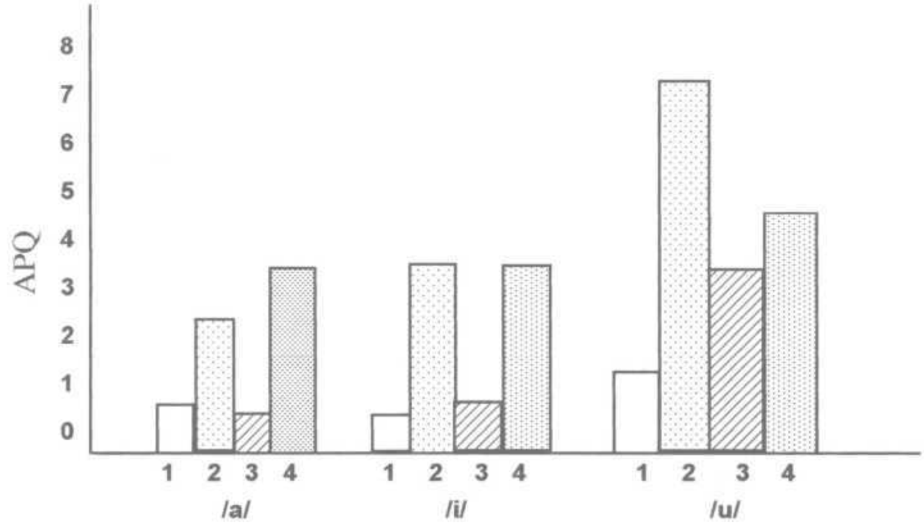
It was also found that while measured with MDVP females had lower values of shimmer (in 96) when compared to males for both groups. This finding is in agreement with the study done by Sorensen and Horii (1983). However the analysis of this parameter with Dr.Speech software revealed that females had higher values of Shimmer (in %) when compared to males for both groups, which is in disagreement with the study done by Sorensen and Horii (1983).

The variability of this parameter was more in the dysphonic group compared to the normal group. This could be attributed to the inability of the dysphonics to maintain constant intensity during phonation.

*17. Amplitude Perturbation Quotient (APQ) :* APQ is defined as relative evaluation of the period to period variability of the peak to peak amplitude within the analysed voice sample at smoothing of 11 periods.

| Vowels | Subjects | Mean | Range | SD |
|---|---|---|---|---|
| /a/ | NM | 0.56 | 0.51 | 0.14 |
| | NF | 2.26 | 13.84 | 3.73 |
| | DM | 0.46 | 1.25 | 0.30 |
| | DF | 3.28 | 12.57 | 3.52 |
| /i/ | NM | 0.45 | 0.60 | 0.16 |
| | NF | 3.43 | 26.28 | 6.76 |
| | DM | 0.63 | 2.34 | 0.56 |
| | DF | 3.40 | 5.23 | 3.54 |
| /u/ | NM | 1.13 | 3.56 | 1.03 |
| | NF | 7.08 | 21.09 | 6.37 |
| | DM | 3.11 | 13.28 | 4.78 |
| | DF | 4.73 | 15.73 | 3.99 |

Table 34: Indicates the Mean, range SD values of APQ for phonation of /a/ /i/ /u/ in normal males, normal females, dysphonic males and dysphonic females.




**1= Normal male; 2- Normal female 3- Dysphonic male; 4-Dysphonic female**

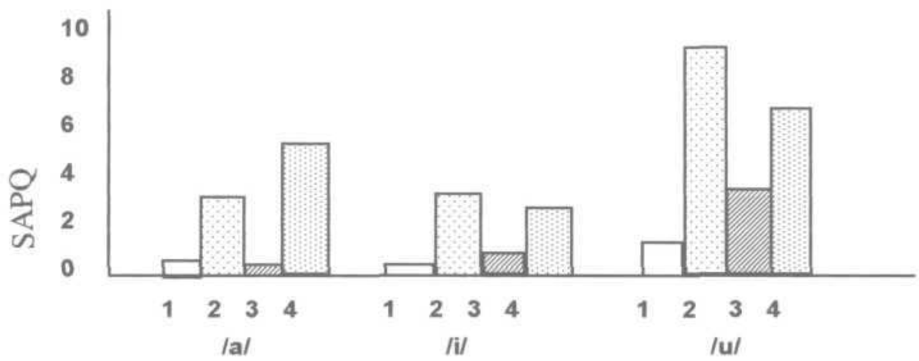
Graph 17 : Mean of APQ in phonation in both normal group (males and females) and dysphonic groups (males and females) as evaluated by MDVP software.

When mean values for APQ were compared for phonation of different vowels, it was observed that the mean value of /u/ was highest followed by /i/ and /a/ for all the groups (normal males, normal females, dysphonic males and dysphonic females). Also it was seen that the mean value of the dysphonics was higher than the normals for all the groups for the phonation of vowels /a/ /i/ and /u/. Also the dysphonics showed greater variability than the normals which can be attributed to the inability of the dysphonics to maintain constant intensity level during phonation.

7 *8. Smoothed Amplitude Perturbation Quotient (SAPQ)* : The mean, SD and range for SAPQ for the four groups (normal males, normal females, dysphonics males and dysphonic females are presented in table 35 as evaluated by MDVP software.

| Vowels | Subjects | Mean | Range | SD |
|---|---|---|---|---|
| /a/ | NM | 0.71 | 1.23 | 0.32 |
| | NF | 3.04 | 16.08 | 4.47 |
| | DM | 0.58 | 1 .409 | 0.34 |
| | DF | 5.73 | 21.93 | 7.01 |
| /i/ | NM | 0.51 | 0.66 | 0.20 |
| | NF | 3.25 | 16.36 | 4.72 |
| | DM | 0.88 | 7.09 | 1.76 |
| | DF | 2.09 | 10.66 | 2.90 |
| | NM | 1.21 | 3.55 | 1.02 |
| | NF | 9.25 | 25.52 | 7.59 |
| | DM | 3.75 | 18.38 | 6.17 |
| | DF | 6.54 | 18.64 | 5.49 |

Table 35: Indicates the mean, range, SD values of SAPQ for phonation in normal males, normal females, dysphonic males and females as evaluatd by MDVP software.




**1= Normal male; 2- Normal female 3- Dysphonic male; 4-Dysphonic female**

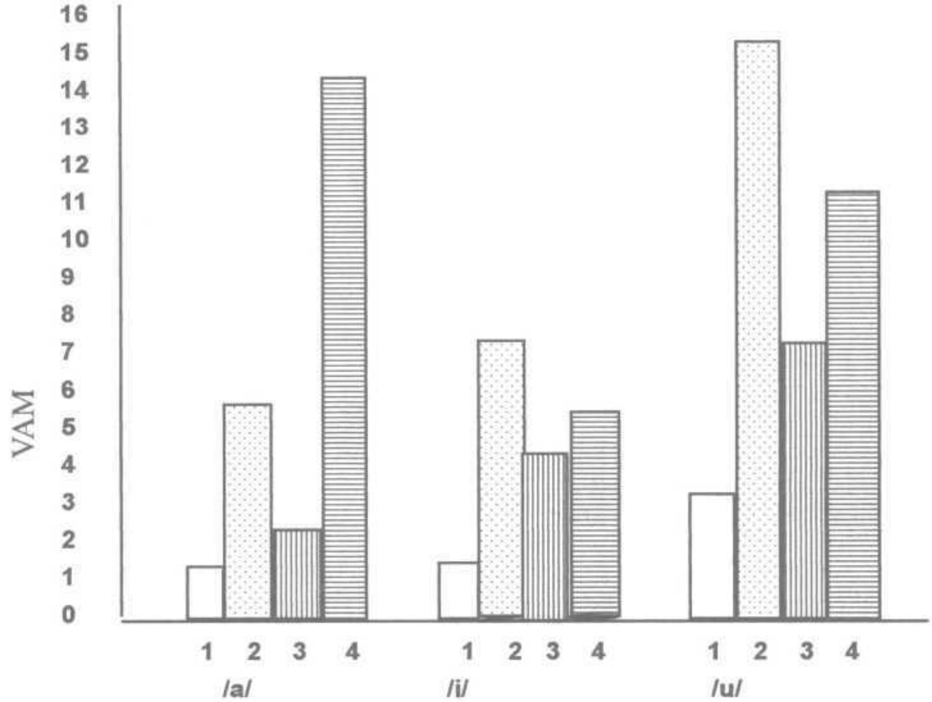
Graph 18 : Mean values of Shimmer in dB in phonation and in both normal group (males and females) and dysphonic groups (males and females) as evaluated by MDVP software.

While comparing the mean values of SAPQ for phonation of different vowels it was found that in normal males, the order of increase in the mean of SAPQ were for /u/ (1.21); /a/ (.7 17) and/i/(.505). In case of dysphonic males and normal females the value of/u/ was highest compared to/a/ and /i/. The dysphonic females group showed the highest value of/u/ (6.54), followed by/a/ (5.73) and /i/ (2.09). It was also seen that the mean values of dysphonic males and females for the vowels were higher than for normal males and females and that the dysphonic group exhibited greater variability than the normal group.

*19. Variation in Amplitude (VAM)* : VAM is defined as relative standard deviation of the peak to peak amplitude. The mean, SD and range are presented for normal males, normal females, dysphonic males and dysphonic females as measured using MDVP software.

| Vowels | Subjects | Mean | Range | SD |
|---|---|---|---|---|
| /a/ | NM | 1.07 | 3.20 | 0.74 |
| | NF | 5.93 | 38.57 | 9.94 |
| | DM | 2.02 | 8.60 | 2.68 |
| | DF | 14.37 | 36.08 | 13.01 |
| /i/ | NM | 1.24 | 6.18 | 1.50 |
| | NF | 6.98 | 36.26 | 9.20 |
| | DM | 4.14 | 22.36 | 10.28 |
| | DF | 5.32 | 39.55 | 7.14 |
| /u/ | NM | 3.00 | 10.96 | 3.13 |
| | NF | 15.20 | 38.48 | 12.41 |
| | DM | 7.04 | 32.00 | 9.78 |
| | DF | 11.00 | 34.31 | 10.51 |

Table 36: Mean, SD and range values of coefficient of amplitude variation in phonation for normal males, normal females dysphonic males and dysphonic females as measured using MDVP software.




**1- Normal male; 2- Normal female 3- Dysphonic male; 4-Dysphonic female**

Graph 19 : Mean of VAM in phonation and in both normal group (males and females) and dysphonic groups (males and females) as evaluated by MDVP software.

In the present study, while comparing the mean values of VAM for phonation of different vowels, in case of normal males and females the mean values of /a/ and /i/ was lower than /u/. In case of dysphonic males also the mean value of /u/ was higher than /a/ and */i/.* However, in case of dysphonic females, the mean value of /a/ was highest followed by /u/ and */i/.* The dysphonic group (males and females) indicated higher mean values than the normal group. The variability was also higher in the dysphonic group compared to the normal group.

**Discussion**

In all the following parameters - Shimmer dB, Shimmer 96, Amplitude Perturbation Quotient (APQ), Smoothed Amplitude Perturbation Quotient (SAPQ) and Coefficient of Amplitude Variation, the mean values for dysphonic males and females were higher than for normal males and females which is in agreement with the study done by VonLeden et al. (1960); Venkatesh et al. (1992) and Kitajima and Gould (1976). This could be attributed to the inability of the dysphonics to maintain a constant intensity in phonation.

However, it was seen that pitch extraction errors may affect voice very well with a smoothing factor of 11, SAPQ is identical to the APQ introduced by Koike (Koike, 1973, Koike and Calcatera, 1977). Because of smoothing, APQ is not as sensitive to pitch extraction errors. While it is less sensitive to the period to period amplitude variations, it still describes the short term amplitude perturbation of the voice very well.

At high smoothing factors, SAPQ correlates with the intensity of the long term peak to peak amplitude variations. The studies of patients with spasmodic dysphonia (Deliyski et al. 1991) show that SAPQ with a smoothing factor set in the range 45-65 periods has increased values in case of regular long term amplitude variations.

The SAPQ smoothing factor set up is 55 periods - SAPQ (55). This set up allows using SAPQ as an additional evaluation of the amplitude tremors in the voice. The intensity and the regularity of the amplitude tremors can be assessed using SAPQ ' .5) in combination with VAM. The manufacturers suggest the use of APQ/SAPQ with VAM instead of shimmer in order to avoid the influence of pitch extraction errors. Hence the mean values of SAPQ and VAM were compared, for the dysphonic males and females.

It was found that the mean SAPQ (55) values were lower for dysphonic males and females when compared to VAM. This indicates that the short term variations were more in case of dysphonics (both males and females).

*20. Noise to Harmonic Ratio (NHR) :* Defined as the average ratio of the inharmonic spectral energy inthe frequency range 1500 - 4500 Hz to the harmonic spectral energy in the frequency range 70 to 4500 Hz.

The mean, SD and range for NHR are presented in the table 37 for normal males, females, dysphonic males and females respectively.

| Vowels | Subjects | Mean | Range | SD |
|---|---|---|---|---|
| /a/ | NM | 0.16 | 0.56 | 0.13 |
| | NF | 0.22 | 0.58 | 0.16 |
| | DM | 0.18 | 0.52 | 0.16 |
| | DF | 0.27 | 0.20 | 5.2 |
| /i/ | NM | 0.12 | 0.12 | 3.28 |
| | NF | 2.36 | 14.79 | 5.1 1 |
| | DM | 1 .09 | 0.10 | 3.18 |
| | DF | 2.12 | 0.17 | 3.44 |
| /u/ | NM | 0.17 | 0.22 | 6.39 |
| | NF | 0.37 | 0.65 | 0.17 |
| | DM | 0.24 | 0.50 | 0.14 |
| | DF | 2.24 | 29.05 | 7.44 |

Table 37: Indicates mean, range, SD values of NHR for/a//i/ and/u/ in both normal groups and dysphonic groups as measured by MDVP software.

While comparing the mean values of NHR for phonation of different vowels, it was found that the mean value of /u/ was greater when compared to mean values of /a/, /i/ in males and females of both the groups. It was also seen that the value of NHR increased in both dysphonic males and females when compared to normal males and females which is in agreement with the study done by Kitajima (1981).

The increase in the value of NHR for phonation of vowel /u/ could be discussed as follows:

Unlike for phonation of other vowels, during the phonation of /u/, lip rounding takes place. This results in directing a stream of air directly on the microphone resulting in an increase in the noise energy picked up by the microphone which eventually increases the value of NHR. However for other vowels like /]/ and /a/ there is a lip spreading and an open mouth position respectively which results in the movement of the airstream in different directions and thereby only a fraction of it will be received by the microphone and eventually the NHR value goes low.

*2I. Voice Turbulence Index (VT/) :* Defined as the average ratio of the spectral inharmonic high frequency energy in the range 2800 - 5800 Hz to the spectral harmonic energy in the range 70-4500 Hz inthe areas of the signal where the influenceof the frequencyand amplitude variations. Voice breaks and subharmonic components are minimal.

VTI mostly correlates with the turbulence caused by incomplete or loose adduction of the vocal folds. It analyses high frequency components to extract an acoustic correlate to "breathiness".

| Vowels | Subjects | Mean | Range | SD |
|---|---|---|---|---|
| /a/ | NM | 0.56 | 7.63 | 1.93 |
| | NF | 0.69 | 9.36 | 2.41 |
| | DM | 0.64 | 8.05 | 2.06 |
| | DF | 0.21 | 6.60 | 5.07 |
| /i/ | NM | 4.62 | 1.07 | 1.06 |
| | NF | 0.15 | 0.49 | 0.16 |
| | DM | 6.02 | 0.01 | 2.12 |
| | DF | 0.13 | 5.01 | 2.94 |
| /u/ | NM | 5.02 | 0.06 | 2.06 |
| | NF | 0.09 | 1.05 | 3.20 |
| | DM | 1 .27 | 6.67 | 4.76 |
| | DF | 0.51 | 18.48 | 1.70 |

Table 38: Mean, range, SD values of VTI during phonation for normal groups (males and females and dysphonic groups) as evaluated by MDVP software.

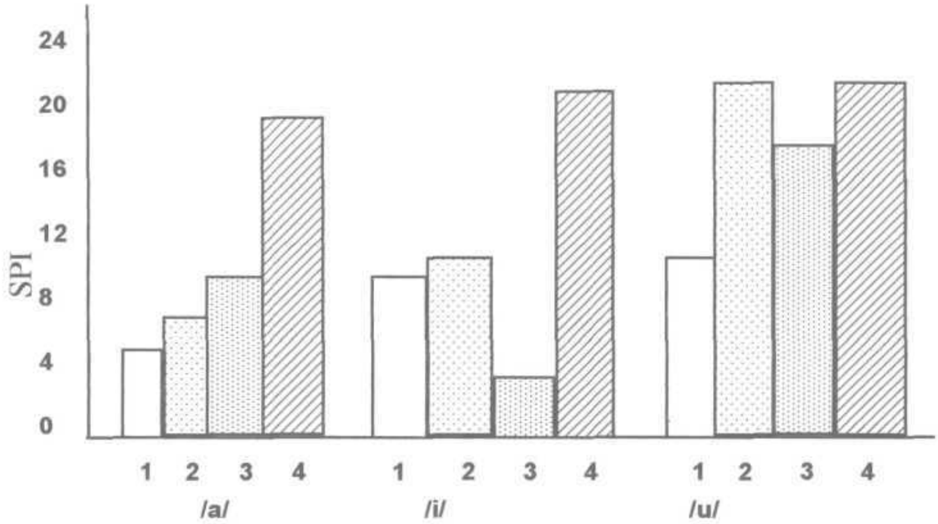
While comparing the mean values of VTI for phonation of different vowels, it was found that the mean value of /u/ was highest compared to /a/ and /i/ in case of normal males. In case of dysphonic males the mean value of /i/ was highest compared to /a/ and /u/. In case of normal females the mean value of /u/ was highest compared to /i/ and /u/ and in case of dysphonic females the mean value of /u/ was highest compared to /a/ and /i/.

*22. Soft Phonation Index (SPI)* defined as the average ratio of the low frequency harmonic energy in the range of 70-1600 Hz to the high frequency harmonic energy in the range 1600 - 4500 Hz.

The mean, SD and range of the SPI are presented in table 39 for normal males and females and dysphonic males and females.

| Vawels | Subjects | Mean | Range | SD |
|---|---|---|---|---|
| /a/ | NM | 4.63 | 12.70 | 2.88 |
| | NF | 7.41 | 25.88 | 6.84 |
| | DM | 8.96 | 27.18 | 9.56 |
| | DF | 19.39 | 61.13 | 19.95 |
| /i/ | NM | 7.07 | 13.33 | 3.71 |
| | NF | 10.88 | 65.25 | 16.66 |
| | DM | 2.95 | 0.04 | 1.09 |
| | DF | 19.65 | 57.12 | 18.75 |
| /u/ | NM | 10.77 | 27.40 | 6.90 |
| | NF | 21.10 | 49.78 | 16.25 |
| | DM | 17.64 | 52.57 | 16.21 |
| | DF | 21.15 | 83.13 | 20.68 |

Table 39: Mean, range and SD values of SPI for/a//i//u/ for normal groups (males and females) and dysphonic groups (males and females) as measured using MDVP software.




**1= Normal male; 2- Normal female 3- Dysphonic male; 4-Dysphonic female**

Graph 20 : Mean of SPI in phonation in both normal group(males and females) and dysphonic groups (males and females) as evaluated by MDVP software.

In the present study, when comparing the mean values of SPI for phonation of different vowels, it was found that in case of normal females, the mean values of /i/ was lesser when compared to /a/ and /u/. In case of normal males, dysphonic males and dysphonic females, the mean value of /u/ was highest when compared to /a/ and /i/.

The increase in the value of SPI for phonation of the vowel /u/ could be discussed as follows:

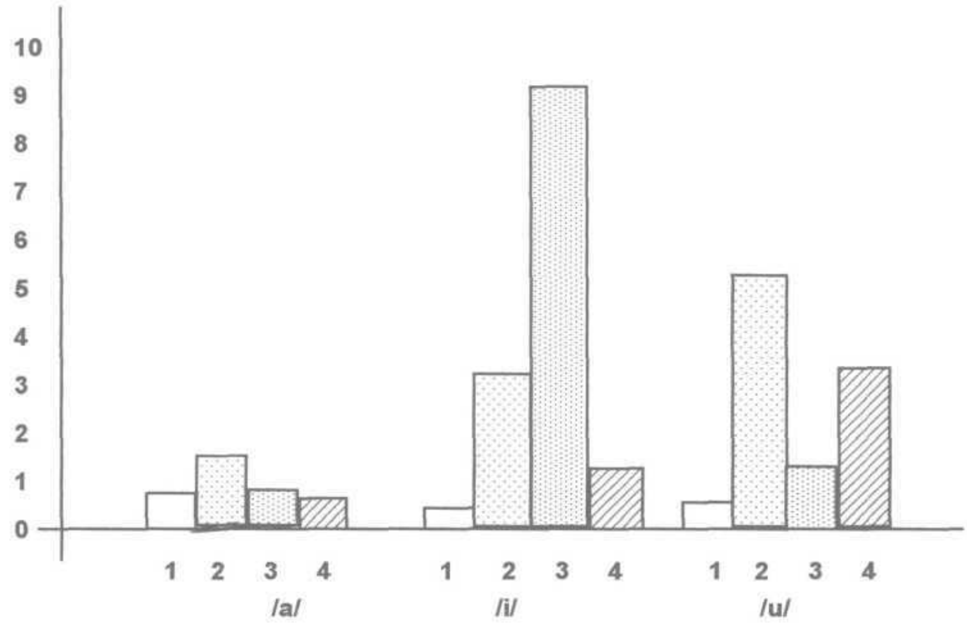
Unlike for phonation of other vowels, during the phonation of /u/ lip rounding takes place, resulting in directing a stream of air directly on the microphone. This in turn results in an increase in the noise energy picked up by the microphone which eventually increases the values of SPI. However for other vowels like/i/, /a/ there is a lip spreading and an open mouth position respectively which results in the movement of the air stream in different directions and thereby only a fraction of it will be received by the microphone and eventually the SPI value goes low.

*23. Frequency Tremor Intensity Index (FTRI) :* It is defined as the average ratio of the frequency magnitude of the most intensive low frequency modulating component to the total frequency magnitude of the analyzed voice sample.

The mean, SD and range are presented for normal males, normal females dysphonic males and dysphonic females in table 40.

| Vowels | Subjects | Mean | Range | SD |
|---|---|---|---|---|
| /a/ | NM | 0.98 | 8.44 | 2.08 |
| | NF | 1.78 | 9.56 | 2.66 |
| | DM | 0.90 | 8.40 | 2.12 |
| | DF | 0.86 | 3.54 | 1.03 |
| /i/ | NM | 0.41 | 1.06 | 0.25 |
| | NF | 3.08 | 18.90 | 4.95 |
| | DM | 19.33 | 42.53 | 14.81 |
| | DF | 1.18 | 4.62 | 1.34 |
| /u/ | NM | 0.53 | 1.15 | 0.28 |
| | NF | 5.09 | 27.20 | 8.41 |
| | DM | 1.38 | 8.47 | 1.76 |
| | DF | 3.35 | 15.29 | 4.93 |

Table 40: Mean, range, SD values of FTRI for/a//i//u/, for normal groups (males and females) and dysphonic groups (males and females) as measured by MDVP software.




**1= Normal male; 2- Normal female 3- Dysphonic male; 4-Dysphonic female**

Graph 21 : Mean of FTRI in phonation in both normal group (males and Females) and dysphonic groups (males and females) as evaluated by MDVP software.

In the present study, while comparing the mean values of FTRI for phonation of different vowels it was indicated that for normal males, the mean value of /a/ was highest followed by /u/ and */i/.* In case of dysphonic males, the mean value of/u/ was the highest for normal females the mean value of/i/was highest and fordysphonic females the mean value of/u/ was the highest compared to /a/ and */i/.*

*24. Amplitude Tremor Intensity Index (ATRI)* : defined as the average ratio of the amplitude of the most intense low frequency amplitude modualting component to the total amplitude of the analysed voice signal.

The mean, SD and range for ATRI are presented for four groups : normal males, dysphonic males, normal females and dysphonic females.

| Vowels | Subjects | Mean | Range | SD |
|---|---|---|---|---|
| /a/ | NM | 0.25 | 0.55 | 0.17 |
| | NF | 1.75 | 16.95 | 4.34 |
| | DM | 0.66 | 5.31 | 1.31 |
| | DF | 3.09 | 16.79 | 4.86 |
| /i/ | NM | 0.22 | 0.63 | 0.20 |
| | NF | 1.79 | 13.20 | 3.34 |
| | DM | 0.29 | 0.51 | 0.13 |
| | DF | 1.37 | 11.49 | 2.95 |
| /u/ | NM | 1.32 | 8.81 | 2.28 |
| | NF | 4.69 | 15.79 | 5.33 |
| | DM | 1.65 | 7.60 | 2.61 |
| | DF | 3.75 | 15.72 | 4.80 |

Table 41 : Indicates mean, range, SD values for normal males, normal females, dysphonic males and dysphonic females for ATRI for phonation of /a/ /i/ and /u/ as measured by MDVP software.

Comparision of mean values of ATRI for phonation of different vowels indicated that in normal males the mean value of /u/ was highest compared to /a/ and /i/. In case of dysphonic males, the mean value of /u/ was highest. In case of normal females and dysphonic females the mean value also the mean value of /u/ was highest compared to /a/ and /i/. The mean values of the dysphonic group was higher than that of the normal group for phonation of /a/ /i/ and /u/ and the dysphonics exhibited greater variability than the normal group.

*25. Degree of Voice Breaks (DVB)* : It is defined as ratio of the total length of areas representing voice breaks to the time of the complete voice sample. It measures the ability of voice to sustain uninterrupted voicing. The normative threshold is zero, because a normal voice during the task of sustained voice will not have any voice breaks.

| Vowels | Subjects | Mean | Range | **SD** |
|---|---|---|---|---|
| ***/a/*** | **NM** | 0 | 0 | 0 |
| | **NF** | 0 | 0 | 0 |
| | **DM** | 0 | 0 | 0 |
| | **DF** | 3.88 | 49.29 | 12.69 |
| ***/i/*** | **NM** | 0 | **0** | 0 |
| | **NF** | 0 | 0 | 0 |
| | **DM** | 0 | 0 | 0 |
| | **DF** | 0.11 | 1.22 | 0.32 |
| ***/u/*** | **NM** | 0 | 0 | 0 |
| | **NF** | 0 | 0 | 0 |
| | **DM** | 0 | 0 | 0 |
| | **DF** | 4.34 | 14.28 | 4.77 |
| ***/sp/*** | **NM** | 6.52 | 33.50 | 10.73 |
| | **NF** | 8.65 | 51 .65 | 13.32 |
| | **DM** | 7.81 | 32.71 | 10.52 |
| | **DF** | 13.32 | 58.39 | 14.17 |

Table 42: Indicates mean, range and SD values for/a/ /V and /u/ and sentence for DVB for both normal groups (mates *&TKI* females) and dysphonic groups (males and females) as measured by MDVP software.

In the case of dysphonic males and females, the mean value of DVB for sentence was higher than for phonation. The result of this parameter is consistent with the study done by Anitha (1 994) and Biswajit (1 995).

The results of this parameter can be discussed as follows:

It was seen that the "DVB" in normal females and males were more in sentence than in phonation. This could be due to the presence of pause in the speech sample which increases the value of degree of voice breaks in sentence but it is not so in case of phonation. In the case of dysphonic males and females the "degree of voice breaks" were higher in phonation and sentence. This is because of the irregular vibration of the vocal folds caused due to the pathological conditions of the larynx. However, the mean value of DVB were higher in case of sentences due to the presence of pauses in between in the sentence.

*26. Number of Subharmonic Segments (NSH)* : The mean, SD and range of NSH are presented in table 43.

| Vowels | Subjects | Mean | Range | SD |
|---|---|---|---|---|
| **/a/** | **NM** | 0 | 0 | 0 |
| | **NF** | 0.33 | 6.00 | 1.29 |
| | **DM** | 0 | 0.30 | 0.82 |
| | **DF** | 0.66 | 4.00 | 1 .17 |
| **/i/** | **NM** | 0 | 0 | 0 |
| | **NF** | 0.46 | 7.00 | 1.80 |
| | **DM** | 0 | 0 | 0 |
| | **DF** | 0.13 | 1 .00 | 0.35 |
| **/u/** | **NM** | 0 | 0 | 0 |
| | **NF** | 3.13 | 16 | 4.34 |
| | **DM** | 0 | 0 | 0 |
| | **DF** | 5.40 | 18.00 | 6.19 |

Table 43: Mean, range, SD values of NSH during phonation for normal groups (males and females) and dysphonic groups (males and females) as measured by MDVP software.

The inspection of the table showed that for normal females the NSH values are negligible for phonation of /a/ /i/ and /u/. The dysphonic group exhibited higher NSH values than the normal group. The results of this parameter are consistent with the findings reported by Anitha (1994) and Biswajit (1995).

The above result can be discussed as follows:

The mean values of NSH for dysphonic group were higher than in the normal group which could be due to irregular vibratory pattern of the vocal folds, which is seen in dysphonics which would result in more than one frequency of vibration at a given instance leading to an increase in the value of number of subharmonic segments.

*27. Degree of Subharmonic Components (DSH)* : It is defined as the relative evaluation of sub-harmonic to Fo components in the voice sample. The mean, SD and range for DSH are presented in the table 44 for normal males, normal females, dysphonic males and dysphonic females.

| Vowels | Subjects | Mean | Range | SD |
|---|---|---|---|---|
| /a/ | NM | 0 | 0 | 0 |
| | NF | 0.68 | 10.20 | 2.63 |
| | DM | 0 | 0 | 0 |
| | DF | 0.96 | 4.88 | 1.57 |
| /i/ | NM | 0 | 0 | 0 |
| | NF | 2.91 | 43.75 | 11.29 |
| | DM | 0 | 0 | 0 |
| | DF | 0.19 | 2.27 | 0.60 |
| /u/ | NM | 0 | 0 | 0 |
| | NF | 6.47 | 19.44 | 7.30 |
| | DM | 0 | 0 | 0 |
| | DF | 9.40 | 42.85 | 11.90 |

Table 44: Indicates mean, SD, range for DSH for phonation in normal groups of males and females and dysphonic groups (males and females)as measured by MDVP Software

From the above table, it can be noted that the dysphonic group (males and females) exhibited higher mean values and higher variability than the normal group.

In case of dysphonic males, the mean value of/a/ was highest compared to /i/ and /u/, while in dysphonic females, the mean value of /u/ was highest compared to /i/ and /a/. The results of this parameter are consistent with the studydone by Anitha (1994) and Biswajit (1995).

*28. Degree of Voiceless (DUV)* : DUV is the estimated relative evaluation of nonharmonic areas in the voice sample. Table 45 represents the mean, SD and range of DUV for normal males and females, dysphonic males and females.

| Vbwels | Subjects | Mean | Range | SD |
|---|---|---|---|---|
| /a/ | **NM** | 0 | 0 | 0 |
| | **NF** | 12.51 | 92.11 | 27.84 |
| | **DM** | 0.21 | 3.125 | 0.81 |
| | **DF** | 14.35 | 77.50 | 23.02 |
| **/i/** | **NM** | 0 | 0 | 0 |
| | **NF** | 10.11 | 97.47 | 25.58 |
| | **DM** | 0 | 0 | 0 |
| | **DF** | 4.25 | 38.23 | 9.98 |
| **/u/** | **NM** | 0 | 0 | 0 |
| | **NF** | 18.27 | 75.30 | 22.02 |
| | **DM** | 0 | 0 | 0 |
| | **DF** | 10.92 | 45.69 | 16.03 |
| Sentence | **NM** | 33.97 | 55.11 | 13.55 |
| | **NF** | 35.88 | 58.12 | 14.68 |
| | **DM** | 41.83 | 57.62 | 16.96 |
| | **DF** | 43.64 | 64.14 | 14.17 |

Table 45: Mean, range, SD values of DUV in normal groups and dysphonic groups in phonation of vowels and in speech.

While comparing the mean values of DUV for phonation of different vowels and sentence it was found that in normal males and females the mean value of sentence was highest compared to /a/ /i/ and /u/. In the case of dysphonic males, the mean values of /u/ was the highest when compared to/a/ and /i/. The mean value of sentence was highest when compared to /a/ /i/ and /u/.

In case of dysphonic females, the mean value of sentence was highest compared to /a//i/ and /u/.

. The above results show that the degree of voiceless was higher for sentence than in phonation of vowels /a/ /i/ and /u/ in all the four groups normal males, normal females, dysphonic males and dysphonic females. This is because of the presence of pauses in between the words in the speech sample but in phonation, this is not so. The rsult of this parameter is in accordance with the study done by Anitha (1 994) and Biswajit (1 995).

*29. Number of Voice Breaks (NVB)* : NVB is the number of times the fundamental period was interrupted during the voice sample. The mean, SD and range values are presented in the table 46.

| Vowels | Subjects | Mean | Range | SD |
|---|---|---|---|---|
| /a/ | NM | 0 | 0 | 0 |
| | NF | 6.66 | 1.00 | 0.26 |
| | DM | 0 | 0 | 0 |
| | DF | 1.53 | 20.00 | 5.13 |
| /i/ | NM | 0 | 0 | 0 |
| | NF | 0.60 | 3.00 | 1.12 |
| | DM | 0 | 0 | 0 |
| | DF | 0.13 | 1.00 | 0.35 |
| /u/ | NM | 0 | 0 | 0 |
| | NF | 3.40 | 12.00 | 4.59 |
| | DM | 0 | 0 | 0 |
| | DF | 3.26 | 12.00 | 3.61 |
| /sp/ | NM | 18.53 | 13.00 | 3.64 |
| | NF | 20.20 | 112.00 | 23.30 |
| | DM | 21.20 | 19.00 | 5.63 |
| | DF | 21.53 | 115.00 | 20.60 |

Table 46: Indicates the mean, range, SD values of NVB in phonation for normal groups (males and females) and dysphonic groups (males and females) as measured by MDVP software

The voice break areas in phonation of vowels were zero in case of normal subjects but in sentences it was present. The number of voice breaks in sentences were greater than those in phonation. In case of dysphonic males and females the voice breaks were present in phonation and sentences. The results of this parameter and in accordance with the study done by Anitha (1994) and Biswajit (1995).

The above results are discussed as follows:

The number of voice break areas in the phonation of vowels were zero, but in sentence it was present. This is because the speech sample has pauses between the words which increases the value of "number of voice breaks" and this is not so in case of phonation. In case of dysphonic males and females, the voice breaks were present in phonation and sentence. This is due to the irregular vibration of vocal folds caused due to pathological conditions of the larynx. However, the mean value of "number of voice breaks "were higher in sentence than in phonation which could be attributed to the reason mentioned earlier.

The results and discussion of parameters evaluated using Vaghmi software are presented as follows:

1 . Fundamental Frequency in Phonation

This was the common parameter also evaluated by MDVP software and Dr.Speech software.

The results of the above parameter are already discussed along with the parameter Average Fundamental Frequency" evaluated by MDVP software.

2. Maximum Fundamental Frequency

This parameter was also evaluated by Dr.Speech software and MDVP software.

The results of this parameter are already discussed along with the parameter "Highest fundamental frequency" evaluated by MDVP software.

3. Minimum Fundamental Frequency

This parameter was also evaluated by Dr.Speech software program and MDVP software program.

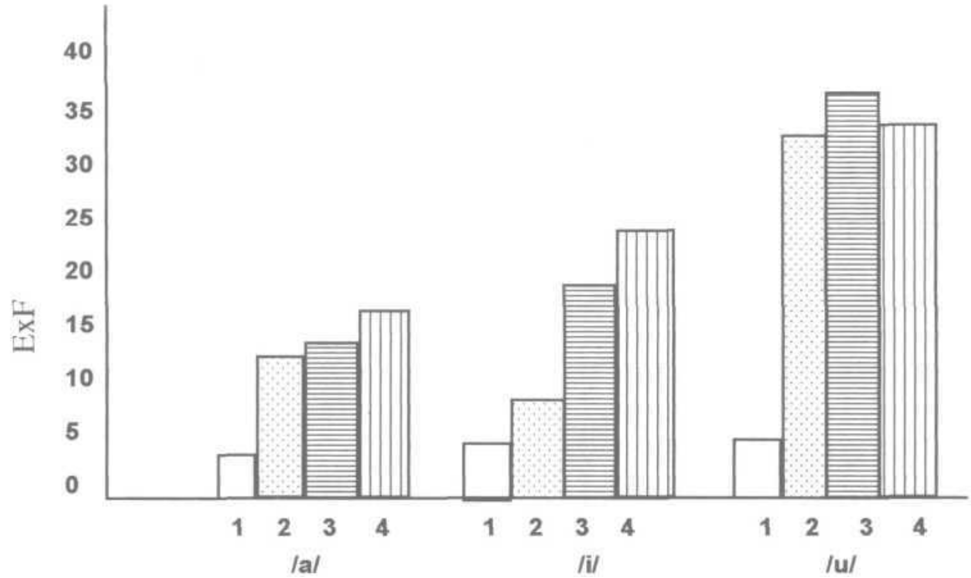
The results of this parameter are already discussed under the parameter "Lowest fundamental frequency' evaluated by MDVP software.

4. Extent of Fluctuation in Fundamental Frequency

Fluctuations in fundamental frequency in phonation, in terms of extent of fluctuation has been defined as the average of deviations in fundamental frequency +/-3 and beyond a sample of 1 sec.

| Vowels | Subjects | Mean | Range | SD |
|---|---|---|---|---|
| /a/ | NM | 3.47 | 0.74 | 0.22 |
| | NF | 12.41 | 33.06 | 10.07 |
| | DM | 14.31 | 48.94 | 16.53 |
| | DF | 15.21 | 74.91 | 16.41 |
| /i/ | NM | 3.99 | 2.57 | 0.83 |
| | NF | 8.32 | 21.64 | 7.30 |
| | DM | 18.80 | 75.99 | 24.36 |
| | DF | 23.72 | 50.23 | 18.84 |
| /u/ | NM | 3.91 | 5.63 | 1.58 |
| | NF | 32.74 | 66.55 | 16.51 |
| | DM | 36.40 | 16.96 | 35.44 |
| | DF | 34.51 | 108.57 | 34.44 |

Table 47: Shows mean values of dysphonic males and females and their normal counter parts for extent of fluctuation in fundamental frequency in phoiiation as measured by Vaghmi software.




**1= Normal male; 2- Normal female 3- Dysphonic male; 4-Dysphonic female**

Graph 22: Mean of extent of fluectuation in fundamental frequency in phonation in both normal group (males and females) and dysphonic groups (males and females) as evaluated by Vaghmi software.

In the present study it was found the dysphonic males and females had a greater number of fluctuation than the normal males and females for all the vowels /a/ /i/ and /u/. The abnormal extent of fluctuation in fundamental frequency in dysphonics suggested irregular vocal fold vibrations indifferent types of voice disorders. Similar results have been found by Kim et al. (1982), Vbon et al. (1984), Imiazumi et al. (1 980), Nataraja (1 986) and others.

| Investigators | Male | Female |
|---|---|---|
| Vanaja (1986) | 3.87 | 3.56 |
| Rajashekar(1991) | 3.0 | — |
| Tharmar(1991) | 2.75 | 3.59 |
| Suresh (1991) | 3.44 | 4.12 |
| Krishnan (1992) | 19.13 | 8.55 |
| Prabha(1997) | 1 .94 | 2.36 |
| Pradeep(1997) | 2.95 | 3.41 |
| Rajkumar(1998) | 3.89 | 4.64 |
| David (1998) | 1.58 | 4.87 |
| Present Study | 3.47 | 12.41 |

Table 48: The values of extent of fluctuation in frequency (in Hz) for phonation of vowel /a/ in normal male and females as reported by various investigators.

| Investigators | Male | Female |
|---|---|---|
| Nataraja (1986) | 28.90 | 24.79 |
| Prabha (1997) | 3.64 | 2.37 |
| David (1998) | 5.04 | 4.92 |
| Present study | 14.31 | 15.21 |

Table 49: The values of extent of fluctuation in frequency (in Hz) for phonation of vowe|/a/ in the "dysphonic" male and female as given by various investigators.

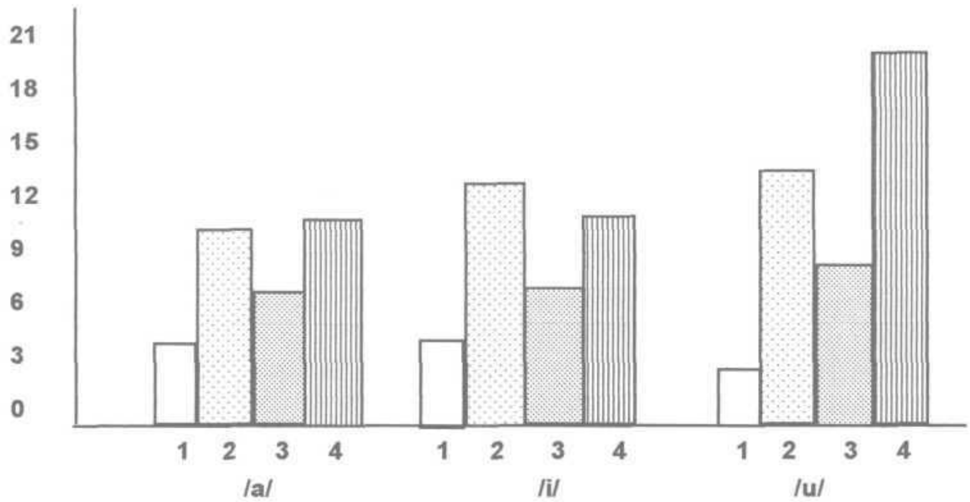
Thus the results of this parameter correlates with those of the earlier investigators.

5. Speed of Fluctuation in Fundamental Frequency

Speed of fluctuation has been defined as the number of fluctuations in fundamental frequency (+/- 3 Hz) and beyond) in a phonation of 1 sec.

| \4>wels | Subjects | Mean | Range | SD |
|---|---|---|---|---|
| ***/a/*** | **NM** | 2.76 | 4.65 | 1.50 |
| | **NF** | 10.19 | 28.65 | 10.4 |
| | **DM** | 6.24 | 14.15 | 4.23 |
| | **DF** | 10.99 | 1 5.77 | 4.38 |
| ***/i/*** | **NM** | 3.03 | 5.48 | 1.55 |
| | **NF** | 12.32 | 61.59 | 17.31 |
| | **DM** | 6.34 | 15.10 | 5.18 |
| | **DF** | 10.74 | 25.38 | 6.95 |
| ***/u/*** | **NM** | 2.06 | 3.19 | 0.87 |
| | **NF** | 13.55 | 58.68 | 8.18 |
| | **DM** | 8.42 | 39.03 | 1 1.13 |
| | **DF** | 20.20 | 27.55 | 21.34 |

Table 50: Mean, range, SD of speed of fluctuation in fundamental frequency of different vowels for normal males, dysphonic males, normal females and dysphonic females as evaluated by Vaghmi software.




**1= Normal male; 2- Normal female 3- Dysphonic male; 4-Dysphonic female**

Graph 23 : Mean of speed of fluctuation in fundamental frequency in phonation in both normal group (males and females) and dysphonic groups (males and females) as evaluated by Vaghmi software.

The males of the normal group and dysphonic groups showed low scores in terms of fluctuations compared to the females.

Studies reported in the literature also showed a lower values for males than for females.

| Investigators | Male | Female |
|---|---|---|
| Nataraja (1986) | 5.60 | 6.18 |
| Tharmar (1991) | 0.80 | 1.95 |
| Rajashekar(1991) | 5.70 | - |
| Suresh (1991) | 2.54 | 4.49 |
| Krishnan(1992) | 8.73 | 15.42 |
| Prabha(1997) | 1.12 | 4.14 |
| Pradeep(1997) | 6.20 | 8.37 |
| Rajkumar (1998) | 14.86 | 22.01 |
| David (1998) | 1.30 | 6.31 |
| Present study | 2.76 | 6.24 |

Table 51 : The values of speed of fluctuation in frequency (in fluctuation/sec) for sustained phonation of /a/ shown by normal males and females as reported by various investigators.

| Investigators | Male | Female |
|---|---|---|
| Nataraja (1986) | 47.59 | 48.37 |
| Prabha (1997) | 5.86 | 9.31 |
| Pradeep(1997) | 17.33 | 19.24 |
| Present study | 10.19 | 10.99 |

Table 52: The values of speed of fluctuation in frequency (in flue/sec) for sustained phonation of/a/ shown by 'dysphonic' males and females as reported by various investigators.

The dysphonic males and females had greater number of fluctuation than normal males and females.

This parameter is hence useful in differentiating dysphonics and normals of both the sexes.

These results have correlated with the findings of earlier investigators. Nataraja (1 986) had concluded that SFF was an important parameter in differentiating normals from dysphonics. Shobha (1 996) had reported speed of fluctuations in good voices (in Indian professional voice uses) and got similar results.

From the above results, it was considered that SFF was one of the parameter useful in differentiating normals from dysphonic.

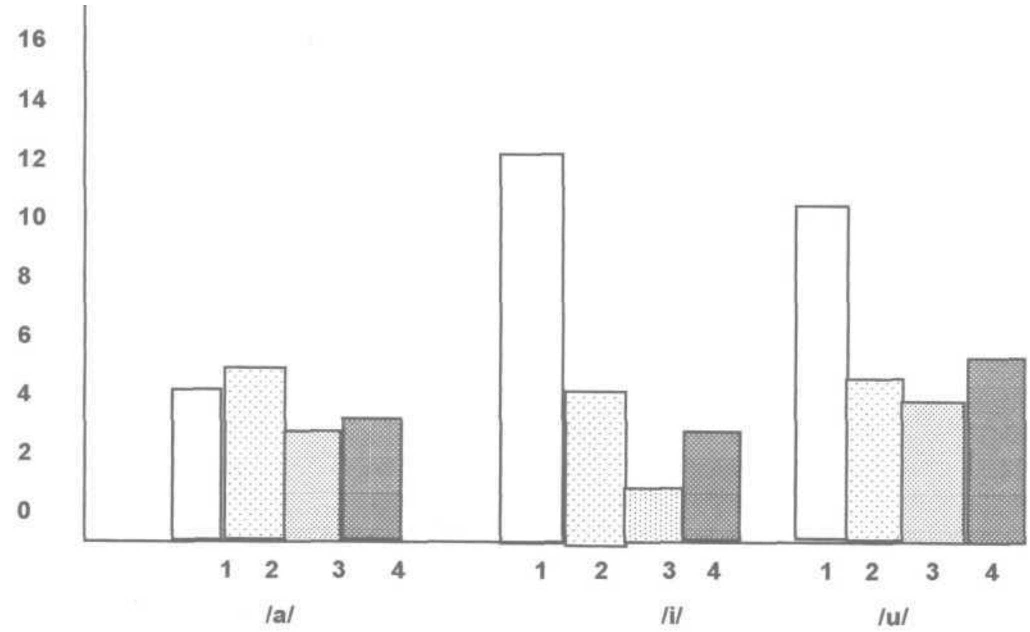
6. Extent of Fluctuation in Intensity

This has been defined as the average fluctuation in intensity (+/- 3 dB) or beyond) in phonation of one second.

| Vowels | Subjects | Mean | Range | SD |
|---|---|---|---|---|
| **/a/** | **N M** | 4.08 | 8.81 | 2.71 |
| | **NF** | 4.78 | 7.30 | 2.25 |
| | **DM** | 2.81 | 6.58 | 4.49 |
| | **DF** | 3.80 | 8.81 | 1.98 |
| **/i/** | **N M** | 12.07 | 17.05 | 13.43 |
| | **NF** | 3.88 | 9.46 | 2.77 |
| | **DM** | 1.5 | 6.52 | 2.01 |
| | **DF** | 3.04 | 8.83 | 2.90 |
| **/u/** | **N M** | 10.38 | 13.39 | 11.69 |
| | **NF** | 4.76 | 7.10 | 2.47 |
| | **DM** | 3.62 | 7.34 | 4.24 |
| | **DF** | 4.31 | 11.33 | 1.87 |

Table 53: Mean, range, SD values of the vowels/a//i/ and/u/ for "Extent of fluctuation in intensity" in case of both normal groups (males and females) and the dysphonic groups (males and females) as measured by Vaghmi software.

The males of the dysphonic group had a greater value of extent of fluctuation in intensity compared to females for the vowels /a/ /i/ and /u/. The females in the normal group obtained a lower extent of fluctuation in intensity compared to their male counterparts.




**1= Normal male; 2- Normal female 3- Dysphonic male; 4-Dysphonic female**

Graph 24: Mean of extent of fluctuation in intensity in phonation in both normal group (males and females) and dysphonic groups (males and females) as evaluated by Vaghmi software.

Some of the studies which had obtained high extent of fluctuation in intensity values were :

| Investigators | Male | Female |
|---|---|---|
| Nataraja (1986) | 2.45 | 1.59 |
| Rajashekar (1991) | 1.80 | - |
| Suresh (1991) | 2.39 | 1.32 |
| Krishnan (1992) | 4.81 | 5.58 |
| Pradeep(1997) | 2.13 | 1.71 |
| Rajkumar (1998) | 1.16 | 1.21 |

Table 54: Shows the values of extent of fluctuation in intensity obtained by various investigators.

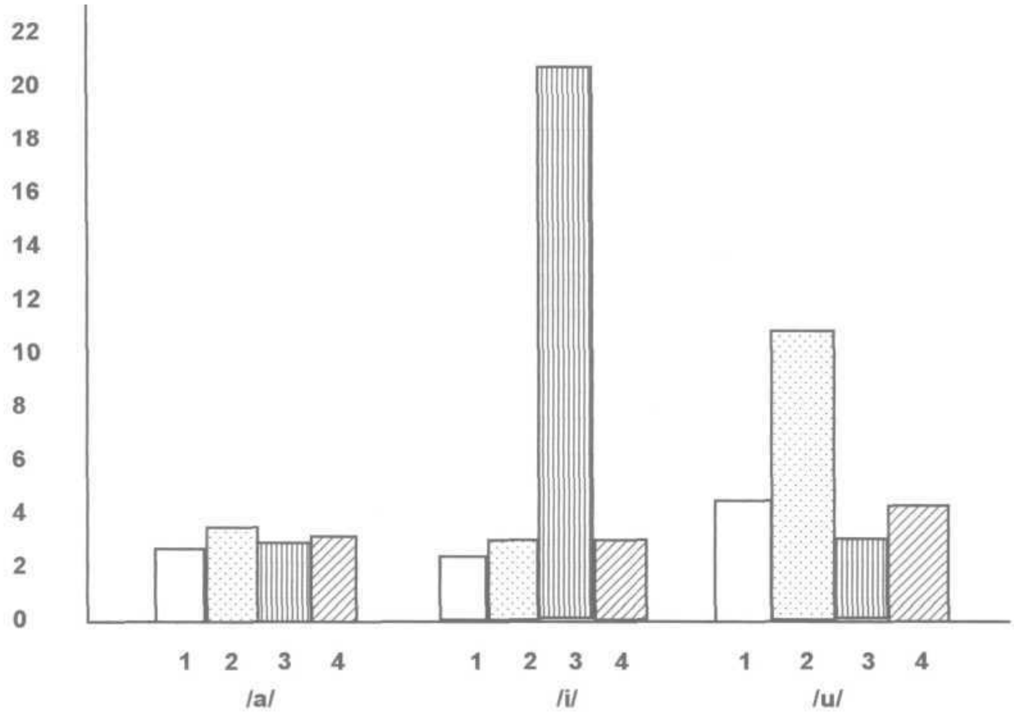
The results of this study approximated that of Tharmar (1991) who found extent of fluctuation in intensity values for males and females as 0.26 and 0.46 and of Prabha (1 997) as 0.42 and 0.54. Shoba (1 986) also found the extent of fluctuation in intensity in professional voice users as 0.66.

7. Speed of Fluctuation in Intensity

Defined as the number of fluctuations in intensity (+/- 3 dB) or beyond) in phonation of 1 second. Table 55 shows the mean, range and SD values for/a/ /i/ and /u/ in both normal groups and dysphonic groups.

| Vowels | Subjects | Mean | Range | SD |
|---|---|---|---|---|
| **/a/** | **NM** | 2.52 | 4.65 | 1.19 |
| | **NF** | 3.80 | 7.30 | 1.98 |
| | **DM** | 2.76 | 4.17 | 1.51 |
| | **DF** | 3.46 | 1 1.84 | 3.68 |
| **/i/** | **NM** | 2.26 | 3.92 | 1.41 |
| | **NF** | 2.92 | 12.36 | 3.46 |
| | **DM** | 21.52 | 8.83 | 2.90 |
| | **DF** | 3.04 | 36.52 | 14.01 |
| **/u/** | **NM** | 4.31 | 7.10 | 1 .67 |
| | **NF** | 10.38 | 33.39 | 1 1.30 |
| | **DM** | 2.90 | 3.00 | 0.87 |
| | **DF** | 3.96 | 13.73 | 3.85 |

Table 55: Mean, range, SD values of/a//i/ and /u/ of speed of fluctuation in intensity for both normal groups (males and females) and dysphonic group as measured by Vaghmi software.




**1= Normal male; 2- Normal female 3- Dysphonic male; 4-Dysphonic female**

Graph 25 : Mean of speed of fluctuation in intensity in phonation in both normal group (males and females) and dysphonic groups (males and females) as evaluated by Vaghmi software.

The males of the normal group had lower values for all vowels compared to males of the dysphonic group. This could be attributed to the limited ability on the control of the vocal mechanism in case of dysphonics, who exhibited higher values for all vowels.The variability was more in the dysphonic group (both males and females) compared to the normal group (males and females) for all the vowels.

The values of speed of fluctuation in intensity obtained in the present study were in agreement with the following reports.

| Investigators | Male | Female |
|---|---|---|
| Nataraja (1986) | 1.40 | 1.00 |
| Rajashekar (1991) | 0.40 | - |
| Tharmar(1991) | 1 .06 | 2.44 |
| Suresh (1991) | 1.43 | 0.45 |
| Krishnan(1992) | 2.23 | 0.30 |
| Prabha (1997) | 0.22 | 0.88 |
| Pradeep(1997) | 2.43 | 1.20 |
| Rajkumar(1998) | 0.80 | 1.62 |
| David (1998) | 0.03 | 0.05 |
| Present study | 2.52 | 2.76 |

Table 56: Values of speed of fluctuation in intensity (in flue/sec) for the phonation of/a/ in normal males and females as shown by various investigators.

The values of speed of fluctuation of intensity in dysphonic Indian population measured by Rajkumar (1998) was 3.45 and 2.56 for males and females respectively. This parameter was thus considered to be one of the useful parameters in differentiating normals and dysphonics.

The parameters evaluated using Dr.Speech software are :

1. Habitual Fundamental Frequency
2. Maximum Fo
3. Minimum Fo
4. Fundamental Frequency Tremor.
5. Amplitude Tremor

6. Standard Deviation of Fo

7. Jitter Percent

8. Shimmer Percent

9. Harmonics to Noise Ratio

10. Normalised Noise Energy

11. Signal to Noise Ratio

The results and discussion of the above parameters are presented as foilows:

Parameters 1 -4 have already been discussed.

5. *Amplitude tremor.* Defined as the average ratio of the amplitude of the most intense low frequency amplitude modulating component to the total amplitude of the analysed voice signal.

Table 57 indicates the mean, SD and range values of "Amplitude Tremor" for both normal and dysphonic groups for phonation of */a/, /i/* and /u/as evaluated by Vaghmi software.

| Vowels | Subjects | Mean | Range | SD |
|---|---|---|---|---|
| **/a/** | NM | 1.57 | 0.71 | 1-3.42 |
| | NF | 2.03 | 1.76 | 1-8.58 |
| | DM | 1.51 | 0.58 | 1-2.83 |
| | DF | 2.35 | 2.44 | 1-10.01 |
| **/i/** | NM | 1.85 | 1.66 | 1-7.01 |
| | NF | 2.03 | 1.76 | 1-8.58 |
| | DM | 1.51 | 0.58 | 1-2.83 |
| | DF | 2.19 | 3.22 | 1-13.72 |
| **/u** | NM | 1.39 | 0.58 | 1-3.04 |
| | NF | 4.27 | 4.42 | 1.01-12.89 |
| | DM | 1.64 | 0.60 | 1-2.49 |
| | DF | 3.29 | 3.23 | 1.03-1 1.21 |

Table 57: The mean, SD and the range of amplitude tremor in phonation of vowels /a/, /i//u/ for both the groups as evaluated .by Dr.Speech software.

While comparing the mean values of amplitude tremor for the phonation of different vowels, it was found that mean amplitude tremor values were higher for dysphonics than normals. Also, the variation is higher in the dysphonic group (males and females) compared to normals. Comparision of mean values of percent jitter revealed that the dysphonic group (both males and females) had higher mean values than the normal group) for all the vowels. The variability was also higher in the dysphonic group compared to the normal group.

This results and discussion and parameter has already been presented.

6. *Standard deviation of Fo.*

7. *Jitter percent:*

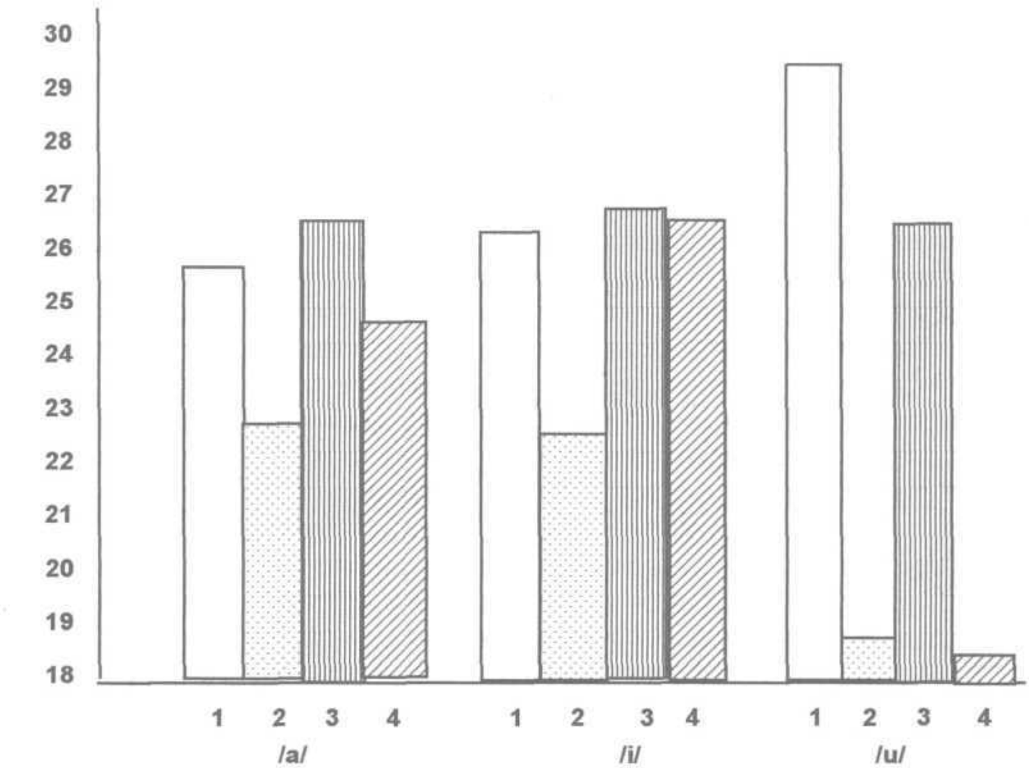
8. *Shimmer Percent*

9. Harmonics to Noise Ratio

Table 58 indicates the HNR for both the normal and dysphonic groups for the phonation of/a/, /i/ and/u/ vowels.

| Vowels | Subjects | Mean | Range | SD |
|---|---|---|---|---|
| /a/ | NM | 25.81 | 5.12 | 13.84-25.81 |
| | NF | 22.87 | 7.19 | 6.32-32 |
| | DM | 26.53 | 2.78 | 22.91-32.21 |
| | DF | 24.82 | 6.85 | 9.73-38.64 |
| **/i/** | NM | 26.48 | 6.95 | 3.27-32.04 |
| | NF | 22.58 | 6.77 | 7.07-34.9 |
| | DM | 26.98 | 3.86 | 20.39-32.47 |
| | DF | 26.50 | 5.97 | 15.49-34.36 |
| /u/ | NM | 29.48 | 2.32 | 25.64-33.83 |
| | NF | 18.93 | 6.63 | 8.42-30.99 |
| | DM | 26.42 | 6.41 | 13.46-33.48 |
| | DF | 18.289 | 5.83 | 9.03-28.05 |

Table 58: Shows the mean, SD, range for both normals and dysphonic group for phonation of different vowels.




**1= Normal male; 2- Normal female 3- Dysphonic male; 4-Dysphonic female**

Graph-26 : Mean of harmonic to noise ratio in phonation in both normal group (males and females) and dysphonic groups (males and females) as evaluated by Dr. Speech software.

Comparision of the mean values of HNR for normal and dysphonic males and females for phonation indicated that the dysphonic group showed lower HNR values than the normal group. The values of HNR obtained for this study correlated with those of earlier studies.

| Investigators | Male | Female |
|---|---|---|
| Rajashekar (1991) | 26.51 | 27.82 |
| Rajkumar (1998) | 24.92 | 27.33 |
| David (1998) | 25.97 | 27.89 |
| Present study | 25.81 | 26.53 |

Table 59: Shows the values of HNR for phonation of/a/ on normal Indian population as reported by various investigators.

In case of normal males, the HNR values of vowel /u/ was more than /a/ and /i/. The dysphonic male group had the lowest HNR value for vowel /u/ followed by /i/ and /a/. The normal females and dysphonic female group indicated the highest HNR value for vowel /i/ followed by /a/ and /u/.

10. Normalized Noise Energy

The values obtained by the two groups are given below:

| Vowels | Subjects | Mean | Range | SD |
|---|---|---|---|---|
| **/a/** | **NM** | -9.87 | 3.42 | -15.65-(-4.58) |
| | **NF** | -3.87 | 3.90 | -13.67-(2.95) |
| | **DM** | -14.42 | 2.56 | -18.36-(-9.22) |
| | **DF** | -7.18 | 6.25 | -17.72-(-O.88) |
| **/i/** | **NM** | -10.41 | 3.78 | -17.16-(3.71) |
| | **NF** | -4.55 | 3.26 | -11.35-(-.19) |
| | **DM** | -10.02 | 3.81 | -16.89-(-3.89) |
| | **DF** | -4.99 | 4.32 | -14.82-(-2.56) |
| **/u/** | **NM** | -5.53 | 3.49 | -14.29-(-1.94) |
| | **NF** | -.43 | 3.92 | -1 1 .85-C5.14) |
| | **DM** | -2.48 | 2.06 | -7.03-(.45) |
| | **DF** | 1.09 | 3.36 | -4.32- (-7.08) |

Table 60: The mean, SD and range obtained for NNE in normals and dysphonics for different vowels /a//i/ and /u/ as measured by Dr.Speech software.

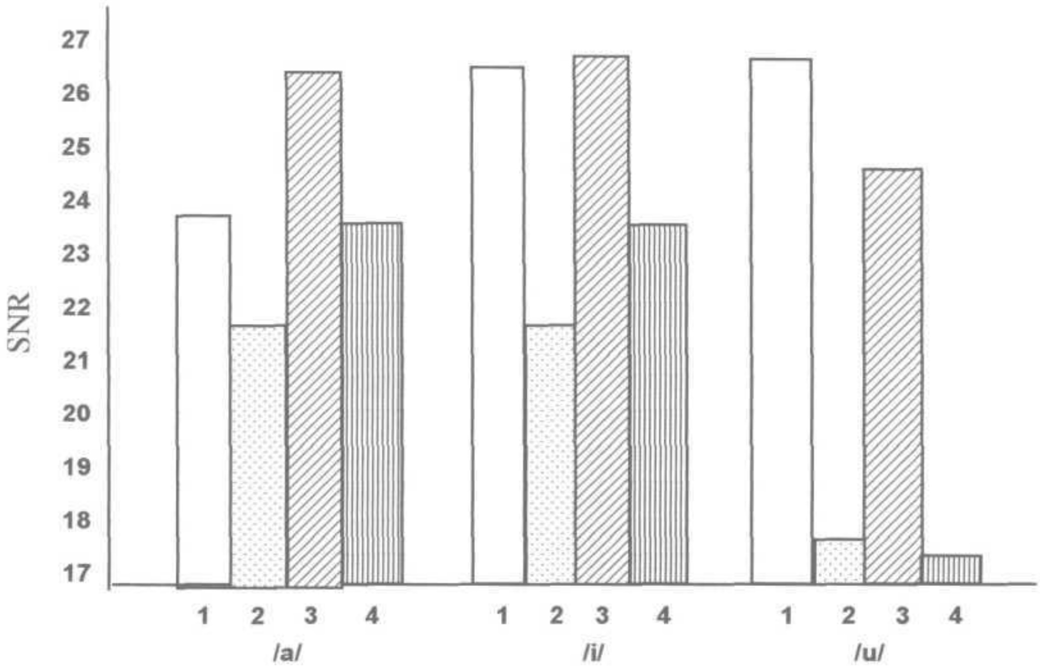
Examination of table 60 revealed that the dysphonic group Cboth males and females) exhibited higher mean values than the normal group Cboth males and females) for phonation of different vowels.

1 1 . Signal to Noise Ratio (SNR)

The results of the above parameter is presented in the table 61

| Vowels | Subjects | Mean | Range | SD |
|---|---|---|---|---|
| /a/ | NM | 23.83 | 6.18 | 6.26-30.52 |
| | NF | 21.84 | 6.96 | 6.37-31 .57 |
| | DM | 26.14 | 2.62 | 21 .75-29.84 |
| | DF | 23.69 | 2.64 | 9.49-37.13 |
| /i/ | NM | 26.46 | 2.99 | 21.25-31.30 |
| | NF | 21.84 | 6.96 | 6.37-31 .57 |
| | DM | 26.84 | 4.25 | 19.38-34.40 |
| | DF | 23.69 | 6.61 | 9.49-37.13 |
| /u/ | NM | 26.66 | 6.42 | 9.1 1-32.53 |
| | NF | 17.82 | 6.30 | 8.14-29.84 |
| | DM | 24.85 | 6.41 | 12.19-32.78 |
| | DF | 17.14 | 5.56 | 8.59-26.69 |

Table 61 : Mean, SD, range values of SNR for phonation of vowels in both normal and dysphonic group as measured by Dr.Speech software.




**1= Normal male; 2- Normal female 3- Dysphonic male; 4-Dysphonic female**

Graph 27 : Mean values of signaltonoise ratio in phonation in both normal group (males and females) and dysphonic groups (males and females) as evaluated by Dr. Speech software.

Comparision of mean values of SNR in both normal and dysphonic groups indicate that the values of the normal group (males and females) are than the dysphonic group (males and females) for all the vowels] The variability is higher in the dysphonic group compared to the normal group.

Thus the results presented till now show that the results are valid and reliable as theyare similar to the others which have been presented in the review of literature.

Results of Discriminant Analysis

The canonical discriminant analysis was conducted individually for all the three softwares i.e. MDVP, Dr.Speech and Vaghmi.

The data was processed using the stepwise discriminant analysis with statistical package for social sciences (7.5 Windows version)

Results of discriminant analysis for Dr.Speech software are presented in the table 62.

| Group | Predicted group members | | Total |
|---|---|---|---|
| | 1 | 2 | |
| 1 | 30 | 0 | 30 |
| 2 | 0 | 30 | 30 |
| % 1 | 100 | 0 | 100 |
| **2** | 0 | 100 | 100 |

100% of cases correctly classified,

Table 62 : Results of discriminant analysis of normal (group-1 ) and dysphonics (group 2) for the vowels /a/ */i/* and /u/

As can be seen from the table 62, in the normal group all the 30 cases have be classified as normal and in the dysphonic group all the 30 individuals as dysphonic.

100% of cases have been correctly classified.

The parameters which had been used for this classification are

1.i Habitual fundamental frequency of/a/.

2. Habitual frequency of/u/.

3 Absolute jitter of */a/.*

| Group | Predicted group Membership | | Total |
|---|---|---|---|
| | 1 | 2 | |
| 1.00 | 30 | 0 | 30 |
| 2.00 | 5 | 25 | 30 |
| % 1.00 | 100 | 0 | 100 |
| 2.00 | 16.7 | 8.33 | 100 |

91 .796 of original cases correctly classified.

Table 63: Results of discriminant analysis of normal (group 1) and dysphonics (group 2) for the vowels /a//i/ and /u/ for MDVP software.

In the discriminant analysis not all the parameters were considered for analysis. The parameters which were considered for analysis were :

1. Number of voice breaks of /u/
2. Voice turbulence index of/i/.
3. Degree of voice breaks of /u/
4. Noise to harmonic ratio of/u/.
-5. Standard deviation of frequency

As can be seen from the table 63 all the individuals in the normal group have been classified as normal but in the dysphonic group, 25 individuals have been classified as dysphonics and 5 in the normal group 91 .7% of the cases have been correctly classified. 91 .7% of the cases have been correctly classified.

Discriminant analysis for Vaghmi software

| | Group | Predicted group Membership | | Total |
|---|---|---|---|---|
| | | 1 | 2 | |
| | 1 | 26 | 4 | 30 |
| | 2 | 3 | 27 | 30 |
| % | 1 | 86.7 | 13.3 | 100 |
| | 2 | 10.0 | 90.0 | 100 |

(88.3%) of cases correctly classified.

Table 64: Results of discriminant analysis of normals (groupi ) and dysphonics (group 2) for vowels /a//i//u/ combined for Vaghmi software.

In discriminant analysis, not all the parameters were considered for analysis. The parameters which were used for analysis are :

1. Minimum fundamental frequency of /a/
2. Range of fundamental frequency of/u/

From the table 64 it can be noted that out of 30 normal subjects, 26 subjects were correctly classified as belonging to the normal group while 4 subjects were classified as belong to the dysphonic group. 86.7% of cases were correctly classified as belonging to the normal group.

Out of 30 dysphonic subjects, 7 cases were correctly classified as belonging to dysphonic group while 3 subjects were classified as belong to the normal group. 90% of the cases were correctly classified as belonging to the dysphonic group while 1 096 of dysphonics were wrongly classified as belonging to normal group.

On the whole 88.396 of cases were correctly classified are using VAGHMI software. Thus, this software was able to classify the dysphonics from normals using just two parameters. Though the software has facilities to measure other parameters like harmonic to noise ratio, jitter, shimmer and related measures, it was not possible to include them in the present investigation due to time limitations. Further it may be that parameters measured using this software may be using much strict criteria than that of other parameters measured and used.

To summarise, 100% of the cases have been correctly classified using Dr.Speech software based on three parameters, 91 .7% using MDVP based on five parameters and 88.396 using VAGHMI software based on two parameters. In terms of parameters used, Vaghmi has been economical but has yielded only 88.3%, whereas Dr.Speech has yielded 100% classification using three parameters. However, regarding the use of software one should keep in mind the user friendliness, cost, time used and hardware involved and other factors.

Thus all the three softwares have been found to be effective in distinguishing normals from dysphonic group and a combination of them can be effectively used in distinguishing normal from dysphonic group.

Thus the hypothesis stating that there is no difference among the three softwares MDVP, Vaghmi and Dr.Speech in terms of their efficacy of evaluation of voice disorders has been accepted and the purpose of the study to evaluate the efficacy of the three softwares has been achieved.

# SUMMARY AND CONCLUSION

Acoustic analysis of voice is presently gaining more importance. Hirano (1981) states that "____this may be one of the most attractive method of assessing the phonatory function or laryngeal pathology because it is non-invasive and provides objective and quantitative data". Acoustic analysis can be done by using methods such as spectrography, peak picking, inverse filtering, computer based methods and others.

In the present study, three software programs namely: (i) Multidimentional voice program (MDVP) (ii) Vaghmi software (iii) Dr.Speech software were used to acquire, analyse and display the following 48 parameters. Both vocalisations (of /a/ /i/ and /u/ and speech (Kannada sentences /idu papu/ /idu kot/ /idu kempu banna) were used as the samples for analysis. The extracted parameters was subjected to statistical analysis.

The parameters extracted were :

1. Average Fundamental frequency
2. Average Pitch Period
3. Highest Fundamental Frequency
4. Lowest Fundamental Frequency
5. Standard Deviation of Fundamental Frequency
6. Phonatory Fundamental frequency Range in Semitones
7. Fundamental Frequency Tremor Frequency
8. Amplitude tremor frequency

9. Absolute Jitter
10. Jitter Percent
11. Relative Average Perturbation
12. Pitch Period Perturbation Quotient
13. Smoothed Pitch Period Quotient
14. Fundamental Frequency Variation
15. Shimmer in dB
16. Shimmer Percent
17. Amplitude Perturbation Quotient
18. Smoothed Amplitude Perturbation Quotient
19. Peak Amplitude Variation
20 Frequency Tremor Intensity Index
21. Amplitude Tremor Intensity Index
22. Degree of Voiceless
23 Noise to Harmonic Ratio
24. Voice Turbulence Index
25. Soft Phonation Index
26. Degree of Voice Breaks
27. Number of Voice Breaks
28. Number of Subharmonic Segments
29. Degree of Subharmonic Components
30. Fundamental Frequency in Phonation
31. Fundamental Frequency in Speech
32. Maximum Fundamental Frequency in Speech
33. Minimum Fundamental Frequency in Speech
34. Extent of Fluctuation in Fundamental Frequency
35. Speed of Fluctuation in Fundamental Frequency

36. Extent of Fluctuation in Intensity
37. Speed of Fluctuation in Intensity
38. Maximum Fundamental Frequency
39. Minimum Fundamental Frequency
40. Habitual Fundamental Frequency in Phonation
41. Fundamental Frequency Tremor
42. Amplitude Tremor Frequency
43. Standard Deviation of Fundamental Frequency
44. Jitter %
45. Shimmer 96
46. Harmonics to Noise Ratio
47. Normalised Noise Energy
48. Signal to Noise Ratio.

Parameters 1-29 were evaluated using MDVP software

Parameters 30-37 using Vaghmi software

Parameters 38-48 using Dr.Speech software.

***Dr. Speech software***

All the 48 parameters were measured in *a* group of 30 normals (1 5 males and 1 5 females) and a group of 30 dysphonics (15 males and 1 5 females).

The results were subjected to statistical analysis (Descriptive statistics and discriminant analysis) using SPSS computer program.

**Discriminant analysis revealed the following :**

From the date of parameters extracted using MDVP software, it was seen that 91 .796 of the cases were correctly classified. The parameters which were used for this analysis were :

1. Number of voice breaks of /u/
2. *Vo\ce* turbulance index of/i/
3. Degree of voice breaks of /u/
4. Noise to harmonic ratio of /u/
5. Standard deviation of frequency

Based on data collected using Vaghmi software, it was found that 88.396 of the cases were correctly classified and for this purpose the following parameters were used for the analysis :

.

1. Minimum fundamental frequency of/a/
2. Range of fundamental frequency of/u/

1 0096 of the cases were correctly differentiated from normals based on data collected using Dr.Speech software. The following parameters were used for analysis.

1. Habitual fundamental frequency of/a/
2. Habitual frequency of/u/
3. Absolute jitter of /a/-

Thus the results of the study show that all the three softwares are capable of distinguishing the dysphonics from normals to a high degree (88.3 to 1 00 96). However, the number of parameters used by each of them varies. The selection of the software may also depend on other factors like, hardware involved, cost, user friendliness of the software, service facilities, training to personnel and others. However, it can be concluded that in terms of their efficacy all the three are of equal status. The study may be repeated with more number of subjects. Further studies can be considered to determine their efficiency in differential diagnosis of dysphonias, the user friendliness and other factors which will be useful in the selection of the software.

# BIBLIOGRAPHY


Anders, L.C., Hollien, H.; Hurme, P., Sonnien, A. and Wendle, J. (1988). Perception of hoarseness by several classes of listeners. Folia Phoniatrica, 40(2), 91-100.

Anderson, V. (1 961). Training the speaking voice. New York: Oxford University.

Anitha, V. (1994). Multidimensional analysis of voice disorders. Unpublished Master's Dissertation, University of Mysore.

Aparna, M. (1996). Multidimensional analysis of voice in hearing-impaired. Unpublished Master's Dissertation, University of Mysore.

Arnold, G.E. (1961). Physiology and pathology of the cricothyroid muscle. Laryngoscope, 71, 687-753.

Aronson,A.E. (1 980). Clinical voice disorders - An interdisciplinary approach. Brain.G Decker, New York.

Arun, B. (1 995). Multidimensional voice analysis in children. Unpublished Master's Dissertation, University of Mysore.

Askenfelt, A.G. and Hammarberg, B. (1986). Speech waveform perturbation analysis : A perceptual acoustical comparision of seven measures. Journal of Speech and Hearing Research, 29(1), 50-64.

Asthana, P. (1977). Relationship between vocal intensity, pitch and nasality in cleft palate speakers. Unpublished Master's Dissertation, University of Mysore.

Biswajith Das (1 995). Multidimensional analysis of voice disorders. Unpublished Master's Dissertation, University of Mysore.

Bohme, G. and Hecker, G. (1970). Gerontologische untersuchungen uber stimmumfang and speechstimmlage. Folia Phoniatrica, 22, 176-1 84.

Boone, D.E. (1971). The voice and voice therapy (1st Ed.). New Jersey : Prentice Hall Inc.

Boone, D.E. (1 977). The voice and voice therapy (2nd Edn). New Jersey : Prentice Hall Inc.

Borden, G.J. and Harris, K.S. (Ed.) (1 980). Speech science primer : Physiology, acoustics and perception of speech. Baltimore : The Williams and Wilkins Company.

Broadnitz, F.S. (1959). Vocal rehabilitation. Rochester : Minn: Whiting Press.

Carr, P.B. and Trill, D. (1964). Long-term excitation spectra. Journal of Acoustical Society of America, 36, 2033- 2040. Cited in Baken, RJ. (1 989). Clinical measurement of speech and voice. London : Taylor and Francis Ltd.

Chandrashekar, K.R. (1987). Electroglottography in Dysphonics. Unpublished Master's Dissertation, University of Mysore.

Coleman, R.F. (1960). Some acoustic correlates of hoarseness. Unpublished master's thesis. Vanderbilt University, Nashiville, T.N. Cited in Childers, D.G. and Lee, C.K. (1 991 ). Vocal quality factors : Analysis, synthesis and perception. Journal of the Acoustical Society of America, 90(5), 2394-241 0.

Cooper, M. (1974). Spectrographic analysis of fundamental frequency and hoarseness before and after vocal rehabilitation. Journal of Speech and Hearing Disorders, 39, 289-296.

Cowan, J. (1936). Pitch and intensity characteristics of stage speech. Arch.Speech.Sub^, 7-85.

Cox, N.B., Ito, M.R. and Morrison, M.D. (1 989). Data labeling and sampling effects in harmonic to noise ratios. Journal of the Acoustical Society of America, 85(5), 2165-2178.

Curry, E.T. (1 940). The pitch characteristics of the adolescent male voice. Speech Monograph, 7, 48-52.

tis, J.F. (1978). Processes and disorders of communication. New York : Harper and Row Publication.

David, T. (1998). Acoustic correlates of hoarse voice. Unpublished Master's Dissertation, University of Mysore.

Davis, S.B. (1 981 ). Acoustic characteristics of laryngeal pathology in Darby, J. (Ed.). Speech evaluation in medicine. Grune and Stratton,77-1 04.

Deem, J.F., Manning, W.H., Knack, J.V. and Matesich, J.S. (1989). The automatic extraction of pitch perturbation using microcomputers : Some methodological consideration. Journal of Speech and Hearing Research, 32(3), 689-697.

Dejonckere, P.H. and Labecq, J. (1 985). Electroglottography and vocal nodule : An attempt to quantify the shape of the signal. Folia Phoniatrica, 37, 195-200.

Deliyski, Orlikoff and Kaharie. (1 991 ). Cited in Rashmi, M. Acoustic aspects of the speech of children. Unpublished Master's Dissertation, University of Mysore,

Eguchi, S. and Hirsh, I.J. (1969). Development of speech sounds in children. Acta Otolaryngology (Suppl), 257-262.

Emanuel, F.W., Lively, M.A. and McCoy, J.F. (1973). Spectral noise levels and roughness ratings for vowels produced by males and females. Folia Phoniatrica, 25, 1 10-120. Cited in Emanuel, F.W. (1991). Spectral noise. Seminar in Speech and Language, 1 2(2), 1 1 5-1 30.

Erickson, C.I. (1959). The basic factors in the human voice. Psychological Monographs. University of IOWA Studies in Psychology, 10, 86-1 1 2.

Fairbanks, G. and Pronovost, W. (1939). An experimental study of pitch characteristics of voice during the expression of emotions. Speech Monograph, 6, 87-104.

Fairbanks, G. (1 940). Recent experimental investigation of voice pitch in speech. Journal of the Acoustical Society of America, 1 1 , 457-466.

Fairbanks, G., Wiley, V.H. and Bassman, F.M. (1 949). An acoustical study of vocal pitch in 7 and 8 year old boys. Child Development, 20, 63-70.

Fairbanks, G. (1 960). Voice and articulation drillbook. New York : Harper and Bros.

Fourcin, A. (1 981 ). Laryngographic assessment of phonatory function. ASH A Reports, 11,11 6-1 27. Cited in Haji, T., Moriguchi, S., Baer, T. and Gould, W. J. (1 986). Frequency and amplitude perturbation analysis of electroglottograph during sustained Phonation. Journal of the Acoustical Society of America, 80(1 ), 58-62.

Freeman (1 982). Cited in Rashmi, M^ Acoustic aspects of the speech of children. Unpublished Master's Dissertation, University of Mysore.

Froeschels, E. (1940). Laws in the appearance and the development of voice hyperfunctions. Journal of Speech Disorders, 5, 1-4.

Frokjaer-Jenson, B. and Prytz, S. (1975). Registration of voice quality. Bruel and Kjaer, Technical Review, 3, 3-17.

Gauffin, J., Hammarberg, B. and Hartegard, S. (1 996). Irregularities in the voice : A perceptual experiment using synthetic voices with subharmonics. Cited in Davis, P.J. and Hetcher, N.H. (Eds.). Vocal Fold Physiology. London : Singular Publishing Group.

George, S. (1973). A study of fundamental frequency of voice and natural frequency of vocal tract on an Indian population of different age ranges. Unpublished Master's Dissertation, University of Mysore.

Gill, J.S. (1961). Automatic extraction of the excitation function of speech with particular reference to the use of correlation methods - proceedinss of the third international congress on acoustics. Amsterdam : Elsevier, 1,21 7-220. Cited in Horii, y. (1979). Fundamental perturbation observed in sustained phonation. Journal of Speech and Hearing Research, 22(1), 5-19.

Gopal, H.S. (1980). Relationship for locating optimum frequency in the age range of 7 to 28 years. Unpublished Master's Dissertation, University of Mysore.

Gopal, N.K. (1986). Acoustic analysis of the speech in normal adults. Unpublished Master's Dissertation, University of Mysore.

Greene, M.C.L. (1 964). The voice and its disorders. Mitma : Medical, London.

Hammerberg, B. (1980). Perceptual and acoustic correlates of abnormal voice qualities. Acta Otolaryngol, 90, 441 -451 .

Hanky (1 949). Cited in Rashmi, M. (1 985). Acoustic aspects of the speech of children. Unpublished Master's Dissertation, University of Mysore.

Hanson, D.G., Gerrat, B.R. and Ward, H.P. (1 983). Glottographic measurement of vocal dysfunction - A preliminary report. Annals of Otorhinolaryngology, 92(5), 413-419.

. and CramOn, D.V. (1 984). Acoustic measurement of voice quality in central dysphonia. Journal of Communication Disorders, 17(6), 425-440.

Hecker, M.H. and Kruel, E.J. (1971). Descriptions of the speech of patients with cancer of the vocal folds. Part I: Measures of fundamental frequency. Journal of the Acoustical Society of America, 49, 1 274-1 282.

Heiberger, V.L. and Horii, Y. (1 982). Jitter and Shimmer in sustained phonation. Lass, N.J. (Ed). Speech and language advances in basic research and practise. New York : Academic Press.

Higgins, B.M. and Saxman, H.J. (1989). A comparision of intrasubject variation across session of three vowel frequency perturbation indices. Journal of Acoustical Society of America, 86(3), 91 1 -91 6.

Hirano, M. (1975). Phonosurgery - Basic and clinical investigations. Otologia (Fukuoka), 21, 239-240.

Hirano, M. (1981). The function of laryngeal muscles in regulating fundamental frequency and intensity of phonation. Journal of Speech and Hearing Research, 12, 616-628.

Hirano (1989). Voice evaluationjclinical aspects. Cited in Kacker, S.K. and Basavaraj,V. (Ed.). Indian Speech, Language and Hearing Tests - The ISHA Battery, 176-195.

Hiraoka, N., Kitazoc, Y, Ueta, H., Tanaka, S. and Tanab, M. (1 984). Harmonic intensity analysis of normal and hoarse voices. Journal of Acoustical Society of America, 76(6), 1648-1651.

Hollien, H. (1 961 ). Vocal fold thickness and fundamental frequency ofphonation. Journal of Speech and Hearing Research, 5, 237-243.

Horii, Y. (1979). Fundamental frequency perturbation observed in sustained phonation. Journal of Speech and Hearing Research, 22, 5-19.

Horii, Y. (1 980). Vocal shimmer in sustained phonation. Journal of Speech and Hearing Research, 23(1), 202-209.

Hudson, A.I. and Halbrook, A. (1981). A study of the reading fundamental vocal frequency of young blacLiedults. Journal of Speech and Hearing Research, 24(2), 197-201.

Imaizumi, S., Hiki, S., Hirano, M. and Mutsushita, H. (1980). Analysis of pathological voices with a sound spectrograph (in Japanese). Journal of Acoustical Society of Japan, 36, 9-16. Cited in Kasuya, H., Ogawa, S., Mashima, K. and Ebihara, S. (1 980). Normalized noise energy as an acoustic measure to evaluate pathologic voice. Journal of the Acoustical Society of America, 80(5), 1 329-1334.

Indira, N. (1982). Analysis of infant cries. Unpublished Master's Dissertation, University of Mysore.

Iwata, S. and VbnLeden, H. (1 970). Phonation quotient in patients with laryngeal diseases. Phoniatrica, 22, 117-128.

Iwata, S. (1 972). Periodicities of pitch perturbations in normal and pathological larynges. Laryngoscope, 82, 87-96.

Jayaram, K. (1975). An attempt at differential diagnosis of dysphonia. Unpublished Masters Dissertation, University of Mysore.

Jensen, P.J. (1965). Adequacy of terming. for clinical judgement of voice quality deviation. Eye, Ear, Nose and Throat monthly, 44, 77-82. Cited in Aronson, A.E. (1 985). Clinical voice disorders - An interdisciplinary approach. Second edition, New York : Thieme Inc.

Kasuya, H., Masubuchi, K., Ebihara, S. and Toshiba, L. (1986). Preliminary experiments on voice screening. Journal of Phonetics, 1 4, 463-468.

Kasuya, H., Ogawa, S., Mashima, K. and Soihara, S. (1986). normalized noise energy as an acoustic measure to evaluate pathologic voice. Journal of the Acoustical Society of America, 80(5), 1329-1334.

Kent, R.D. (1976). Anatomical and Neuromuscular maturation of speech mechanism, evidence from acoustic studies. Journal of Speech and Hearing Research, 19, 412-445.

Kim, K.M., Kakita, Y. and Hirano, M. (1 982). Sound spectrographic analysis of the voice of patients with recurrent laryngeal nerve paralysis. Folia Phoniatrica, 34, 124-133.

Kitajima, K. and Gould, W.J. (1 976). Vocal shimmer in sustained phonation of normal and pathologic voice. Annals of Otology, Rhinology and Laryngology, 85, 377-381.

Klatt, D.H. and Klatt, L.C. (1 990). Analysis, synthesis and perception of voice quality variations among female and male talkers. Journal of the Acoustical Society of America, 87, 820- 857.

Klingholtz, F. (1990). Acoustic recognition of voice disorders : A comparative study of running speech versus sustained vowels. Journal of the Acoustical Society of America, 87(5), 2218-2224.

Koike, Y. (1 969). \£>wel amplitude modulations in patients with laryngeal diseases. Journal of the Acoustical Society of America, 45, 839-844.

Koike, Y. (1 973). Application of some acoustic measures for the evaluation of laryngeal dysfunction. Studia Phonologica, 7, 17-23.

Koike, y. (1 986). Cepstrum analysis of pathologic voices. Journal of Phonetics, 14, 501-508.

Krieman, J., Gerratt, B.R., Precoda, K. and Berke, G.S. (1992). Individual differences in voice quality perception. Journal of Speech and Hearing Research, 35(3), 512-520.

Krieman, J., Gerratt, B.R., Kempster, G.B., Erman, A. and Berke, G.S. (1 993). Perceptual evaluation of voice quality ; Review, tutorial and framework for future research Journal of Speech and Hearing Research, 36(1 ), 21 -40.

Krishnan, B.T. and Vareed, C. (1992). The vital capacity of 103 medical students in South India. Indian Journal of Medical Research,! 9, 1 1 65-1 1 83.

Kushal Raj, P. (1983). Acoustic analysis of speech of children. Unpublished Master's Dissertation, University of Mysore.

Lass, N.J., Brong, G.W., Ciccolella, S.A., Walters, S.C. and Maxwell, F.I. (1980). An investigation of speaker height and weight discriminations by means of paired comparision judgements. Journal of Phonetics, 8.

Lieberman, P. (1961). Perturbation in vocal pitch. Journal of Acoustical Society of America, 33, 597-603. Cited in Askenfelt, A.G. and Hammarberg, B. (1986). Speech waveform perturbation analysis : A perceptual acoustical comparision of seven measures. Journal of Speech and Hearing Research, 29(1), 50-64.

Lieberman, P. (1963). Some measures of the fundamental periodicity of normal and pathological larynges. Journal of the Acoustical Society of America, 45, 1537-1543.

Ludlow, C.L. (1981). Research needs for the assessment of phonatory function. ASH A Reports, 11,3-8.

Michel, J.F. and Wendahl, R. (1971). Correlates of voice production, in Travis, L.E.(Ed.). Handbook of Speech Pathology and Audiology. New Jersey : Prentice Hall Inc, 465-480.

Moore, G.P. (1971). Organic voice disorders. Englewood Cliffs, New Jersey : Prentice Hall Inc.

Moore, G.P. (1 97 5). Observation on the physiology of hoarseness - proceedings of the fourth international congress of phonetic science. Helsinki : Fin, 92-95. Cited in Childers, D.G. and Lee, C.K. (1 991 ). Vocal quality factors : Analysis, synthesis and perception. Journal of the Acoustical Society of America, 90(5), 2394-2410.

Moore, G.P. and Thompson,C.L. (1965). Comments on physiology of hoarseness. Archives of Otolaryngology, 81 , 97-1 02.

Morris, H. and Spriestersbach, H. (1 978). Appraisal of respiration and phonation. In Darley, F. and Spriestersbach, D.C. (Eds.). Diagnostic methods in speech pathology, 200-21 2, New York : Harper and Row Publishers.

Murphy, A.J. (1 964). Functional voice disorders. New Jersey : Prentice Hall Inc.

Murray, T. (1978). Speaking fundamental frequency characteristics associated with voice pathologies. Journal of Speech and Hearing Disorders, 43(3), 374-379.

Mysak, E.D. (1959). Pitch and duration characteristics of older males. Journal of Speech and Hearing Research, 2, 46-54.

Mysak, E.D. (1 966). Phonatory and Resonatory problems, in Riebber, R.W. and Brubaker, R.S. (Ed.). Speech pathology. Amsterdam : North Holland Publishing Company.

Nataraja, N.P. (1 972). Objective method of locating optimum pitch. Unpublished Master's Dissertation, University of Mysore.

Nataraja, N.P. (1986). Differential diagnosis of dysphonias. Unpublished Doctoral thesis, University of Mysore.

Nataraja, N.P. and Jagadish, A. (1984). Vocal duration and fundamental frequency. Journal of All India Institute of Speech and Hearing, 1 5.

Nataraja, N.P. and Jayaram, M. (1 982). A new approach to the classification of voice disorders. Journal of All India Institute of Speech and Hearing, 8, 21 - 28.

Noll, A.M. (1964). Short term spectrum and cepstrum techniques for vocal pitch detection. Journal of the Acoustical Society of America, 36, 296-302. Cited in Baken, R.J. (1 987). Clinical measurement of speech and voice. London: Taylor and Francis Ltd.

Pandita, R. (1 988).Objective analysis of Efficient voice.Unpublished Doctoral thesis.University of Mysore.

Pannbacker, M. (1 984). Classification of voice disorders : A review of literature. Language, Speech and Hearing Services in Schools, 1 5, 169-1 74.

Perkins, W.H. (1971). Speech pathology - An applied behavioural science. St.Louis : The C.V.Mosby Co.

Plomp, R. (1967). Pitch of complex tones. Journal of the Acoustical Society of America, 41, 1526-1534.

Prabha, S. (1997). Analysis of dysphonic voice before and after therapy. Unpublished Master's Dissertation, University of Mysore.

Pradeep (1997). Effects of pitch change on vocal fatigue. Unpublished Master's Dissertation, University of Mysore.

Preethi, K. (1998). Factors in normal and abnormal voice. Unpublished Masters Dissertation, University of Mysore.

Qi, y. and Shipp, T. (1992). An adaptive method for tracking voicing irregularities. Journal of the Acoustical Society of America, 91 (6), 3471-77.

Rashmi, M. (1985). Acoustic aspects of the speech of children. Unpublished Master's Dissertation, University of Mysore.

Read, C, Buder, E.H. and Kent, R.D. (1990). Speech analysis systems : A survey. Journal of Speech and Hearing Research, 33, 363-374.

Samuel, G. (1973). A study of fundamental frequency of voice and natural frequency of vocal tract on an Indian population of different age ranges. Unpublished Masters Dissertation, University of Mysore.

Sanjay, K. (1991). Electroglottography in dysphonics. Unpublished Master's Dissertation, University of Mysore.

Schoentgen, J. (1 989). Jitter in sustained vowels and isolated sentences produced by dysphonic speaker. Speech Communication, 8(1 ), 61 -80.

Scripture, E.W. (1906). Researches in experimental phonetics : The study of speech curves. Washington, D.C. : Carngegie Institute, 41 . Cited in Horii, Y. (1979). Fundamental frequency perturbation observed in sustained phonation. Journal of Speech and Hearing Research, 22(1 ), 5-1 9.

Sederholm, E., McAllister, A., Sundberg, J. and Dallcvist, J. (1 992). Perceptual analysis of child hoarseness using continuous scales. Speech transmission laboratory quarterly progress and status report, 1 , 99-1 1 3.

Sheela, E.V. (1974). Comparative study of vocal parameters of trained and untrained singers. Unpublished Master's Dissertation, University of Mysore.

Shipp, T. and Huntington, D. (1 965). Some acoustic and perceptual factors in acute-laryngitic hoarseness. Journal of Speech and Hearing Disorders, 30, 3 50-359.

Skinner, E. (1 93 5). A calibrated recording and analysis of pitch, force and quality of vocal tones expressing happiness and sadness. Speech Monograph, II, 81-137.

Snidecor, T. (1 943). A comparative study of the pitch and duration characteristics of imprompt speaking and oral reading. Speech Monograph, 10, 50-56.

Sonninen, A. and Sonninen, E. (1976). Logopedic-phoniatric observation on eight year old school children in the town of Jyskyla, Nord, Tidskr, Logopedinog Foniatri, 2, 74-82. Cited in Sedrholm, E., McAllister, A., Sundberg, J. and Dallcvist, J. (1992). Perceptual analysis of child hoarsensss using continuous scales. Speech Transmission Laboratory Quarterly Progress and Status Report, 1, 99-113.

Sorenson, D. and Horii, Y. (1983). Frequency and amplitude perturbation in the voice of female speakers. Journal of Communication Disorders, 16, 57-61.

Sreedevi, H.S. (1987). Acoustic characteristics of optimum voice. Unpublished Master Dissertation, Univrsity of Mysore.

Suresh, T. (1991). Acoustic analysis of voice in geriatric population. Unpublished Masters Dissertation, University of Mysore.

Tharmar, S. (1991). Acoustic analysis of voice in children and adults. Unpublished Master's Dissertation, University of Mysore.

Titze, I.R., Horii, *Y.* and Scherer, R.C. (1987). Some technical considerations in voice perturbation measurements. Journal of Speech and Hearing Research, 30(2), 252-260.

Usha, A.A. (1978). A study of fundamental frequency in Indian population. Unpublished Master's Dissertation, University of Mysore.

Vanaja, C.S. (1986). Acoustic parameters of normal voice. Unpublished Master s Dissertation, University of Mysore.

Venkatesh, C.S. (1985). Cited in Rashmi, M. Acoustic aspects of the speech of children. Unpublished Master's Dissertation, University of Mysore.

Venkatesh, C.S., Satya, K. and Jenny, E.P. (1992). Normative data on pitch and amplitude measurements in normal and discriminant analysis of pitch and amplitude perturbation measurements in dysphonics. Paper presented at XXIV National Conference of Indian Speech and Hearing Association

Von Leden, H. and Koike, Y (1970). Detection of laryngeal disease by computer technique. Arch. Otol. 91, 3-10.

Wendahl, R.W. (1963). Laryngeal analog synthesis of harsh voice quality. Folia Phoniatrica, 15, 241-250. Cited in Baken, R.J. (1987). Clinical measurement of speech and voice. London : Taylor and Francis Ltd.

Wendahl, R. W. (1966). Laryngeal analogy synthesis of jitter and shimmer auditory parameters of harshness. Folia Phoniatrica, 18, 98-108. Cited in Baken, R.J. (1987). Clinical measurement of speech and voice. London: Taylor and Francis Ltd.

Wendler, J., Doherty, E.T. and Hollien, H. (1990). Voice classification by means of long-term speech spectra. Folia Phoniatrica, 32, 51-60.

West, R., Ansberry, M. and Carr, A. (1957). The rehabilitation of speech (III ed). New York : Harper and Row.

Wewers, M.E. and Lowe, N.K. (1990). A critical review of visual analogue scales in the measurement of clinical phenomena. Research in nursing and health, 13, 227-2367. Cited in Sederholm, E., McAllister, A., Sundberg, J. and Dalkvist. J. (1992). Perceptual analysis of Child hoarseness using continuous scales. Speech transmission laboratory quarterly progress and status report, 1, 99-113.

Wilcox, K. (1978). Age and vowel differences in vocal jitter. Unpublished Master thesis, Purdue University. Cited in Sorensen, D. and Horii, *Y.* (1984). Directional perturbation factors in jitter and shimmer. Journal of Communication Disorder, 17(3), 143-151.

Wilcox, K. and Horii, *Y.* (1980). Age and changes in vocal jitter. Journal of Gerentology, 35, 194-198.

Wilcox, K. and Horii, *Y.* (1980). Age and changes in vocal jitter. Journal of Gerontology, 35, 194-198. Cited in Higgins, M.B. and Saxman (1989). A comparision in intrasubject variation across sessions of three vocal frequency perturbation indices. Journal of the Acoustical Society of America, 86(3), 911-916.

Wilson, D.K. (1979). Voice problems in children. 2nd Edn. Baltimore : The Williams and Willcins Company.

Yanagihara, N. (1967). Significances of harmonic change and noise components in hoarseness. Journal of Speech and Hearing Research, 10, 531-541.

Vbon, M.K., Kakita, Y. and Hirano, M. (1 984). Sound spectrographic analysis of the voice of patients with glottic carcinomas. Folia Phoniatrica, 36, 24-30.

Yumoto, E., Sasaki, Y, Okamura, H. (1984). Harmonics to noise ratio and psychological measurements of the degree of hoarseness. Journal of Speech, Hearing and Mental Research, 27, 92-96.

# APPENDIX

The definitions of the parameters considered in this study are as follows

*Average fundamental frequency (Fo) / Hz/*

Average value of all extracted period-to-period fundamental frequency values. Voice break areas are excluded.

Fo is computed from the extracted period-to-period pitch data as :

*/Absolute jitter/sec/or/jita:*

$$Fo = \frac{1}{N-1} \sum_{i=1}^{N-1} Fo^{(i)}$$

Where

$$Fo^{(i)} = \frac{(1)}{To^{(i)}} \quad \text{Period-to-period fundamental frequency}$$

To, i = 1, 2, ... N extracted pitch period data

N = PER, Number of extracted pitch periods.

***Highest Fundamental Frequency (HFo/Hz)***

The greatest of all extracted period-to-period fundamental frequency values. Voice break areas are excluded. It is computed as

$$Fhi = Max. [Fo^{(I)}), \; i = 1, 2, ..., N$$

*Lowest Fundamental Frequency (LFo/Hz)*

The lowest of all extracted period-to-period. It is computed as :

$$Flo = Min \; (Fo^{(I)}), \; i = 1,2, ...., N$$

The lowest fundamental within the defined period is extracted and displayed as Flo. However, the pitch extracted range is defined to either search for periods from 70-625 Hz or 200-1 000 Hz. Therefore, the high' range will not determine a fundamental under 200 Hz.

*Standard Deviation of Fundamental Frequency (SFD)/Hz)*

Standard deviation of all extracted period-to-period fundamental frequency values. Voice break areas are excluded.

$$STD = \sqrt{\frac{1}{N} \sum_{i=1}^{N-1} (Fo - Fo^{(i)})^{(2)}}$$

where,

$$Fo = \frac{1}{N} \sum_{i=1}^{N-1} Fo^{(i)}$$

$$Fo^{(i)} = \frac{1}{To^{(i)}}$$ - Period to period of values

$To^{(i)}$, $i = 1, 2, ...., N$ extracted pitch period data

$N$ = Number of extracted pitch periods.

*Phonatory Fundamental Frequency range (PFR)/Semitones)*

The range between Fhi and Flo expressed in number of semitones. The ratio of two consecutive semi-tones is equal to 1 2th root of 2.

First all frequencies of semitones $Fst^{(k)}$ - F1 , k= 1, 2, ... are computed within the frequency range 55 Hz to 1055 Hz.

where $a = \sqrt[12]{2}$

$$f1 = 55\ Hz,\ f2 = 1055\ Hz \text{ and } f1 < Fst^{(k)} \leq f2$$

*Fo - Fremor Frequency (FFFR)/Hz)*

The frequency of the most intensive low frequency Fo- modulating component in the specified Fo-tremor analysis range. If the corresponding FTRI value is below the specified threshold, the Fftr value is zero.

The method for frequency tremor analysis consists of the following :

A. Division of the fundamental frequency period-to-period (Fo) data **into** 2 sec windows at 1 sec step between. For every window, the following procedures apply.

1 . Low-pass filtering of the Fo data at 30 Hz and down sampling at 400 Hz.
2. Calculation of the total energy of the resulting signal.
3. Subtraction of the DC component.
4. Calculation of an auto correlation function on the residue signal.
5. Division by the total energy and conversion to (%)

6. Extraction to the period of variation.

7. Calculation of Fftr corresponding to the period of variation found.

B. Computation of the average autocorrelator curve and average Fftr for all processed window.

Amplitude tremor frequency (FATR

The frequency of the most intensive low-frequency amplitude modulating component in the specified amplitude tremor analysis range. If the corresponding ATRI value is below the specified threshold, the Fatr value is zero.

The method for amplitude tremor analysis consists of the following :

A. Division of the peak-to-peak amplitude data at 30 Hz. and down sampling to 400 Hz.

1 . Calculation of the total energy of the resulting signal.

2. Subtraction of the DC component.

3. Calculation of an autocorrelation function of the residuence signal.

4. Division by the total energy and conversion to percentage.

5. Extraction of the period of variation.

6. Calculation of Fatr corresponding to the period of variation found.

B. Computation of the average autocorrelation curve and average Fatr for all processed windows.

*Absolute Jitter (Jita/usec)*

An evaluation of the period to period variability of the pitch period within the analysed voice sample. Voice break areas are excluded. Jita is cbmputed as :

$$\text{Jitta} = \frac{1}{N-1} \sum_{i=1}^{N-1} \left| To^{(i)} - To^{(i+1)} \right|$$

where, $To^{(i)}$, $i = 1, 2, \ldots N$ extracted pitch period data
$N$ = Number of extracted pitch periods.

Absolute jitter measures the very short term (cycle-to-cycle) irregularity of the pitch periods in the voice sample This measure is widely used in the research literature on voice perturbation (Iwata and Vonleden, 1970). It is very sensitive to the pitch variations occurring between consecutive pitch periods. However, pitch extraction errors may affect absolute jitter significantly.

The pitch of the voice can vary for a number of reasons, cycle-to-cycle irregularity can be associated with the inability of the vocal cords to support a periodic vibration for a defined period. Usually this type of variation is random. They are typically associated with hoarse voices.

Both Jittai and Jitt represent evaluations of the same type of pitch perturbation. Jitta is an absolute measure and shows the result in micro seconds which makes it dependent on the average fundamental frequency of voice. For this reason, the normative values on Jita for men and women differ significantly. Higher pitch results into lower Jitla. That's why, the Jitta value of two subjects with different pitch are difficult to compare .

*Jitter Percent (Jitt) (96)*

Relative evaluation of the period-to-period (very short term) variability of the pitch within the analysed voice sample. Voice break areas are excluded. It is computed

$$Jitt = \frac{\frac{1}{N-1}\sum_{i=1}^{N-1}\left|T_o^{(i)} - T_o^{(i+1)}\right|}{\frac{1}{N}\sum_{i=1}^{N} T_o^{(i)}}$$

where, $T_o^{(i)}$, $i = 1, 2, \ldots, N$ extracted pitch period data
$N$ = PER, Number of extracted pitch periods.

Jitter percent measures the very short term cycle-to-cycle irregularity of the pitch period of the voice. Jitt is a relative measure and the influence of the average fundamental frequency of the subject is significantly reduced.

*Relative Average Perturbation (RAP) (96)*

Relative evaluation of the period-to-period variability of the pitch within the analysed voice sample with smoothing factor of three periods. Voice breaks areas are excluded. It is computed as :

$$RAP = \frac{\frac{1}{N-1}\sum_{i=2}^{N-1}\left|T_o^{(i-1)} + T_o^{(i)} + T_o^{(i-1)}\right|}{\frac{1}{N}\sum_{i=1}^{N} T_o^{(i)}}$$

where, $T_o^{(i)}$, $i = 1, 2, \ldots, N$ extracted pitch period data
$N$ = PER, Number of extracted pitch periods.

Relative average perturbation measures the short term (cycle-to-cycle with smoothing factor of three periods) irregularity of the pitch period of the voice. The smoothing reduces the sensitivity of RAP to pitch extraction errors. However, it is less sensitive to the very short term period-to-period variations, but describes the short-term pitch perturbation of the voice very well.

The pitch of the voice can vary for a number of reasons, cycle-cycle irregularity can be associated with the inability of the vocal cords to support a periodic vibration with a defined period. Hoarse and/or breathy voices may have an increased RAP.

*Pitch Period Perturbation Quotient (PPQ 96)*

Relative evaluation of the period-to-period variability of the pitch within the analysed voice sample with a smoothing factor of five periods. Voice break areas are excluded. PPQ is computed as,

$$APQ = \frac{\frac{1}{N-4} \sum_{i=1}^{N-4} \left| \frac{1}{5} \sum_{r=0}^{4} To^{(i-r)} + To^{(i-2)} \right|}{\frac{1}{N} \sum_{i=1}^{N} To^{(i)}}$$

where $To^{(i)}$ , $i = 1, 2, \ldots, N$ extracted peak-to-peak amplitude data

$N$ = Number of extracted impulses.

PPQ measures the short-term (cycle-to-cycle with a smoothing factor o five periods) irregularity of the pitch period of the voice. The smoothing reduces the sensitivity of PPQ to pitch-extraction errors while it is less sensitive to period-to-period variations, it describes the short-term pitch purturbation of the voice very well. Hoarse and/or breathy voices may have an increased PPQ.

***Smoothed Pitch Period Perturbation Quotient (SPPQ 96)***

Relative evaluation of the short or long term variability of the pitch period within the analysed voice sample at smoothing factor defined by the user. The factory setup for the smoothing factor is 55 periods. Voice break areas are excluded.

$$SPPQ = \frac{\frac{1}{N-Sf+i} \sum_{i=1}^{N-Sf+1} \left| \frac{1}{Sf} \sum_{r=0}^{Sf-1} To^{(i-r)} + To^{(i-m)} \right|}{\frac{1}{N} \sum_{i=1}^{N} To^{(i)}}$$

where, $To^{(i)}$ , $i = 1, 2, \ldots, N$ extracted peak to peak amplitude data
$N$ = PER, Number of extracted impulses.
SF = Smoothing factor.

SPPQ allows the experimenter to define his own pitch perturbation measure by changing the smoothing factor from 1 to 99 periods. This is desirable because in the scientific literature researchers use pitch perturbation measures with different smoothing factor or without smoothing.

With a small smoothing factor, SPPQ is sensitive mostly to the short-term pitch variation of the voice impulses. With a smoothing factor of 1 (no smoothing), SPPQ is identical to Jitter percent (Jitt). It is very sensitive to the pitch variations occurring between consecutive pitch periods. Usually this type of variation is random. It is typical for hoarse voices. However, pitch extraction errors may affect jitter percent significantly.

With a smoothing factor of 3, SPPQ is identical to the relative average perturbation introduced by Koike (1973). With a smoothing factor of 5, SPPQ is identical to the pitch perturbation quotient introduced by Koike and Calcatera (1977). At high smoothing factors SPPQ correlates with the intensity of the long-term pitch period variations. The studies of patients with spasmodic dysphonia (Deliyski, et al. 1991) show that SPPQ with smoothing factor set in the range 45-65 period has increased values in case of regular long-term pitch variations (frequency voice tremors).

The SPPQ smoothing factory set-up is 55 periods. This set up allows using SPPQ as an additional evaluation of the frequency tremors in the voice. The intensity and the regularity of the frequency tremors can be assessed using SPPQ (55) in combination with VFo. The difference between VFo and SPPQ (55) is that VFo represents a general evaluation of the fundamental frequency (pitch) variation of the voice signal. The VFo value increases regardless of the type of pitch variation. Either random or regular short-term or long-term variations increase the value of VFo. However, SPPQ (55) is more sensitive to regular long term variations with a period near and above 55 pitch periods. If both SPPQ (55) and VFo are low, the intensity of pitch variations in the voice signal is very low. If VFo is high but SPPQ (55) is low, there are pitch variations but not a long-term periodic one. If both SPPQ (55) and VFo are high, there is a long-term periodic pitch variation (most likely a frequency tremor).

*Coefficient of Fo Variation VFo /%/*

Relative standard deviation of the fundamental frequency. It reflects, in general, the variation of Fo (short to long- term), within the analysed voice sample.

Vtiice break areas are excluded.

$$VFo = \frac{\frac{1}{N}\sum_{i=1}^{N}\sqrt{\frac{1}{N}\sum_{j=1}^{N}\left|\frac{1}{N} Fo^{(i)} - Fo^{(i)}\right|^{(2)}}}{\frac{1}{N}\sum_{i=1}^{N} Fo^{(i)}}$$

Where, $Fo^{(i)}$, = 1, 2,.... N extracted peak-to-peak amplitude data

N = PER, number of extracted pitch periods

VFo reveals the variations in the fundamental frequency. The VFo value increases regardless of the type of pitch variation. Either random or regular short-term or long-term variations increases the value of VFo. Because the sustained phonation normative thresholds assume that the Fo should not change, any variations in the fundamental frequency are reflected in VFo. These changes could be frequency tremors or non-periodic changes, very high jitter or simply rising a falling pitch over the analysis length.

*Shimmer in dB (ShdB)/dB/*

Evaluation in dB of the period-to-period (very short-term) variability of the peak-to-peak amplitude within the analysed voice sample. Voice break areas are excluded. ShdB is computed as,

$$ShdB = \frac{1}{n-1}\sum_{i=1}^{N-1}\left|20 \log \left(\frac{A^{(i+1)}}{A^{(i)}}\right)\right|$$

where, $A^{(i)}$, i = 1,2, .... N extracted peak-to-peak amplitude

N = Number of extracted impulses.

Shimmer in dB measure the very short term (cycle-to-cycle) irregularity of peak-peak amplitude of the voice. This measure is widely used in the research literature on voice perturbation (Iwata and Von Leden, 1970). It is very sensitive to the amplitude variation occurring between consecutive pitch periods. However, pitch extraction errors may affect shimmer percent significantly.

The amplitude of the voice varies due to several factors. Cycle-to-cycle irregularity of amplitude can be associated with the inability of the vocal folds to support d periodic vibrdtion for a defined period and with the presence of turbulent noise in the voice signal usually, this type of variation is random. It is typically associated with hoarse and breathy voices. APQ is the preferred measurement for Shimmer because it is less sensitive to pitch extraction errors while still providing a reliable indication of short-term amplitude variability in the voice.

Both shim and ShdB are relative evaluations of the same type of amplitude perturbation but they use different measures for the result-percent an dB.

*Shimmer Percent (96)*

Relative evaluation of the period-to-period (very short term) variation of the peak-to-peak amplitude within the analysed voice sample. Voice break means are excluded.

$$Shim = \frac{\frac{1}{n-1}\sum_{i=1}^{N-1}\left| A^{(i)} - A^{(i+1)} \right|}{\frac{1}{N}\sum_{i=1}^{N} A^{(i)}}$$

where, $A$ , $i = 1, 2,\_\_\_\_N$ extracted peak-to-peak amplitude
N = Number of extracted impulses.

Shimmer percent measure the very short term (cycle-to-cycle) irregularity of the peak-to-peak amplitude of the voice.

*Amplitude Perturbation Quotient (APQ) (%)*

Relative evaluation of the period-to-period variation, variability of the peak-to-peak amplitude within the analysed voice sample at smoothing of 1 1 periods. Voice break areas are excluded.

$$APQ = \frac{\frac{1}{N-4} \sum_{i=1}^{N-4} \left| \frac{1}{5} \sum_{r=0}^{4} A^{(i+r)} - A^{(i+2)} \right|}{\frac{1}{N} \sum_{i=1}^{N} A^{(i)}}$$

where, $A$ , $i = 1, 2,\_\_\_\_N$ extracted peak-to-peak amplitude
N = Number of extracted impulses.

APQ measures the short-term (cycle-to-cycle with smoothing factor of 1 1 periods) irregularity of the peak-to-peak amplitude of the voice. While it is less sensitive to the period to period amplitude variations it still describes the short-term amplitude perturbation of the voice very well breathy and hoarse voice usually have an increased APQ. APQ should be regarded as the preferred measurement for shimmer in MDVP.

*Smoothed Amplitude Perturbation Quotient (SAPQ (96)*

Relative evaluation of the short or long term variability of the peak-to-peak amplitude within the analysed voice sample at smoothing factor defined by the user. The factory set up for the smoothing factor is 55 periods (providing relatively long-term variability,- the user can change this value as desired). Voice break areas are excluded.

$$SAPQ = \frac{\frac{1}{N-Sf+1} \sum_{i=1}^{N-Sf+1} \left| \frac{1}{Sf} \sum_{r=0}^{Sf-1} A^{(i+r)} - A^{(i+m)} \right|}{\frac{1}{N} \sum_{i=1}^{N} A^{(i)}}$$

where, $A^{(i)}$ ,i = 1, 2,_____N extracted peak-to-peak amplitude

N = Number of extracted impulses.

SF = Smoothing factor.

*Coefficient of Amplitude Variation (VAm) /%/*

Relative standard deviation of peak-to-peak amplitude. It reflects in general to peak-to-peak amplitude variations (short to long term) within the analysed voice sample, voice break areas are excluded. VAm is computed as the ratio of the standard deviation to the average value of the extracted peak-to-peak amplitude data as

$$VA_m = \frac{\sqrt{\frac{1}{N} \sum_{i=1}^{N} \left| \frac{1}{N} \sum_{j-1}^{N} A^{(j)} - A^{(i)} \right|^{(2)}}}{\frac{1}{N} \sum_{i=1}^{N} A^{(i)}}$$

where, $A^{(i)}$ , i = 1 , 2 ,____N extracted peak-to-peak amplitude

N = Number of extracted impulses.

VAm reveals the variations in the cycle-to-cycle amplitude of the voice. The VAm value increases regardless of the type of amplitude variation. Either random or regular short-term or long-term variation increases the value of VAm.

Noise to harmonic Ratio (NHR)

Average ratio of the inharmonic spectral energy in the frequency range 1 500-4500 Hz to the harmonic spectral energy in the frequency range 70-4500 Hz. This is general evaluation of noise present in the analysed signal.

NHR is computed using a pitch synchronous frequency domain method. In general terms the algorithm functions as follows:

A. Divides the analysed signal into windows of 81 .92 ms (4096 points at 50 kHz sampling rate or 2048 at 25 kHz). For every windows the following steps apply.

1 . Low pass filtering 6 kHz (order 22) with Hamming window, down sampling of the single data down to 1 25 kHz and conversion of the real signal into an analytical one using the Hilbert transform.

2. 1 024 points complex fast Fourier Transform (FFT) on the analytical signal corre sponding to a 2048 - points FFT on real data.

3. Calculation of the power spectrum from the FFT.

4. Calculation of the average fundamental frequency within the window synchronously with the pitch extraction results.

5. Harmonic/inharmonic separation of the current spectrum synchronously with the current window fundamental frequency.

6. Computation of the noise-to-harmonic ratio of the current window. NHR is the ratio of the inharmonic (1 500-4500 Hz) to the harmonic spectral energy (70-4500 Hz).

B. Computes the average values of NHR for all previously processed windows.

Increased values of NHR are interpreted as increased spectral noise which can be due to amplitude and frequency variations (i.e. Shimmer and Jitter) turbulent noise, subharmonic components and/or breaks which affects NHR. NHR globally measures the noise in the signal (includes contributions of jitter, shimmer and turbulent noise,).

Voice turbulence index (VTI)

Average ratio of the spectral inharmonic high frequency energy in the range 2800-5800 Hz to the spectral harmonic energy in the range 70-4500 Hz in the areas of the signal where the influence of the frequency and amplitude variations, voice breaks and subharmonic components are minimal. VTI measures the relative energy level of high frequency noise.

VTI is computed using a pitch synchronous frequency domain method. The algorithm consists of the following steps :

For every window, the following steps apply:

A. Selects upto four but atleast two 81 .92 msec windows where the frequency and amplitude perturbations are lowest for the signal. These windows are located in different as of the signal and donot include voice break and subharmonic components.

1 . Low-pass filtering at 6 kHz.
2. Down sampling 12.5 kHz.
3. Conversion of the real signal to analytical one.
4. Computation of a 1024 points complex fast Fourier transform on the analytical signal.
5. Computation of power spectrum from the FFt.
6. Calculation of the average fundamental frequency within the window.
7. Harmonic/inharmonic separation of the current spectrum synchronously with the current window Fo.
8. Computation of the VTI for every window, VTI is the ratio of the spectral inharmonic high frequency energy (2800-5800 Hz) to the spectral harmonic energy (70-4500 Hz).

B. Calculate the average VTI values for all processed windows. VTI measures the relative energy level of high- frequency noise.

VTI mostly correlates with the turbulence caused by incomplete or loose adduction of the vocal folds. VTI, unlike NHR, analyses high frequency components to extract an acoustic correlate to "breathiness". However it is unlikely that users will Find a one-to-one correspondence between their perceptual impression of a voice and this acoustic analysis. However, VTI is a new attempt to compute a parameter which correlates with breathiness. Because VTI is a new parameter, normative values cannot be found in the professional literature.

Soft phonation index (SPI)

Average ratio of the lower-frequency harmonic energy in the range of 70-1 600 Hz to the higher frequency harmonic energy in the range 1 600-4500 Hz.

SPI is computed using a pitch synchronous frequency domain method. The algorithm does the following procedures.

A. Divides the analysed signal into windows of 81 -92 ms.

For every one of these windows, the following steps apply:

1. Low-pass filtering at 6 kHz order 22 with Hamming window, down sampling of the signal data down to 12.5 Hz and conversion of the real signal ratio analytical one using Hilbert transform.
2. 1 024 points complex fast Fourier transform on the analytical signal.
3. Computation of the power spectrum from the FFT.
4. Calculation of the average Fo with in the window synchronously with the pitch extraction results.
5. Harmonic/inharmonic separation of the current spectrum synchronously with the current window Fo.

6. Computation of SPI of the current window. SPI is a ratio of the lower- frequency (70-1 600 Hz) to the higher frequency (1 600-4500 Hz) harmonic energy.

B. Computes the average of SPI for all previously processed windows.

SPI can be thought of as an indicator of how completely or tightly the vocal folds adduct during phonation. Increased value of SPI is generally an indication of loosely or incompletely adducted vocal folds during phonation. However, it is not necessarily an indication of a voice disorder. Similarly, patients with "pressed" phonation may likely have a "normal" SPI though their pressed voice characteristic may not be desirable. Therefore, a high SPI value is not necessarily bad, nor a low SPI value necessarily good. Subjects with glottal chinks (determined stroboscopically) or with high phonatory air flow rates often exhibit an increased SPI. Spectral analysis will show a well defined higher formants when SPI is low, and less well defined when SPI is high.

SPI is very sensitive to the vowel formant structure because vowels with lower high frequency energy will result in higher SPI, only values computed for the same vowel can be compared.

Increased SPI values may be due to a number of factors. The subject may have a "soft" phonation because of a voice or speech disorder and may not be able to strongly adduct his vocal folds. However, the subject may naturally speak with a softer "attach" and hence have an elevated SPI. Psychological stress could also be a factor that may increase SPI. Another important factor is the amplitude of the sustained vowel. If the subject phonates softly, SPI may be high.

Frequency tremor intensity index (FTRI) /%/

Average ratio of the frequency magnitude of the most intensive low-frequency modulating component (Fo-tremor) to the total frequency magnitude of the analysed voice signal.

The method for frequency tremor analysis consists of the following steps

A. Division of the fundamental frequency period to period (Fo) data into 2 sec. windows. For every window, the following procedures apply :

1. Low-pass filtering of the Fo data at 30 Hz and down sampling at 400 Hz.
2. Calculation of the total energy of the resulting signal.
3. Subtraction of the DC component.
4. Calculation of an autocorrelation function on the residue signal.
5. Division by total energy and conversion to per cent.
6. Extraction of the period of variation.
7. Calculation of Fftr and Ftri corresponding to the period of variation found.

B. Computation of the average autocorrelation curve and *average* Ftri for all processed windows.

The algorithm for tremor analysis determines the strongest periodic frequency and amplitude modulation of voice. Tremor has both frequency and amplitude components (i.e. the Fo may vary and/or the amplitude of the signal may vary in a periodic manner). Tremor frequency provides the rate of change with Fftr providing the rate of periodic tremor of the frequency and Fatr providing the rate of change of the amplitude. The program will determine the Fftr and Fatr of any signal if the magnitude of these tremors is above a low threshold of detection. Therefore, the magnitude of the frequency tremor and the magnitude of the amplitude tremor are more significant than the respective frequencies of the tremor.

Amplitude tremor intensity index (ATRI) /%/

Average ratio of the amplitude of the most intense low-frequency amplitude modulating component to the total amplitude of the analysed voice signal.

The method for computation is same as Ftri except that here the peak-to-peak amplitude data has been taken into consideration instead of Fo data.

Degree of voice breaks (DVB) /%/

Ratio of the total length of areas representing voice breaks to the time of the complete voice sample.

$$DVB = \frac{t1 + t2 + .... + tn}{Tsam}$$

Where, T1 , T2, ... tn - lengths of the 1st, 2nd, ... voice breaks.

Tsam - Length of analysed voice data samples.

DVB does not reflect the pauses before the first and after the last voiced areas of the recording. It measures the ability of the voice to sustain uninterrupted voicing. The normative threshold is o' because the normal voice, during the task of sustaining voice, should not have any voice break areas. In case of phonation with pauses (such as running speech, voice breaks, delayed start or earlier end of sustained phonation) DVB evaluates only the pauses between the voiced areas.

of sub-harmonic components (DSH) / % /

Relative evaluation of sub-harmonic to Fo components in the voice sample. DSH is computed as a ratio of the number of autocorrelation segments where the pitch was found to be subharmonic of the real pitch (NSH) to the total number of autocorrelation segments.

The *degree* of sub-harmonic components in normal voices should be equal to zero. It is expected to increase invoices where double or triple pitch periods replace the fundamental in certain segments over the analysis length. These effects are typical for diplophonic voices and voices with glottal fry. The experimental observation of patients with functional dysphonia or neurogenic voice disorders may show increased values of DSH.

Degree of voiceless (DOV) / % /

Estimated relative evaluation of non-harmonic areas (where Fo cannot be detected) in the voice samples.

DOV is computed as a ratio of the number of autocorelation segments where an unvoiced decision was made to the total number of autocorrelation segment. DOV measures the ability of the voice to sustain uninterrupted voicing. The normative threshold is 'o' because of normal voice, in the defined task of sustaining voicing, should not have any voiceless segments. In case of phonation with pauses (such as running speech, voice breaks, delayed start or *earlier* end of sustained phonation), DOV also evaluates the pauses before, after and/or between the voiced areas.

Number of voice breaks (NVB)

Number of times the fundamental period was interrupted during the voice sample (measured from the first detected period to the last period).

NVB does not reflect the pauses before the first and after the last voiced areas of the recording. However, like NVB, it measures the ability of the voice to sustain uninterrupted voicing. The normative threshold is 'o' because of normal voice, during the task of sustaining voice, should not have any voice breaks. In cases of phonation with pauses (such as running speech, voice breaks, delayed start or earlier end of sustained phonation), NVB evaluates only the pauses between the voiced areas.

Number of subharmonic segments (NSH)

Number of autocorrelation segments where the pitch was found to be a subharmonic of Fo.

The number of sub-harmonic components in normal voices should be equal to zero. It is expected to increase invoices where double or triple pitch period replaces the fundamental in certain segments over the analysis length. These effects are typical for diplophonic voices and voices with glottal fry.

Extent of fluctuation in fundamental frequency

The extent of fluctuation as defined as the percent score of the ratio of the peak to peak value of fluctuation (Fo) to the mean fundamental frequency (Fo).

This has been defined the peak to peak value in decibels measured on an average amplitude display.

Extent of fluctuation in intensity

This has been defined as the peak to peak value in decibels measured on an average amplitude display.

Speed of fluctuation of intensity

This was defined as the number of positive peaks on an amplitude display within one sec. Peaks of 3 dB or greater from adjacent trough have been counted.

Normalised Noise Energy

The normalised noise energy measures primarily the turbulence noise caused by the closing insufficiency of the glottis during the phonation.