MULTI DIMENSIONAL PROFILE OF VOICE IN DYSPHONICS

Shameen Taj (D.K)
Reg. No. M 9719

A DISSERTATION SUBMITTED AS PART
FULFILMENT OF FINAL YEAR M.Sc. (Speech and Hearing)

All India Institute of Speech and Hearing
Mysore

May 1999

Dearest Daddy and Mummy

You helped me see the invisible,
feel the intangible and achieve
the impossible

with love always........

Certificate

This is to Certify that the disseration entitled **"MULTI DIMENSIONAL PROFILE OF VOICE IN DYSPHONICS"** *is the bonafied work in part fulfilment for the degree of 'Master of Science (Speech and Hearing) of the student with Register No. M. 9719.*

Mysore
May, 1999

Dr. S. Nikam
Director
All India Institute of Speech & Hearing
Mysore - 570 006

Certificate

This is to Certify that the disseration entitled **"MULTI DIMENSIONAL PROFILE OF VOICE IN DYSPHONICS"** *has been prepared under my supervision and guidance.*

Mysore
May, 1999

Dr. N.P. Nataraja
Professor and H.O.D
All India Institute of Speech & Hearing
Mysore - 570 006

DECLARATION

This dissertation entitled ***"MULTl DIMENSIONAL PROFILE OF VOICE IN DYSPHONICS"*** *is the result of my own study under the guidance of Dr. N.P. Nataraj, Professor and H.O.D. 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, 1999 ***Reg. No. M 9719***

ACKNOWLEDGMENTS

The Dictionary's full words of every length and kind, but I am at a loss of words to convey my gratitude to my teacher and guide Dr. N.P. Nataraja, Professor and Head, Department of Speech Sciences, All India Institute of Speech and Hearing, Mysore.

Thank you Sir, for you constant support, encouragement and guidance.

I would like to thank Dr. (Miss). S.Nikam, Director, All India Institute of Speech and Hearing, Mysore, for permitting me to carry out this study.

Mrs. Sreedevi, Ms Yashoda, Mrs Sangeetha, Ms Lalitha and Mrs Rohini - Thank you for your patience and help.

Dear grand parents - Thank you for your inspiration and prayers.

Zubair, its good to know that you are always there with advice when its asked for and with encouragement when its needed.

Sa-ad - being with you lifts up my spirit. Just to let you know that you mean a lot to me.

Shireen, Seema, Riaz and Jaleel - an informal thanks to all in encouraging me and giving me a helping hand whenever needed.

Zain, Simran, Shihab, Saman - you all add colour to my life. Here wishing you all "All the best".

Sarah - Thanks for being such a good friend and making my life exciting in A.I.I.S.H. I shall miss you ton es.

Hats off to you - Sammer and Joby - Thank you for having help me in my work.

Renu, Shivani, Narayan, Deepak - Thanks for being such lovely friends.

Radhika, Mili, Sanyukta - Thank you for your friendship which has added colour to my life.

Achala, Amala, Sonal, Mili, Vimi - Thanks for lending me a helping hand whenever needed.

A special thanks to Mr. Gopalkrishna, Mr. Nandakumar, Mrs. Rasita, for bringing out the best in me.

All my dearest classmates - Thanks a ton for accepting the way I am and making my stay at A.I.I.S.H a memorable one.

CONTENTS

PAGE NO.

LIST OF TABLES

LIST OF GRAPHS

LIST OF FIGURES

INTRODUCTION

The importance of human voice in modern society cannot be overstated. It is the primary instrument through which most of we project our personalities and influence our compatriots. Voice is a "Window" to many functions performed by the body.

Voice has been defined as "the laryngeal modulation of the pulmonary air stream, which is further modified by the configuration of vocal tract" (Micheal and Wendahl, 1971).

Fant (1960) defines voice by using the formula p = ST, where "p" is the product of source 's' and the transfer function of the vocal tract T. An attempt has been made by Nataraj and Jayaram (1975) to review the definitions of the normal voice critically. They conclude that each of the available definitions of voice have used subjective terms, which are neither defined or measurable. They have suggested a possibility of defining good voice operationally by stating that "the good voice is one which has optimum frequency as its fundamental (habitual) frequency".

The production of voice is a complex process. It depends on the •synchrony between respiratory, phonatory and resonatory systems which in turn requires precise control by the central nervous system.

Hirano (1981) states that, "during speech and singing the higher order centers in the cerebral cortex control voice production, and all activities of the central nervous system is finally reflected in the muscular activity of the voice organs". Because of the interdependence of the respiratory, phonatory and resonatory systems during voice production, disturbances in any one of the systems may lead to deviant or abnormal voice quality.

Since voice plays a major role in speech and hence communication, it needs to be constantly monitored and in the event of abnormal functioning voice an immediate assessment should be undertaken.

The ultimate aim of the studies on normality and abnormality of voice, and assessment and diagnosis of voice disorders is to enforce a procedure which will eventually bring back the voice of an individual to normal or optimal level.

There are various means of analyzing voice, developed by workers, to note the factors which are responsible for creating an impression of a particular "voice" (Hirano, 1981; Nataraj, 1979; Rashmi, 1985).

The human ear has a remarkable capacity to identify and discriminate varying sound complex. One can identify the speakers simply by listening to the voice. Well trained voice clinicians are frequently able to determine the causative pathologies on the basis of psyshoacoustic impression of voice (Takhashi, 1974; Takhashi et al, 1974; Hirano, 1975). The psycho-acoustic evaluation of voice is done based on pitch, loudness and quality of voice sample. Due to its subjectivity, the perceptual judgment of voice has been considered less worthy than the objective measurement.

There are objective methods like EMG, stroboscopy, ultra-sound glottography, ultra-high speed photography, photo-electric glottography, electroglottography, aerodynamic measurements, acoustic analysis etc.

Micheal and Wendahl (1971) consider voice as a series of measurable events and suggest twelve parameters for assessing the voice and voice disorders. Nataraj (1986) studied 22 acoustic and aerodynamic parameters in normals and dysphonics. He concluded that only six parameters were sufficient to differentially diagnose dysphonics from normals.

Precently 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 funciton 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 clipping inverse filtering, computer based methods and others.

In computer based techniques there are many software programs which are designed to extract different parameters of voice. However, the software program used in the present study "Multidimensional voice program - model 4305" developed and marketed by Kay Elemetrics Inc., New Jersey, acquires, analyses and displays thirty three voice parameters from a single vocalization. This program uses the computerised speech Lab hardware system for the signal acquisition, analyses and playback. Thirty three extracted parameters are available as numerical file or they can be displayed graphically in comparison with a data base. This comparison graphically provides a visual "snapshot" of clients vocalization. This graphic analysis can also be printed for a patient's file. The analysis is completed in just 4 seconds.

The advantage of a multiple parameters extraction is that different parameters are important for the diagnosis of different vocal pathologies. For example, a breathy voice may have normal jitter values but the degree of breathiness is likely to be revealed in the extrcated turbulance parameter.

The tremor parameters will measure the modulation of the voice by analysing the voice and extracting amplitude and frequency tremor rate and amplitude. A patient with parkinson's disease may have a normal voice except for the tremor.

After the measurements of parameters, the clinician has to make a judgement regarding the diagnosis based on the values of parameters. Since the computers have been put into use in almost all spheres of life, attempts have been made to use them for classifying the disorders, based on the symptoms and values of parameters measured.

An Artificial Neutral Network is a computer branch which has been widely applied for such activities.

Neural networks are simply a new way of analyzing data. The revolutionary aspect of neural networks which makes them so unique is their ability to learn complex patterns and trends in date. Neural networks acquire knowledge by training a set of data. After the network has been trained and validated, the model may be applied to data it has not seen previously for prediction classification, time series analysis or data segmentation.

Artificial Neural Networks (ANNs) are 'biologically' inspired networks. They have ability to learn from empirical data or information (Raol and Mankame, 1996). ANNs have apparent ability to imitate the brain's activity to make decisions and draw conclusion when presented with complex and noisy informations (Raol and Mankane, 1996).

Studies using neural network in the field of speech and hearing are scanty and ANN's applied to the field of voice and voice disorders are still less.

Few studies have used self organised Map (SOM) which is an algorithm of Kohenen network to classify normal and dysphonics voices to find objective indices for dysphonics voice (Leinonen, Kangas, Torkkola and Juvas, 1992), to find acoustic pattern recognition of fricative-vowel coarticulation by self organizing map (Leinonen, Mujumen, Kangas, Torkkola and Juvas, 1993), acoustic pattern recognition of /s/ misarticulation by SOM (Mujunen, Leinonen, Kangas and Torkkola, 1993), model for control of voice fundamental frequency by a neural network. (Farley, 1994) and categorization of voice disorders on perceptual dimensions (Leinonen, Kangas, Torkkala and Munjunen, 1997).

Hence the present study attempts to classify voice disorders in terms of its acoustic parameters by using a neural network.

In the present study, multi dimensional voice program software was used to analyse the acoustic parameters.

The acoustic parameters which was measured include

1. Average fundamental frequency (Fo).
2. Highest fundamental frequency (Fhi).
3. Lowest fundamental frequency (FLo).
4. Standard duration of Fo (STD).
5. Phonatory Fo - range in semi tones (FPR).
6. Fo - tremor frequency (FFtr).
7. Amplitude tremor frequency (Fatr).

8. Absolute jitter (Jita).
9. Jitter percentage. (Jitt)
10. Relative Average perturbation. (RAP)
11. Pitch perturbation quotient. (PPQ)
12. Smoothed pitch perturbation quotient. (SPPQ)
13. Fundamental frequency variation. (vFo)
14. Shimmer in dB. (ShdB)
15. Shimmer in percentage. (Shim)
16. Amplitude perturbation quotient. (APQ)
17. Smoothened amplitude perturbation quotient. (SAPQ)
18. Peak -amplitude variation. (vAm)
19. Noise to harmonic ratio. (NHR)
20. Voice turbulence index. (VTI)
21. Soft phonation index. (SPI)
22. Fo - tremor intensity index. (FTRI)
23. Amplitude - tremor intensity index. (ATRI)
24. Degree of voice breaks. (DVB)
25. Degree of sub harmonics. (DSH)
26. Degree of voiceless. (DUV)
27. No: of sub -harmonic segments. (NSH)
29. No: of un voiced segments. (NUV)

These acoustic parameters were fed into a neural network. The neural network used for the present study was multi layer perception (MLP).

Hypothesis :

I. There is no significant difference across the five sessions in terms of the different acoustic parameters measured.

II. a. There is no significant difference between the two groups - normal males and normal females interms of the different acoustic parameters measured.
    b. There is no significant difference between the two groups - normal males and dysphonics males interms of the different acoustic parameters measured.
    c. There is no significant difference between the two groups - normal females and dysphonics females interms of the different acoustic parameters measured.
    d. There is no significant difference between the two groups - dysphonics males and dysphonics females interms of the different acoustic parameters measured.

Implications of the study :

1. The voice is analysed objectively which is then fed into artificial neural network which has the capability to learn and understand complex data and hence helps in objectively diagnosing vocal pathology.

2. Both analysis and interpretation is done objectively leaving no corner for bias.

3. Along with being highly objective, this procedure is time saving also. Hence this can be used as a regular clinical activity.

4. Using MDVP the specific parameter of voice which is affected in dysphonics can be tapped. Hence more efficient treatment can be aimed by treating these specific aspect of voice.

Limitations of the study :

1. The number of subjects in the study was limited.

2. The number of subjects were not uniformly distributed in terms of age.

REVIEW OF LITERATURE

"Man's need for communication with his fellowmen is possibly the greatest need and the fulfillment of his other needs and desires is largely dependent upon, or at the last greatly facilitated by his ability to satisfy his basic one", (Louise Tracy, 1970).

Voice is the primary means of expression. Voice is more than a means of communicating verbal messages clearly. It serves as a powerful conveyer of personal identity, emotional state, education and social status. It is because of this that impairment of vocal function or complete loss of voice is so distressing to any individual. Voice constitutes the matrix of verbal communication, infusing all parameters of human speech and the unique self we present to the world.

Voice is a "window" to many functions performed by the body. Hoaeshess can be an indication of viral infection, weakness and tremor in the voice can be an early sign of neurologic disease, a change in pitch can suggest a fluid imbalance, and a "strangled voice can result from severe emotional trauma.

A voice disorder exists when a persons quality, pitch and loudness differ from those of similar age, sex cultural background and geographic locations (Aronson, 1980; Boone, 1977; Greene, 1972; Moore 1971).

According to Wilson (1979), analysing voice is an important step in the management of voice disorders in children and adolescents. He states that "the voice analysis should be done in detail adapted to the type of problem presented. Following this, the goals of therapy are listed, the voice therapy schedule planned and a prognosis assigned. "Diagnosis is intended to define the parameters of the problem, determine etiology and outline a logical course of action" (Emerick and Hatten, 1979).

The nature of sound generated is chiefly determined by the vibratory pattern of the vocal folds. It can be specified both in acoustic terms and in psycho-acoustic terms. The psycho-acoustic parameters are naturally dependent on the acoustic parameters. The acoustic parameters are fundamental frequency, intensity, acoustic spectrum and their time-related variations. The psychoacoustic parameters are pitch, loudness and quality of the voice and their time related changes.

The crucial event essential for voice production is vibration of the vocal folds. Accordng to Hirano (1981) "the parameters which regulate the vibratory pattern of the vocal folds can be divided into two groups: physiological and physical. The physiological factors are those related to the activity of the respiratory, phonatory and articulatory muscles. The physical factors include the expiratory force, the conditions of the vocal folds and the state of the vocal tract".

The production of voice depends upon three primary factors:

1. Pulmonic pressure - supplied by the respiratory system.
2. Laryngeal vibration - phonation.
3. Transfer function of the vocal tract-resonance.

The production of voice depends on the synchrony or co-ordination between the above systems.

The process of respiration involves inspiraiton and expiration of air. For the purpose of voice production, the air during exhalation is used.

The essential function of the folds as vibrators is now commonly accpeted. Divergence among the present day "theories" of voice production arises when the way the vocal folds are set into vibration is questioned. The classical theory of voice production, knwon as, Myoelastic-aerodynamic theory, is in direct oppostion to the more recent neurochronaxic theory.

The proponent of the Myoelastic-aerodynamic theory was Muller (1943). Minor modifications were suggested by Tandorf (1955); Smith (1954) and Titze (1994). However, the salient features of Muller's version have remained unchaged. The vocal folds are set into vibration by the air stream from the lungs which are in adducted position by pushing them apart. This reduces the air pressure below the glottis. Because of the elasticity of the muscles and the sudden decrease in pressure at the level of the glottis, the vocal cords are sucked back to the midline. This causes an obstruction of air passage and leads to building of pressure at the subglottal region. As the subglottal air pressure reaches a level, it agian pushes the vocal cords apart, thus setting the vocal folds into to and fro motion. The vibratory movements of the vocal cords are performed at a frequency determined by the mass, length and tension of the vocal cords. Their frequency of vibration in turn determines the frequency of the air puffs which are the primary source of sound. Thus the vocal folds are set into vibration by the air flow and the myoelastic conditions. The frequency of sound is determined by the mass, length and tension of the vocal cords.

Husson (1950) the proponent of the neurochronaxic theory considers the vocal fold vibration as an active neuromuscular process. He believes that the periodic opening of the vocal cords is controlled by action potentials of equal frequency, which are conducted to the fibers of the vocalis muscle through the recurrent laryngeal nerve. Each new vibratory cycle is initiated by a nerve impulse transmitted from the brain to the vocalis muscle by means of the recurrent laryngeal nerve. The frequency of vocal, fold vibration is thus dependent upon the rate of impulses delivered to the laryngeal muscles.

Zemlin (1968) concludes with reference to the studies in support and against these two theories by stating that whereas most normal laryngeal functions may be accounted for quite easily by the myoelastic-aerodynamic theory, very little conclusive evidence may be found to support the neurochronaxic theory of vocal fold vibration. Many experimenters had little difficulty in rejecting the neurochronaxic theory. However, a crucial test of either theory has not yet been performed". According to Fant (1960) the myoelastic-aerodynamic theory of voice production is most commonly accepted.

While commenting on the myoelastic aerodynamic theory of phonation, Titze (1980) states that "the myoelastic aerodynamic theory of phonation has been quantified and tested with mathmatical models. The models suggest that vocal fold oscillation is produced as a result of asymmetric forcing functions over closing and opening portions of the glottal cycle". He concludes that "it is possible, however, that the aerodynamic theory, as initially developed by Vanden Berg (1958) refined by Ishizaka and Matsudaira (1972) and fruther illuminated by Stevens (1977) and Broad (1979), is still insufficiently precise to predict quantitatively an aerodynamically controlled intonation pattern.

.......Experimental and theoretical work is in progress to provide further validation of the aerodynamic theories".

"When vibrating the vocal folds provide a wide spectrum of quasi-periodic molulations of the air stream, accounting for various tonal qualities, reflecting the different ways the vibrator behaves" (Brackett 1971). This tone produced by the larynx consists of frequencies approximately ranging from 80 Hz to 8KHz and includes fundamental and hormonic frequencies (Fletcher, 1959).

According to Judson and Weaver (1965) "the sounds produced by the laryngeal vibrator do not of themselves contribute voice. The laryngeal tone, cord tone or voice complex generated by this mechanism consists of a fundamental tone and a rich supply of overtones.....only as it is intensified and as its various partials are resonated does it become upon issuing from the mouth and nose, the human voice".

Some doubts have been raised as to whether resonance really amplifies sound or not. Berry and Eisenson (1956) talking about the determinants of voice state that "the physical tone and density of the resonating surfaces of the pharyngeal and oral cavities and the absorbent index of the resonators, both subglottal and supraglottal, also may alter the tonal character, although their precise influence is still debatable. The traditional concept of the resonators as amplifiers of sound has been quesitoned by some voice scientists. They claim that only 20 to 30% of the energy produced at the larynx is found in the oral cavity".

Boone (1971) reports an interesting case history wiht reference to this. "A patient with an open wound immediately superior to his laryngeal cartilages was found producing thin, soft voice with a voice quality which was inhuman".

Thus the tone generated by the laryngeal vibrator, gets amplified and modified as it passes through the vocal tract and becomes audible and gets a human quality.

Some investigators like Curtis (1968) do not consider the coupling between the generator and resonator important whereas some others like Michel and Wendahl (1971), take the view that the golttis is more closely coupled to the vocal tract. The degree of tuning existing between the resonator and the generator establishes not only the pitch, but also the loudness and the quality of the resultant tone.

It has been demonstrated by Wendahl and Page (1967) and Stevens (1967) that the frequency of vibration of the vocal cords is at least partially determined by the supraglottal resonators. The coupling therefore determines the pitch.

The resonance apparatus in human beings is made up of a series of connected cavities in the neck and head. The vocal tract may be thought of as a series of cylindrical sections with acoustical mass and compliance uniformly distributed along each section: (Gray and Wise, 1959). The present thinking is that different sections of the vocal tract may behave as tube or cavity resonators under varying conditions.

"A resonant voice is desirable not only for aesthetic but for practical reasons. Resonance, apart from rendering the voice pleasing, has the great advantage of giving it richness and carrying power (loudness or intensity) without, however, making any extra demands upon the physical or neural energy of the speaker. Voice teachers have consistently referred to the resonators as the chief, if not the sold determiners of voice quality" (Berry and Eisenson, 1956).

Emrickson (1959) is of the opinion that the vocal cords are the ultimate detemriners of pitch.

"...........both quality and loudness of voice are mainly dependent upon the frequency of vibration. Hence, it seems apparent that frequency is an important prarameter of voice" (Anderson', 1961).

Acoustic analysis has been considered as the basic tool in the investigation of voice disorder. It has been considered vital in the diagnosis and the management of patients with voice disorders. Hirano (1981) has pointed out that the acoustic analysis of the voice signal may be one of the most attractive methods for assessing phonatory function or laryngeal pathology because it is non-invasive and provides objective and quantitative data.

Further, a clinician will not really know what to expect with a medical diagnosis having a complete physical description of the larynx together with some adjectives like "hoarse" a "rough" until he actually sees the case (Micheal and Wendahl, 1971). On the other hand, if the clinician receives a report which includes measures of frequency ranges, respiratory function, jitter, shimmer, their related variations, noise and harmonic components etc., in the form of a voice profile, the clinician can then compare these values to the norms for each one of the parameters and thus have a relatively good idea as to how to proceed with therapy even before seeing the patient. Moreover, periodic measurement of these parameters during the course of therapy may well provide an useful index so as to the success of the treatment. (Micheal and Wendahl, 1971).

An objective method of locating optimum pitch was undertaken by Nataraja (1972). This was done by stimulating the vocal tract by an external sound source. A relation between the natural frequency of the vocal tract and the fundamental frequency was developed and it was found to be 8:1 for males in the age range of 20-25 years (Nataraja, 1972). A ratio of 5:1 was found between the two, in the same age range of female population (Shantha, 1973).

Jayaram (1975) has made an attempt to compare some of the parameters of voice between normals and dysphonics. A significant difference in the habitual frequency measures were got between the subjects of both the groups. Nataraja (1972), Samuel (1973), Shanta (1973), Sheela (1974) and Asthana (1977) have used stroboscope with Tacho unit and SPL meter to determine fundamental frequency of voice in their studies. The subjects were instructed to phonate a vowel in their normal speaking voice and this phonation was fed to the stroboscope through the SPL meter and Tacho unit. The Fo was read directly from the Tacho unit.

Deliyski (1990) presented an acoustic model of pathological voice production which describes the non-linear effects occurring in the acoustic wave-form of disordered voices. The noise components such as fundamental frequency and amplitude irregularities and variations, sub-harmonic components, turbulent noise and voice breaks are formally expressed as a result of random time function influence on the excitation function and the glottal filter. Quantitative evaluation of these random functions is done by computation of their statistical characteristics which can be useful in assessing voice in clinical practice. This set of parameters, which corresponds to the model, allows a multidimensional voice quality assessment. Since any single acoustic parameter is not sufficient to demonstrate the entire spectrum of vocal function or of laryngeal pathology, multidimensional analysis using multiple acoustic parameters has been attempted by some investigators. Davis (1976) used parameters such as pitch perturbation quotient, amplitude perturbation quotient, pitch amplitude, co-efficient and spectral flatness of the residue signal spectrum and performed multidimensional analysis aiding at differentiation of pathological voices from normal voices. The detection probability was 95.2% in a closed test and 67.4% in an open test.

Nataraja (1981), Studied spectrograms of hoarsness and normal voice. He found the presence of aperiodic vibration of vocal cords, noise components, variation in frequency and amplitude as contributing to the hoarsness of voice. Bhat (1981), Studied 27 dysphonics and 9 normals using spectrograms. He found that irregular vertical striations, absence of higher harmonics and presence of noise components in the higher frequency range and the variation in fundamental frequency as indicators of hoarsness.

Hirano (1989) did an international survey and has recommended the follwoing measures for clinical voice evaluation.

1. Air flow
   Phonation Quotient (PQ)
   Vocal Velocity Index (VVIO)
   Maximum Phonation Time (MPI)

2. Fo range
   SPL range

Habitual Fo

Habitual SPL

3. Electroglottography
4. Tape recording

Pitch perturbation

Amplitude perturbation

S/N ratio

LTAS

Inverse filtering

VOT

Perceptual evaluation

5. Laryngeal Mirror

Fibroscopy of larynx

Microscopy of larynx

6. X-ray Laryngography
7. Vital capacity

Ribcage and abdominal movements

8. Audiometry.

There are various objective methods to evaluate these parameters. Stroboscopic procedure, pardue, pitch meter, high speed cinematography, electroglottography, digi pitch, pitch computer, ultrasonic recordings and the high resolution signal analyser.

But at present various computer based methods are being evolved which are very fast in terms of analysisng the voice samples and giving the values of the parameters as such. Recently these methods are being used mostly in clinical and research work because they are time saving and they don't need interpretation on the part of experimenter since the parameters are automatically analysed and given.

1a. FUNDAMENTAL FREQUENCY

The fundamental frequency often loosely called the pitch of the voiced speech sounds varies considerably in the speech of a given speaker and the average or characteristic fundamental frequency varies over speakers.

The fundamental frequency of the voice is determined by vocal fold length, tension and mass in combination with subglottic pressure. Vocal fold vibration increases as vocal fold tension is increased and mass is reduced, thus producing a note of higher pitch than when the fold are relaxed, bulky and vibrating more slowly.

According to Stevens and Davis (1935), the pitch of complex tone depends upon the frequency of its dominant component, that is, the fundamental frequency in a complex tone. Plomp (1967) states that 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. Emrickson (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 are produced.

Amongst the three major attributes of voice, both quality and loudness of voice are mainly dependent upon the frequency of vibration. Hence, it seems apparent that frequency is an important parameter of voice". (Anderson, 1961).

Joseph and George (1962) studied the intralaryngeal relationships during pitch and intensity changes and found that pitch is independent of length of the vocal bands but not of the vertical position of larynx.

There are various objective methods to evaluate the fundamental frequency of the vocal cords. Stroboscopic procedure, high speed cinematography, Electroglottography, Ultrasonic recordings, stroboscopic Laminography (STROL), Cepstrum pitch detection; Digi pitch, the 3M plastiform Magnetic Tape Viewer, Spectrography, Pitch computer, the high resolution signal analyser frequency meter, visipitch, vocal II, computer with speech interface unit and software etc.

Average fundamental frequency decreases with increasing age until adulthood for both males and females. The average drop fo Fo in females is roughly 75 Hz (from about 270-300 Hz to about 200-225 Hz) from prepubescence to adulthood. For males the drop over the same period is likely to be about 150 Hz (275-300 Hz to 100-150 Hz) about 100 Hz of which may occur abruptly as a result of the adolescent voice break (Curry, 1940; Fairbanks, 1940).

Lowering in the fundamental frequency is gradual till the age of 10 years (Gopal, 1980), 15 years (Samuel, 1973), 13 years (Usha, 1978), 14 years (Rashmi, 1985) after which there is a sudden marked lowring in the fundamental frequency. The fundamental frequency values are distinguished by sex only after the age of 11 years, although small sex difference might occur before that age Kent (1976), Usha (1978), George (1973), Gopal (1980).

Gopal (1980) reported gradual a lowering of the fundamental frequency as a function of age from the age of 7 years to 17 years, for the vowel /a/ in both mals and females. The fundamentals frequency drops slightly during the first moth of life, after which it stabilizes for a period of approximately five months.

Beginning with the first year, Fo decreases sharply until about three years of age, then it makes a more gradual decline, reaching the onset of pubertay at 11 or 12 years of age. A sex difference is apparent by the age of 13 years. The decrement in Fo from infancy to adulthood among females is somewhat in excess of an octave, where as males exhibit an overall decrease approaching two octaves (Kent, 1976).

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

The study of fundamental frequency has important clinical implications. Cooper (1974) has used spectrographic analysis as a clinical tool to describe and compare the Fo and hoarseness in dysphonic patient before and after vocal rehabilitation. Jayaram (1975) found a significant difference in habitual frequency measures between normals and dysphonics. Boone (1977): estimated habitual pitch level as 128 Hz for men, 225 Hz for women and 265 Hz for children.

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 (Nataraja and Jayaram, 1982). Most of the therapies aim to alter the habitual pitch level of the patients or make the patient to use his optimum pitch (Cowan, 1936; West et al, 1957; Anderson, 1961; Van Riper and Irwin, 1966). It is therefore apparent that the measurement of the fundamental frequency of voice has important applications in both the diagnosis and treatment of voice disorders and also reflects the neuromuscular development in children (Kent, 1976).

1b. Fundamental frequency in speech

Evaluation of the Fo in phonation, may not reflect the true fo used by individuals in speech. Hence, it becomes important to evaluate the speaking fundamental frequency. The fundamental frequency in speech is estimated subjectively by matching or it is determined objectively with a pitch meter or digipitch. For more precise measurement, Fo histograms are obtained with the aid of a computer. Mean fo in speech is measured as a clinical test values (Hirano, 1981).

Lushisinger and Arnold (1965) estimated that a child's and a small womans (i.e soprano) modal fundamental frequency will be in the region of 300 Hz, whereas a bass will be in the range of 100 Hz. They give the following estimates.

1. Men : Median speaking range.
    - Bass - 98 - 110 Hz.
    - Baritone - 117 - 133 Hz
    - Tenor - 147 - 165 Hz

2. Women : Median speaking range
(One octave higher than men)

Contralto	-	220 Hz
Mezzo-soprano	-	226 Hz
Soprano	-	262 Hz

Nataraja and Jagadeesh (1984) - 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.

Table A(i) : shows the results of the study

Group		Optimum frequency	Fundamental frequency in			
			Phontion	Speaking	Reading	Singing
Male	M	127.91	141.98	166.00	192.25	211.17
	SD	8.48	29.72	33.84	43.27	60.64
Female	M	241.13	237.03	266.26	272.91	304.04
	SD	18.39	29.30	53.72	44.98	55.89

Thus the results of the above study cautions that fundamental frequency has to be measured under different conditions in the evaluation of voice disorders, i.e. it may not be enough if one considers one conditions to determine the means fundamental frequency used by the case for evaluation of voice.

Table A(ii) : Showing the Fo in males and females as reported by different investigators on Indian Population. (Based on studies conducted at AIISH)

Age group in years	Normal Fundamental Frequency in Hz	
	Males	Females
4-7	233	248
7-11	255	238
11-13	247	240
14-15	177	244
16-25	139	224
26-35	142	230
36-45	147	243
46-55	148	258
56-65	150	235

Many investigators have studied the speaking fundamental frequency as a function of age and in various pathological conditions. Bohme and Hecker (1970) indicate that the mean speaking fundamental frequency decreases with age by the end of adolescence. A marked lowering takes place during adolescence in men. In advanced age, mean fundamental frequency in speech becomes higher in men but is slightly lowered in women.

A study of the pitch level in speech in two groups of females, between 65 and 75 years and between 80 and 94 years, indicated no significant difference in the pitch level between the two groups. Therefore, speaking pitch level of women probably varies little throughout adult life.

Hollien and Shipp (1972) present data on the mean speaking fundamental frequency in 175 males ranging in age from 29-89 years. Mean frequency levels by age decode shows a progressive lowering of speaking fundamental frequency from age 20-40 with a rise in level from age 60 through the 80s. In the 20-29 years range the mean frequency is reported to be 120 Hz. This value (120 Hz) obtained for the 20-29 year olds agrees.

1. Best with the data reported by Hanley (N=27, median frequency = 128 Hz) for a population of about the same size and age.
2. Reasonably well for a larger and younger group studied by Hollen and Jackson (N = 157, Mean frequency = 128 Hz, Mean age = 21 years), and
3. Poorest (but no in conflict with studies by Provrouest (N = 6, mean frequency = 132 Hz and Hilhour (N = 24, Mean frequency = 132 Hz).

Michel, Hollin and Moore (1965), studied the speaking fundamental frequency characteristics of 15-, 16-, and 17- years old girls, in order to determine the age at which adult female speaking fundamental frequencies are established, show that females attain adult speaking fundamental frequency by 15 years of age.

Rashmi (1985) based on her study reported that there was very little change in the speaking fundamental frequency (SFF) as a function of age in males upto the age of 14 years after which a sudden decrease in SFF in the males was noticed.

Gilbert and Campbell (1980) studied the speaking fundamental frequency in three groups (4 to6 years, 8 to 10 years and 16 to 25 years) of hearing impaired individuals, and reported that the values were higher in the hearing impaired groups when compared to values reported in the literature for normally hearing individuals of th same age and sex.

Murry (1978) studying the fundamental frequency in speech of four groups of subjects, namely vocal fold paralysis, benign mass lesion, cancer of the larynx and normals noted that the parameters of mean fundamental frequency in speech failed to separate the normals from the three groups of pathologic subjects. In a parallel study, Murry and Doherty (1980) reported that along with other voice production measures such as directional and magnitudinal perturbation, the fundamental frequency in speech improved the discriminant function between normal voices and malignancy of the larynx.

Sawashima (1968) reported a rise in mean fundamental frequency in speech in cases of sulcus vocalis and a fall in mean fundamental frequency in speech in cases of polypoid vocal folds and virilism. Very high mean fundamental frequency in speech values result from disturbances of mutation in males.

Nataraja and Jagadeesh (1984) 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. Hence, fundamental frequency has to be measured under different conditions in evaluation of voice disorder ie., it may not be enough if one considers one condition to determine the mean fundamental frequency used by the case for evaluation of voice.

Thus the review of literature shows that the measurement of Fo both in phonation and speech are important in assessing the neuromuscular development and diagnosis and treatment of voice disorders. However, the present study is also considering the measurement of fundamental frequency both in phonation and speech as it would be helpful in assessing the voice disorders.

II 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 the critical prosodic feature, namely intonation (Freeman, 1982).

Variations in fundamental frequency and the extent of range used also relate to the intent of the speaker. (Fairbanks and Pronovast (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 use than serious, thoughtful speech.

Hudson and Holdbrook (1981) studied the fundamental frequency range in reading, in a group of young black adults, age ranging from 18 to 29 years. Their results indicated a mean range from 81.95 to 158.50 Hz in males and from 139.05 Hz to 266.10 Hz in females. Compared to a similar white population studied by Fitch and Holdbrook (1970), the black population had greater mean frequency ranges. Fitch's (1970) white subjects showed a greater range below the mean than above. This behaviour was reversed for the black subjects. Hudson (1981) pointed out that such patterns of vocal behaviour may be important clues which alert the listener to the speaker's racial identity.

Frequency Range in Phonation and Speech in Normals and Dysphonics (Based on Studies conducted at AIISH)

Frequency Range in (Hz)	Normal		Dysphonics	
	Mean	Range	Mean	Range
Phonation	9.00	1-29	210	117-470
Speech	295	117-427	332	121-496

Nataraja and Savithri, 1990

Sheela, (1974) found that the pitch range was significantly greater in trained singers than in untrained singers. Jayaram (1975) reported that in normal males the frequency range ranged from 90 Hz to 510 Hz; and it ranged from 30 Hz to 350 Hz in dysphonic males. The females of the normal and dysphonic groups presented 140 Hz to 710 Hz; and 60 Hz to 400 Hz as their range of frequency range respectively. He also reported that as a group, dysphonics, both males and females presented a restricted frequency range as compared to normals. Thus, the measure of frequency range gains importance in differential diagnosis of dysphonics.

Shipp and Huntington (1965) indicated that laryngitic voices had significantly smaller ranges than did post laryngitic voices. The results of a study by Murry (1978) showed a reduced semitone range of fundamental frequency in speech in patients with vocal fold paralysis, as compared to normals. Murry and Doherty (1980) reported that the variability in fundamental frequency in speech, along with directional and magnitudinal perturbation factors, enhaced the ability to discriminate between talkers with no known laryngeal vocal pathology and talkers with cancer of the larynx.

Adams (1981) discovered that stuterers and non-stuterers used a greater range of fundamental frequencies while reading at a higher than normal pitch as compared to reading in their habitual pitch. Moreover, reading in a lower than normal pitch produced less fundamental frequency variability than reading at habitual pitch levels.

Nataraj (1986) found that the frequency range did not change much with age ie, 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 in advanced age.

General conclusions about the diagnostic value of fundamental frequency variability are difficult to make because such measurements are helpful in diagnosing certain pathological conditions but not in others (Kent, 1976).

Thus review indicates, that it is important to have extensive data on the pitch variations, before it can be applied to the clinical population.

III. Pertubration measures in fundamental frequency and intensity

A number of high speed laryngoscopic motion pictures reveal that the laryngeal structures (the vocal folds) were not totally symmetric. Different amounts of mucous accumulates on the surface of the vocal folds during vibration. In addition turbulent air flow at the glottis also causes some perturbations. Limitations of laryngeal servo mechanism through the articular myotic mucoucsal relfex system (Gould and Okamura 1974; Wyke 1967) may also introduce small perturbations in laryngeal muscle tone. Even without consideration of reflex mechanism, the laryngeal muscle tone have inherent perturbation due to the time staggered activities, which exist in any voluntary muscle contractions.

Aperiodic laryngeal vibrations have been related to abnormal vocal production by various investigators (Carhart, 1938, 1941; Bowler, 1964). Vanleden et al. (1960) report that in a frame by frame analysis of ultra-high speed motion pictures "the commonest observation in pathological conditions is a strong tendency for frequent and rapid changes in the regularity of vibratory pattern".

Henson, Gerratt and Ward (1983), suggested that majority of phonatory dysfunctions are associated with abnormal and irregular vibrations of the vocal folds. These irregular vibrations lead to the generation of random acoustic energy, ie., noise, fundamental frequency and intensity variations. This random energy and aperiodicity of Fo is perceived by human ears as hoarseness. Hence, the spectral, intensity and Fo parameters are more appropriate in quantifying phonatory dysfunctions.

(i) Pitch Perturbation

Cycle to cycle variation in fundamental frequency is called pitch perturbation or jitter. Presence of small amount of perturbation in normal voice has been known (Moore, Vonleden 1958, Von Leden et al 1960).

Different algrithms of pitch perturbation that can be measured include:

1. Absolute Jitter/sec/or Jita
2. Jitter percent or Jitt
3. Pitch period perturbation quotient (%) (PPQ)
4. Smoothed pitch period perturbation Quotient (%) or SPPQ

5. Co-efficient of Fo variation (%) or vFo
6. Relative average perturbation (%) or RAP.

Wyke (1969), Sorenson, Horri and Leonard (1980) have reported the possible role of laryngeal mucosal reflex mechanism in Fo prturbation. This view of possible role of laryngeal mucosal reflex findings gets support from the studies where deprivation or reduction of afferent information from the larynx occurred by anesthesizing the laryngeal muscles. This might have reduced the laryngeal mucosal reflex (Wyke 1967, 1969) and inturn increase the Jitter size in sustained phonation. (Sorenson et al 1970).

Baer (1980) explains vocal jitter as inherent to the method of muscle excitation based on the neuromuscular model of the fundamental frequency and muscle physiology. He has tested the model using EMG from crico-thyroid muscle and voice signals, and claims neuromuscular activities as the major contributor for the occurrance of perturbation.

Heiberger and Horii (1982) stated that "the physiological interpretation of jitter in sustained phonation should probably include both physical and structural variations and myoneurological variations during phonation.

Wilcox (1978), Wilcox and Horii (1980) reported that a greater magnitude of jitter occurs with advancing age which they attributed to the reduced sensory contribution from laryngeal mechanoreceptors. However, these changes in voice with age may also be due to physical changes associated with respiratory and articulatory mechanism.

Lieberman, (1963) found that pitch perturbations in normal voice never exceeded 5msecs in the steady state portion of sustained vowels. Similar variations in fundamental periodicity of the acoustic wave form have been measured by Fairbanks (1940).

Iwata and Vonleden (1970) reported that the 95% confidence limits of pitch perturbations in normal subjects ranged from - 0.19 to +0.2 msec.

Several factors have been found to effect the values of jitter such as age, sex, vowel produced, frequency and intensities.

Sorenson and Horii, (1983) reported that normal female speakers had more jitter than normal male speakers. Robert and Baken, (1984) reported higher jitter values in males than females. They attributed this difference to Fo. As the Fo increases the percentage of jitter values decrease.

Higgins and Saxman (1989) reports higher values of frequency perturbation in males than females. Gender difference may exist not only in magnitude, but also in the variability of frequency perturbation.

Zemlin, (1962) has reported greater jitter values for /a/ than /\/ and /u/ showed lowest value. This result was supported by the studies of Wilcox (1978) and Linville and Korabic (1987). Johnson and Michel, (1969) reported greater jitter value for high vowels than low vowels in 12 ENglish vowels. Wilcox and Horii, (1980) reported that /u/ was associated with significantly smaller jitter (0.55%) than /a/ and /i/ (0.68% and 0.69% respectively. Sorensen and Horii, (1983) studied the vocal jitter during sustained phonation of /a/ /i/ and /u/ vowels. The result showed that jitter values were low for /a/ with 0.71%, high for /\/ with 0.96% and intermediate for /u/ with 0.86%.

Linville and Korabic, (1987) have found that intraspeaker variability tend to be greatest on the low vowel /a/, with less variability on high vowels lil and lul. The values of the measures of jitter are dependent upon the vowels produced during sustained phonation and also the frequency and intensity level of the phonatory sample and also the type of phonatory initiation and termination. Ramig, (1980) postulated that jitter values should increase when subjects are asked to phonate at a specific intensity, and/or as long as possible.

Lieberman, (1961, 1967) has shown that pathological voices generally have large perturbation factors than normal voices with comparable fundamental frequency and that this factor is sensitive to size and location of growth in larynx. Pitch perturbation factor was defined as the relative frequency of occurrence of perturbation larger than 0.5 msec. Koike (1973) investigated the pitch periods of voice produced by pathologic speaker, and found that discrimination between laryngeal tumor and paralysis was possible. The perturbation factors, during sustained vowels, were significant in discriminating normal talkers from those with laryngeal cancer (Murry and Doherty, 1980). Higgins and Saxman (1989) investigated within subject variation of three vocal frequency perturbation indices over multiple sessions for 15 female and 5 male young adults (pitch perturbation quotient and directional perturbation factor). Co-efficient of variation for pitch perturbation quotient and directional perturbation factor were considered indicative of temporal stability of these measures. While jitter factor and pitch perturbation quotient provided redundant information about laryngeal behaviour. Jitter factor and pitch perturbation quotient varied considerably within the individual across sessions, while directional Pertubration factor was a more temporarily stable measure.

Vieira, Mclnnes and Jack (1997) assessed the agreement between jitter obtained through acoustic and EGG measurements in pathological voices, acoustic and EGG waveforms of sustained vowels (/i/, /a/, /u/) produced by 15 dysphonic patients. The agreement between acoustic and EGG derived jitter was poorer for /\/ and /u/ than for /a/ vowels. For vowel /a/, a method of acoustic jitter estimation was proposed that combines peaks and zero crossings and resulted in increased consistency.

(ii) Intensity perturbation

Cycle to cycle variation of amplitude is called intensity perturbation or shimmer. These perturbations in amplitude can be measured using several parameters. Their are different algorithm for measurement of amplitude perturbations. Some of them are

1. Shimmer in dB/dB/ or Sh dB:
2. Shimmer percent (%) or Shim
3. Amplitude perturbation Quotient /%/ - APQ
4. Smoothed Amplitude perturbation Quotient (SAPQ) (%)
5. Co-efficient of Amplitude Variation (%) VAM

Shimmer in any given voice is dependent atleast upon the modal frequency level, the total frequency range and the SPL relative to each individual voice, Michel and Wendahl (1971) and Ramig (1980) postulate that shimmer values should increase when subjects are asked to phonate at a specific intensity and/or as long as possible.

Vowel produced and sex are the two factors affecting shimmer values as reported in the literature. Sorensen and Horii, (1983) reported that normal female speakers have less shimmer than normal male speakers. Wilcox and

Horii, (1980) reported that shimmer values are different for different vowels. Sorensen and Horii (1983) studied the vocal shimmer during sustained phonation of /a/, *IiI* and *Id* vowels. The results showed that shimmer values were lowest for *\/u/* with 0.19 dB, highest for /a/ with 0.33 dB and intermediate for *lil* with 0.23 dB. These results are supported by Horii (1980). McGlove and Hollicn (1963) report that in women 65 to 79 years the speaking pitch variability changes little. However, Stoicheff (1981) reported an increase in variability of fundamental frequency in post-menopausal adults, which was interpreted as indicating decreased laryngeal control over fundamental frequency adjustments.

Koike (1969), showed that relatively short period modulation of vowel amplitude was observed in patients with laryngeal neoplasms. He reasoned from this that the measurement and analysis of such modulation might be useful in assessing laryngeal pathology. Vonleden and Koike (1970) found a significant correlation between subjects with various laryngeal disease and different types of amplitude modulations and affirmed the potential values of short term perturbations in the acoustic signal for diagnostic purposes. THe data suggested different types of amplitude modulations which in turn correlated with clinical groupings.

Kitagima and Gould (1976) studied the vocal shimmer during sustained phonation in normal subjects and patients with laryngeal polyps and found the value of vocal shimmer to range from 0.04 dB to 0.2 dB in normals and from 0.08 dB to 3.23 in the case of vocal polyps. Although some overlap between the two groups was observed they noted that the measured value may be a useful index in screening for laryngeal disorders or for diagnosis of such disorders and differentiation between the two groups.

(iii) The fluctuation in frequency and intensity in a given phonation sample may indicate the physiological or pathological changes in the vocal mechanism. Crystal and Jackson (1970) measured both the fundamental frequency and amplitude perturbations of voice in persons with varying laryngeal conditions and concluded that several purely statistical measures of the data they extracted might be useful as guidelines in detecting laryngel dysfunction. Kim et al. (1982) have analysed the vowel */e/,* using the spectrograph, in 10 patients with recurrent laryngeal nerve paralysis and 10 normals to obtain the following acoustic parameters.

The acoustic parameters obtained from the spectrogrpahs were :

1. Extent of fundamental frequency fluctuation: The extent of fluctuation was defined as the per cent score of the ratio of the peak to peak value of fluctuation to the mean fundamental frequency (Fo).
2. Speed of fundamental frequency fluctuation : This has been defined as the number of positive peaks within 1 sec.
3. Extent of amplitude fluctuation : This has been defined as the peak to peak value in decibel measured on an average amplitude display.
4. Speed of amplitude fluctuation, which was defined as the number of positive peaks on an amplitude display within 1 sec. Peaks of 3 dB or greater from adjacent troughs have been counted.

The results of this study have indicated that among the acoustic parameters studied, significant differences were found between the control and the diseased groups in terms of fluctution of Fo.

Yoon et al. (1984) studied the voice of patients with glotic carcinomas, using the same procedure and the parameters as described by Kim et al (1982). They concluded that significant difference were found between the normals and patients with advanced carcinomas in terms of extent of fluctuation and speed of amplitude fluctuations.

Rashmi (1985) concluded that

1. Fluctuations in frequency of the initial and final segments of phontion of /a/, /i/ and /u/ showed a decreasing trend with age in males.
2. The 14 to 15 years old group shown an increase in the range of fluctuations for all the vowels.
3. In females, three was a decrease in the range of fluctuations in frequency of the initial and final segments upto the age of 9 years, an increase in the range of fluctuations in the 9-11 years old femals, which again drops down till the age of 15 years.
4. The medial segment of phonation, both males and females was quite steady, and the range of fluctuations as a function of age did not show much difference.
5. No difference in the ranges of fluctuations in frequency between males and females was observed in the younger age groups.
6. The males consistently showed greater fluctuations in frequency in the phonation of /a/, /i/ and /u/ than the females of 14-15 year old group.
7. The fluctuations in the initial and final segments of phonation for all the three vowels was greater than the fluctations in the medial segment, for both males nd females.
8. The fluctuations in intensity did not show any systematic trend for any vowels both in males and females. However, the initial segment of phonation showed a significantly larger fluctuation in intensity in the above 12 year groups, in the cases of males, for all three vowels.

Vanaja (1986), Tharmar (1991) and Suresh (1991) have reported that as the age increased there was increases in fluctuations in frequency and intensity of phonation and this difference was more marked in females. Nataraj (1986) found that speed of fluctuation in fundamental frequency and extent of fluctuation in intensity parameters were sufficient to differentiate the dysphonics from the normals.

Sridhara (1986) studied laryngeal wave forms of young normal males and females. The results are given below in the Tables a and b.

Table-A : Mean Values of Jitter (in m. sec)

	/a/	/i/	/u/
Males	0.065	0.11	0.067
Females	0.058	0.03	0.048

Table-B : Mean Values in Shimmer (in m. dB)

	/a/	/\/	/u/
Males	0.033	0.066	0.15
Females	0.07	0.37	0.44

Chandrashekar (1987) found significant difference in Jitter values in /a/ for males and /i/ and /u/ for females when compared with dysphonics. Also, the shimmer values were greater for vocal nodule cases than normals with respect to both male and female groups. But the values were significant only for males. On the whole, he found significant difference in jitter and shimmer values between normals and dysphonics.

Venkatesh et al. (1992) reported the jitter ratio (JR), relative average perturbation 3 point (RAP3), deviation from linear trend (DLT), shimmer in dB (SHIM) and amplitude perturbation quotient (APQ) to be most effective parameters in differentiating between normal males, normal females and dysphonic groups. They also state that pitch perturbation (3 point) and DLT are most useful in differentiating laryngeal disorders.

IV. Measurement of Noise

Several investigators have found high correlations between listres evaluations of roughness and the acoustic feature of spectral noise for simulated, rough vowels produced by normal speakers (Deal and Emanuel, 1978; Emanuel and Sansone, 1969; Sansone and Emanuel, 1970; Whitehead and Emanuel, 1974). Yanagihara (1967) states that in cases with a slight degree of perceived hoarseness, the noise component appears in the formant region and in severe hoarseness, additional noise over 3KHz can be noticed.

On sound spectrographic analysis Yanagihara (1967) has found that the sustained vowels perceived as hoarse has the follwoing characteristics:

1. Noise components in the main formants of various vowels.
2. High frequency noise component.
3. Loss of high frequency harmonic component.

As the degree of judged hoarseness increases more noise appears and replaces the harmonic structure. He also developed a technique for visually evaluating hoarseness based on the spectrogram. Emanuel et al., (1979) estimated noise levels in the spectra of sustained vowels and found a relationship between the spectral noise level (SNL) and th perceived magnitude of the roughness of the voice.

Yumoto, Gould and Baer (1982) developed harmonic to noise ratio (H/N) as an objective and quantitative evaluation of the degree of hoarseness. The result showed a highly significant agreement between H/N calculation and subjective evaluation of the spectrograms. H/N ratio proved useful in quantitative assessment of results of treatment of hoarseness. Yumoto et al (1982) and Yumoto (1983) determined H/N ratio directly from the voice signals. They reported significant agreement between the H/N ratio and subjective spectrographic evaluation, thereby concluding that the H/N ratio would be useful in the assessment of clinical treatments for hoarseness. They have also discussed the importance of both cycle-tocycle periodicity and the wave-form within one pitch period for the evaluation of hoarseness. Objective evaluation of normals and hoarse voices was performed considering that the hoarse voices show a prominent Fo intensity compared with harmonics in the voice spectrum. The relative harmonic intensity (Hr) obtained from a stable position of the sustained vowel /a/, is defined as the intensity of the second and higher harmonics expressed as a percentage of the total vocal intensity. 95% of the normal voices examined have relative harmonic intensity larger than the critical value of 67.2% where as 90% of the hoarse voices have relative harmonic intensity smaller than the critical value. The harmonic intensity analysis thus provides good discrimination between normal and hoarse voices.

Kitajima (1981) did a study in which he obtained a quantitative magnitude of the noise in the sustained vowel /ah/ when uttered by speakers with pathologic voice. The findings indicated that the noise ratio obtained could the used as one of the reliable acoustic parameters of the hoarse voice. Experiments with samples shows that the normalized noise energy is especially effective for detecting glottic cancer, recurrent nerve paralysis and vocal nodules. But 22.6% of patients with glottic-T_1 cancer are incorrectly classified as normal. However, normalized noise energy has been shown effective in discriminating glottic T_2 - T_4 cancer. The detectability of other laryngeal diseases can be improved by incorporating other measures such as jitter and shimmer (Kasuya et al., 1984).

Kasuya, Ogawa, Mashima and Ebihara (1986) devised an adaptive comb filtering method operating in the frequency domain to estimate noise components from a sustained vowel phonation and proposed an acoustic measure of the amount of noise in the pathologic voice signal for the purpose of applying it in the screening of laryngeal diseases by voices.

There are several computer programmes which have been developed to carry out the acoustic analysis of voice. MDVP is one such computer programme.

Mufti Dimensional Voice Programme (MDVP)

The MDVP software computers a set of 29 acoustic voice parameters in about 16 seconds and provides flexible routines for graphical representation of the results. Also a user upgradable voice data base allows automatic comparison of the current results with different nosological groups. The definitions of each parameter are provided in Appendix II.

Deliyski - In order to extract the normative threshold values of the acoustic parameters during sustained phonation of th vowel /a/, he studied 15 persons with normal voice production and 53 patients with laryngeal disease which included laryngeal cancer, benign neoplasms, chronic laryngitis, functional dysphonia and paralysis of a recurrent nerve. The computed normative threshold values for this data base are :

Frequency perturbation measurements

Jita - 83.2us
Jitt - 1.04%
RAP - 0.68%
PPQ - 0.84%
SPPQ - 1.02%
Fo - 1.10%

Amplitude perturbation measurement

ShdB - 0.35dB
Shim - 3.81%
APQ - 3.07%
SAPQ - 4.23%
AM - 8.20%

Voice break, sub-harmonic and voice irregularity measurements

DVB - 0%
DSH - 0%
Duv - 0%
NVB - 0
NSH - 0
Nuv 0

Noise and tremor evaluation measurements

NHR - 0.19
VTI - 0.061
SPI - 14.12
FTRI - 0.95%
ATRI - 3.37%

However, he has recommended that a separate database be created for the different applications because the normative values may vary depending on the nosological groups included.

Anitha (1994) studied sixty normal subjects (30 males and 30 females) and 30 dysphonics (18 males and 12 females) to examine the relationship between various parameters of voice and voice disorders. Phonation of vowels /a/, /iI, Iu/ and sentence (/alli/ /ga:di/ /ide/) was used to extract the acoustic parameters from multi dimentional voice program.

She found that -

1. There was no significant differences between the trials of vowels /a/, /iI, luI and sentence (/alli/ /ga:di/ /ide/) inturms of different parameters.
2. There was significant difference between the vowels /a/, /iI, luI and the sentence (/alli/ /ga:di/ /ide/) interms of different parameters.
3. There was significant differences between the males and females in both normals and dysphonics interms of different parameters.
4. There was significant difference between the normal males and dysphonic males and normal females and dysphonic females interms of different parameters.

She found the following parameters to be useful in differentiating normals and dysphonics with an error ranging from 13-15% and wilks lambda as low as 0.03 to 0.04. The mean values for each parameter for the phonation of /a/, /iI , luI and sentence is listed in Appendix I for both normal and dysphonics.

1. Average fundamental frequency
2. Average pitch period
3. Lowest fundamental frequency

4. Phonatory Fo range
5. Amplitude tremor frequency
6. Absolute Jitter
7. Relative Average Perturbation
8. Pitch perturbation quotient
9. Shimmer in dB
10. Shimmer percent
11. Smoothed amplitude perturbation quotient
12. Peak amplitude variation
13. Voice turbulence index
14. Soft phonation index
15. Amplitude trimor intensity index
16. Degree of voice breaks
17. Degree of sub-harmonics
18. Degree of voiceless
19. Number of sub harmonic segments
20. Number of segments computed.

Her recommendations for further study were ;

1. These parameters to be studied with different laryngeal pathologies before, during and after therapy to find out the exact effect of therapy.
2. To include other parameters (like aerodynamic parameters) and correlate with these parameters.
3. To include more number of subjects for further study.

Biran (1995) studied children in the age range of 5-15 years to examine the changes in acoustic parameters as a function of age and sex. Phonation of the vowel /a/ was used to extract parameters from multi dimensional voice program.

a) A comparison of different age groups in case of males showed that there was a consistent significant difference across the age groups in terms of the following parameters.

1. Average Fundamental Frequency (Fo)
2. Average Pitch Period (To)
3. Highest Fundamental Frequency (Fhi)
4. Lowest Fundamental Frequency (Flo)
5. Absolute Jitter (Jita)
6. Jitterpercent (Jitt)
7. Relative Average perturbation (RAP)
8. Shimmer in percent (Shim)
9. Soft Phonation Index (SPI)

b) A comparison of different age group in females showed that there was a consistent significant difference across the age groups in terms of the following parameters :

1. Average Fundamental Frequency (Fo)
2. Average Pitch Period (To)
3. Highest Fundamental Frequency (Fhi)
4. Lowest Fundamental Frequency (Flo)
5. Absolute Jitter (Jita)
6. Relative Average perturbation (RAP)
7. Pitch Period Perturbation Quotient (PPQ)
8. Shimmer in dB (ShdB)
9. Average Perturbation Quotient (APQ)
10. Smoothed Average Perturbation Quotient (sAPQ)
11. Peak Amplitude Variation (vAm)

12. Soft Phonation Index (SPI)
13. Degree of Sub-Harmonic (DSH)

c) A comparison between males and females across age groups showed that there was a significant difference across the age groups in terms of the following parameters :

1. Average Fundamental Frequency (Fo)
2. Average Pitch Period (To)
3. Highest Fundamental Frequency (Fhi)
4. Lowest Fundamental Frequency (Flo)
5. Amplitude Perturbation Quotient (APQ)

Gopal Krishna (1995) studied five normal subjects (4 males and 1 female) in the age range of 20-45 years to develop susceptibility criteria for vocal fatigue. He determined the acoustic parameter before fatigue (per fatigue condition) after half an hour and one hour of reading.

He found that the following parameters showing significant differences after 30 minutes of reading in males:

1. Fundamental frequency
2. Average pitch period
3. Highest fundamental frequency
4. Pitch pertubration quotient
5. Smoothed Average Perturbation Quotient
6. Frequency Tremor Intensity Index

In females,

1. Fundamental frequency
2. Average pitch period
3. co-efficient of amplitude variation

After one hour of reading, the following parameters showed significant difference in males,

1. Average Pitch Period
2. Highest Fundamental Frequency
3. Co-efficient of Fo Variation
4. Fundamental frequency Tremor Intensity Index

In females,

1. Fundamental Frequency
2. Average Pitch Period
3. Standard Deviation of Fundamental Frequency
4. Absolute Jitter
5. Jitter Percent
6. Relative average pertubration
7. Pitch Pertubration quotient
8. Smoothed pitch perturbation quotient
9. Co-efficient of Fo variation
10. Smoothed amplitude pertubration quotient
11. Co-efficient of amplitude variation
12. Noise to harmonic ratio
13. Voice turbulence index

Thus he concluded that Fo related parameters, short and long term frequency pertubration related measurements could be used to assess the fatiguability from multi-dimensional voice program.

The means obtained for each parameter in normals during the pre-fatigue condition is shown in Appendix I as it is relevant to the present study.

Das (1996) studied 30 male dysphonics to examine the relationship between various parameters of voice and voice disorders. The parameters were extracted using multi dimensional voice program for the phonation of /a/, /i// /u/ and sentence (/alli/ /ga:di/ /ide/. For the sake of comparison, normative data was taken from study done by Anitha (1994). The mean values for each parameter for dysphonic males are summarised in Appendix I.

He concluded that the following parameters are useful in differentiating normals from dysphonics.

1. Highest Fundamental Frequency (HFi)
2. Standard Deviation of Fundamental Frequency (STD)
3. Amplitude tremor frequency (Fatr)
4. Absolute Jitter (Jita)
5. Jitter Percent (Jitt)
6. Relative Average Pertubration (RAP)
7. Pitch Period Pertubration Quotient (PPQ)
8. Smoothed Pitch Period Pertubration Quotient (SPPQ)
9. Fundamental Frequency Variation (vFo)
10. Shimmer in dB (ShdB)

11. Amplitude Pertubration Quotient (APQ)
12. Smoothed Amplitude Pertubration Quotient (sAPQ)
13. Peak amplitude variation (vAm)
14. Soft Phonation Index (SPI)
15. Frequency tremor Intensity Index (FTRI)
16. Degree of voice breaks (DVB)
17. Degree of sub-harmonic breaks (DSH)
18. Degree of voiceless (DUV)
19. Number of sub-harmonic segments (NSH)
20. Number of unvoiced segments (NUV).

His recommendations for future study was -

1. Examining more number of subjects
2. To examine these parameters with different laryngeal pathologies, before, during and after therapy to find out the exact effect of therapy.

Aparna (1996) examined twelve hearing impaired children (6 males and 6 females) in the age range of 5-9 years. Twelve subjects having normal hearing formed the control group. The study aimed at examining the various parameters of voice of hearing impaired children using multi dimensional analysis of voice. Phonation of vowels /a/, /i/ and /u/ was considered for the purpose of analysis.

Significant difference was noted between normals and hearing impaired for the following parameters, in both males and females.

1. Mean fundamental frequency - normals showing lower fundamental frequency than the hearing impaired group.
2. Highest fundamental frequency - The hearing impaired showed larger maximum fundamental than the normal group.

3. The frequency range - greater means in hearing impaired than normals.
4. Amplitude pertubration quotient.

Parameters which showed significant difference between normals and hearing impaired, with respect to females only were ;

1. Lowest fundamental frequency
2. Speed of fluctuations
3. The extent of fluctuations in intensity
4. Jitter ratio

Parameters which showed significant difference between males and females, with respect to both hearing impaired and normals were :

1. Directional pertubration factor
2. Relative amplitude pertubration
3. Shimmer
4. Directional pertubration of amplitude
5. Amplitude pertubration quotient

Hence the present study was undertaken to

1. To check the stability of the various acoustic parameters that are extracted from multi dimensional voice program.
2. To support the findings of earlier studies and strengthen its validity.
3. To device an objective method to differentiate normals from dysphonics.

After the measurement of parameters, the clinician has to make a judgement regarding the diagnosis based on the values of parameters. Since the computers are widely used and attempts have been made to use them for clasifying the disorders. An artificial Network is a computer branch which have been widely applied for such activities.

Artificial Neural Networks

Artificial Neural Networks (ANNs) commonly referred to as Neural Networks (NNs) ar also termed as neuro-computers, connectionist networks, parallel distributed processors, etc. In the most general form, a neural network is a machine designed to model, the way in which the brain performs a particular task or function of interest by using electronic components or simulated in software on a digital computer.

ANN is a massively parallel distributed processor that has a natural propensity for storing experimental knowledge and making it available for use. It resembles the brain in two respects : (1) Knowledge is acquired by the network through a learning process, (2) Interneuron connection strength, known as weights are used to store knowledge. In other words, ANNs are "biologically" inspired networks having the apparent ability to initiate the brains activity to make decisions and draw conclusions when presented with complex and noisy information (Haykins, 1994).

Characteristics of Neural Networks (NNs)

NN derives its computing power through its massively parallel distributed structure and its ability to learn and therefore to generalise. Generalization refers to the NN producing reasonable outputs for input not encountered during training (learning). These two information processing capacities make it possible for NNs to solve complex (large scale) problems that are currently interactble. A complex problem of interest is decomposed into a number of relatively simple tasks and NNs are assigned subsets of tasks (eg. pattern recognition, associative memory, control) that match their inherent capabilities (Haykin, 1995).

The use of NNs offers the following useful properties and capabilities (Zurda, 1992).

1. Non-linearity

A neuron is basically a nonlinear device. Consequently, a NN, made up of interconnections of neurons, is itself nonlinear and is distributed throughout the network. Nonlinearity is a highly important property, particularly if the underlying physical mechanism responsible for the generation of an input signal (eg. speech signal) is inherently nonlinear.

2. Input-output mapping

A popular paradigm of learning called supervised learning involves the modification of the synaptic weight of a NN by applying a set of labelled training samples or task samples. Each example consists of a unique input signal and the corresponding desired response. The network is presented an example picked at random from the set, and the synaptic weights (free parameters) of the network are modified so as to minimize the differences between the desired response and the actual response of the network produced by the signal in accordance with an appropriate statistical criterion. The training of the network is repeated for many examples in the set until the network reaches steady state, where there are no significant changes in the synaptic weights. The previously applied training examples may be reapplied during the training session but in a different order. Thus, the network learns for the examples by constructing an input-output maping for the problem at hand like the supervised learning paradigm which suggests a close analogy between the input-output mapping performed by a NN and non-parametric statistical inference.

3. Adaptivity

NNs have built in capacity to adopt their synaptic weights to changes in the surrounding environment particularly, a NN trained to operate in a specific environment can be easily retrained to deal with minor changes in the operating environmental conditions. When operating in a non-stationary environment (i.e. one whose statistics change with time), a NN can be designed to change its synaptic weights in rule time. The natural architecture of a NN for pattern classification, signal processing and control applications coupled with adaptive capability make it an ideal tool for adaptive pattern classification, adaptive signal processing and adaptive control.

4. Evidential response

In the context of pattern classification, a NN can be designed to provide information not only about which particular pattern to select, but also about the confidence in the decision made. Thus information may be used to reject ambiguous patterns if arised and thereby improve classification performance of the network.

5. Contextual information

Knowledge is represented by the very structure and activation state of the NN. Every neuron in the network is potentially affected by the global activity of all other neurons in the network. Consequently, contextual information is dealt with naturally by a NN.

6. Fault tolerance

A NN implemented in hardware form, has the potential to be inherently tolerant in the sense that its performance is degraded under adverse operating conditions (Bolt, 1992). That is, if a neuron or its connection links are damaged, recall of stored pattern is impaired in quality. However owing to the distributed nature of the information in the network, the damage has to be extensive before the overall response of the network is degraded seriously.

7. VLSI implementability

The massively parallel nature of a NN makes it potentially fast for the computation of certain tasks. This feature makes it ideally suited for implementation using very-large scale integrated (VLSI) technology. This provides a means of capturing truly complex behaviour in a highly hierarchical fashion, which makes it possible to use a NN as a tool for real time applications involving pattern recognition signal processing and control.

8. Uniformity of analysis and design

Basically, NNs enjoy universality as information processors. That is, the same notation is used in all the domains involving the application of NNs, in terms of sharing theories and learning algorithms, etc.

9. Neurobiological analogy

The design of a NN is motivated by analogy with the brain, which is a living proof that fault tolerant parallel processing is not only physically possible but also fast and powerful. Neurobiologists look to ANNs as a research tool for the interpretation of neurobiological phenomena. For eg: NNs have been used to provide insight on the development of promotor circuits in the oculo-motor system (responsible for eye movements) and the manner in which they process the signals (Robinson, 1992).

Structure and function of ANN

The typical structure of an ANN is given in figure 1 below. ANNs are basically three layer system consisting of :

1. input layer of nodes for data entering the network.
2. hidden layer between the input and output layers.
3. output layer of nodes that produce the networks output responses.

Typical structure of ANN

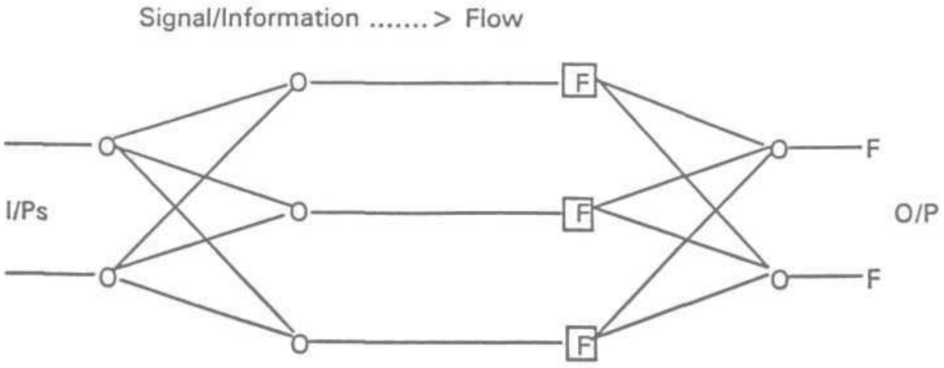


Figure 1

Mecullock and Pitts, regarded as pioneers in the field of NNs outlined the first model of an elementary neuron in 1943 (Roal and Mankame, 1996). However they could not realise the model using the bulky vacuum tubes of the era. A learning algorithm scheme for updating a neurons connections (weights) was proposed by Donald Hebb in 1949. A new and powerful learning law called widrow-Hoff learning rule was developed by Bernard Widrow and Mavrian Hoff in 1960.

The ANN structure is trained with known samples of data, the known pattern or information in the form of digital signals. Then the ANN is ready to recognise a similar pattern when it is presented to the network. The nonlinear characteristics of ANN are due to the non-linear activation function 'F'. This is useful for accurate modeling of non-linear systems. That is, in the non-linear systems the output variables do not depend on the input variables in a linear manner. The dynamic characteristics of the system itself would depend on either one or more of the amplitude of the input signal, its waveform or its frequency.

The structure of an ANN is a set of processing units (nodes) arranged in rows. Inputs nodes are interconnected by simple connections with an internal layer of hidden nodes and a single output mode. Rather than having a fixed algorithmic approach to a classification problems, an ANN is sequentially presented with a set of supervised training cases input data passes with correct output. The ANN modified its behaviour in this process of training by adjusting the strengths or weights of the connections until information "learned" by the ANN is stored in the weight of the network gives to the connections between nodes. Thus, ANNs are designed to realise very specific computational tasks/problems by the highly inter-connected, parallel computational structures with many and relatively simple elements.

Neural net and traditional classifiers

Block diagrams of traditional and neural net classifiers are given in the figure as shown -

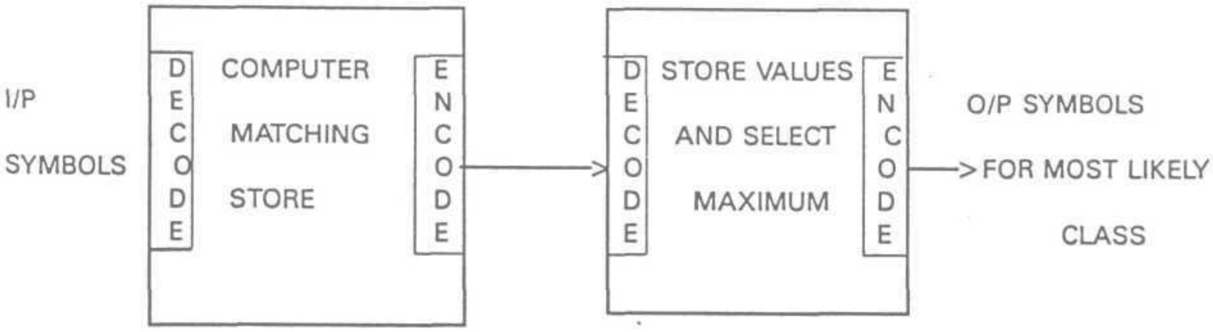


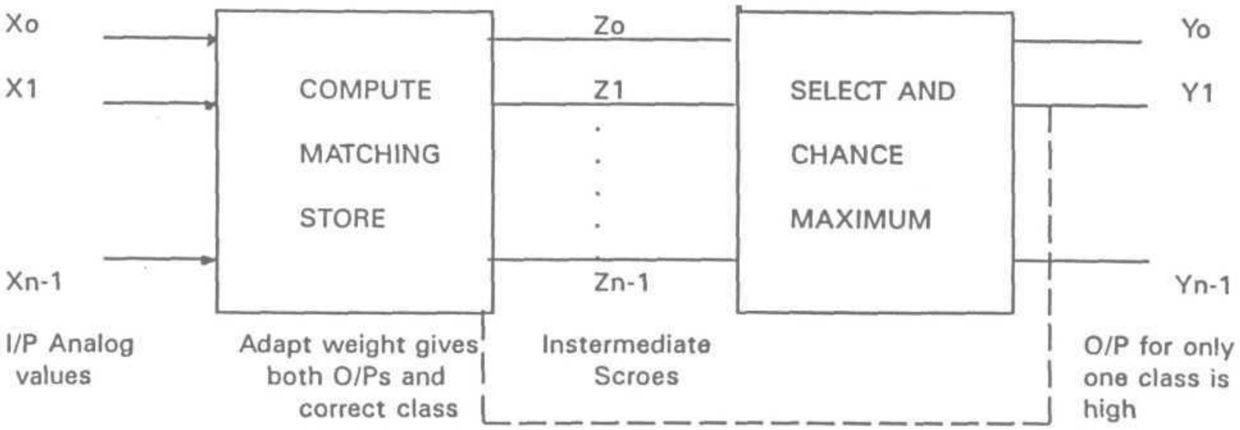
Figure 2

TRADITIONAL CLASSIFIER

Both traditional and NN classifiers determine which of M classes is most representative of an unknown static input pattern containing N input elements. In a speech recogniser the input might be the output envelope values from a filter bank spectral analyzer sampled at one time instant and the classes might represent different vowels.

Inputs and outputs of the traditional classifier are passed serially and internal computations are performed sequentially. Parameters are typically estimated from training data and then held constant.

In the adaptive NN classifier input values are fed in parallel to the first stage via N input connections. Each connection carries an analog value which may take on two levels for binary inputs or may vary over a large range for continuous valued inputs. The first stage computes matching scores and outputs these scores in parallel to the next stage over M analog output lines. Here the maximum of these values is selected and enhanced. The second stage has one output for each M classes. After classification is



NUERAL NET CLASSIFIER

FIG-3

complete, only that output corresponding to the most likely values will be on strongly or 'high'; other inputs will be 'low'. In this design outputs exist for every class and this multiplicity of outputs must be preserved for further processing stages as long as the classes are considered distinct. In the simplest classification system these output lines might go directly to lights with labels that specify class identities. In most complicated cases they may go to further stages of processing where inputs from other modalities or temporal dependencies are taken into consideration of the correct class is provided, then this information and classifier outputs can be fed back to the first stage of the classifier to adapt weights using a learning algorithm as shown in figure 3. Adaptation will make a correct response more likely for succeeding input patterns that are similar to the current pattern.

The parallel inputs required by the neural net classifiers suggest that real time hardware implementations should include special purpose pre-processors. One strategy for designing such processors is to build physiologically based pre-processors modelled after human sensory systems. Pre-processor filter banks for speech recognition that are crude analogs of the cochlea have been constructed (Martin, 1970). More recent physiologically based processor algorithms for speech recognition attempt to provide information similar to that available on the auditory nerve (Delgutte, 1984). Many of these algorithms include filter bank spectral analysis, automatic gain control, and processing which uses timing or synchrony information in addition to information from smoothed filtered output envelope.

Classifiers shown in the figures 2 and 3 perform three different tasks:

1. They can identify which class best represents an input pattern, where it is assumed that inputs have been corrupted by noise or some other

process. This is a classical decision theory problems.

2. They can be used as a content addressable or associative memory, where the class example is designed and the input pattern is used to determine which exempler to produce. This is useful when only part of the information is available and the complete pattern is required as in bibliographic retrieval of journal references from practical information.
3. They can perform vector quantize (Makhoul et al., 1985) or cluster the N inputs into M clusters. Vecotr quantizers are used in image and speech transmission systems to reduce the number of bits to transmit the analog data. In speech and image recognition applications they are used to compress the amount of data that must be processed without loosing important information on either application, the number of clusters can be pre-specified or may be allowed to grow upto a limit determined by the number of nodes available in the first stage.

Multi-Layer Perceptron

These are 'feed forward' nets with one or more layers of nodes between the input and output nodes. These additional layers contain hidden units or nodes that are not directly connected to both input or output nodes. They overcome many of the limitations of single layer perceptions and are more effective with the recent development of new training algorithms and have been successful for many problems of interest.

The capabilities of MLPS stem from the non-linearities used within the nodes. It has been demonstrated that no more than three layers are required in perception like 'feed-forward' nets because three layer net can generate arbitrarily complex decision regions (Lippman, 1987). The number of nodes in the second layer must be greater than one when decision regions are disconnected or meshed and cannot be formed from one convex area. The number of second layer nodes required in the worst case is equal to the number of disconnected regions in input distributions. The number of nodes in the first layer must typically be sufficient to provide three or more edges for each convex area generated by every second layer node. There should thus be more than three times as many nodes in the second as in the first layer. This is centered primarily on multilayer perceptions with multiple output nodes when sigmodial non-linearities are used and the decision rule is to select the class corresponding to the output node with the largest output. The behaviour of these nets is more complex because decision regions are typically bounded by smooth curves instead of by straight line segments and analysis is thus more difficult.

Uses of ANNs

Some of the uses of ANNs (Raol and Mankame, 1996) include :

- Information storage/recall : The recall is a process of decoding the previously stored information by the net work.
- Pattern recognition/classification : To recognise a specific pattern from a cluster of data/to classify sets of data of information.
- Non-linear mapping between high dimensional spaces (mathematical modelling of non-linear behaviour of systems).
- Time-series prediction (like weather forecasting), modelling of non-linear aerodynamic phenomena, detection of faults/failures in systems like power plants and aircraft sensors.

Metz, Schiavetti and Knight (1992) used an Artificial Neural in estimating the speech intelligibility of hearing impaired subjects from known acoustic variables.

Leinonen, Kangas, Torkkola and Juvas (1992) examined the vowel / a:/ in a test word (finnish) which was judged as normal or dysphonic by two speech language pathologists, with the self-organizing map, the artificial neural network algorithm of Kohenen. The algorithm produced a two-dimensional representation (maps) of speech. Input to the acostic maps consisted of 15 -component spectral vectors calculated at 9.83 msec intervals from short-time power spectra. They found that the dysphonic voices deviated from normals both in composition of the short-time power spectra (characterised by the dislocation of the trajectory pattern on the map) and in the stability of the spectrum during the performance (characterised by the pattern on trajectory on the map). They concluded that this method is suitable not only for diagnosis but also for therapeutic purposes.

Mujunen, Leinonen, Kangas and Torkkola (1993a) studied word initial samples of fricative /s/ preceding vowels /a:/, */æ/, /e:/,* /i:*/,* /u:/, /o:/ and /y:/ in finnish words with a self organizing map, the neural network algorithm of koheenen. The /s*/* samples were drawn from 10 women, aged 20-45 years. The subjects were selected on the basis of having normal */s/* articulation. Fifteen component input vectors, which constituted the input to the acoustic map was calculated from short-time FFT (Frequency Fourier Transform) spectra at 9.83 msec intervals. In all the 10 subjects the /s/ samples preceding the round vowels /u:/ and /o:/ clearly differed from the samples in front of unrounded /a:/, /æ/, */e:/* and */i/.* There were no significant differences between the locations of */s/* preceeding /a/, /u/ and /y/ or between those preceding /a/, /e*/,* /i*/* and */æ/.*

Mujunen, Leinonen, Kangas and Torkkola (1993b) further studied whether the self organizing map also distinguishes between misarticulated and normal /s/. They collected speech samples from 11 women aged 16-38 years, who had misarticulation for the /s/ sound as examined by the speech therapist. A psychoacoustic evaluation was also done where the recorded speech samples of the subjects were classified as acceptable or unacceptable by 2 speech pathologists. The results showed that the map patterns of the most deviant /s/ samples differed from those of the normal ones and those judged acceptable by most listeners. The degree of audible acceptability correlated well with the location of the sample on the map. They concluded that the self organised maps are suitable for the extraction and measurement of acoustic features underlying psychoacoustic classifications, and for on-line visual imaging of speech.

Farley (1994) created an artificial neural network model which was supposed to correspond to one component of fundamental frequency (Fo) control by the brain. Good Fo control could be achieved using only seven neurodes. These included three motor nodes, a single inhibitory neurode influencing only the thyroarytenoid (TA) motor neurode, and three exitatory neurodes, one of which excited all motor neurodes and two of which excited the inhibitory neurode. The potential utility of this type of model in the study of mechanisms of vocal control was discussed.

Leinonen, Hihunen, Laakso, Rikhanen and Poppius (1997) obtained a perceptual reference for acoustic feature selection. 94 male and 124 female voices were categorised using the rating of 6 clincicians of visual analog scales for pathology, roughness, breathiness, strain, asthenia and pitch. Partial correlations showed that breathiness and roiughness were the main determinants of pathology. The 6-dimensional ratings (the 6 median scores for each voice) were categorized with the aid of the sammon map and the self-organizing map. The 5 categories created differed with respect to the breathiness or roughness ratio and the degree of pathology.

Sunil, Murthy, Venkatesh and Vijaya (1995) devised a study in the Indian population which attempted to evaluate the automatic recognition of electroglottographic (EGG) patterns for differential diagnosis of 128 laryngeal disorders using Artificial Neural Network. A comparison was also made between identification abilities of automatic pattern recognition methods, trained speech language pathologists, and student clinicians. Results indicate that automatic recognition of EGG patterns is possible with a recognition rate of 63.3%. The recognition rates of the trained speech pathologists and student clinicians were found to be poorer.

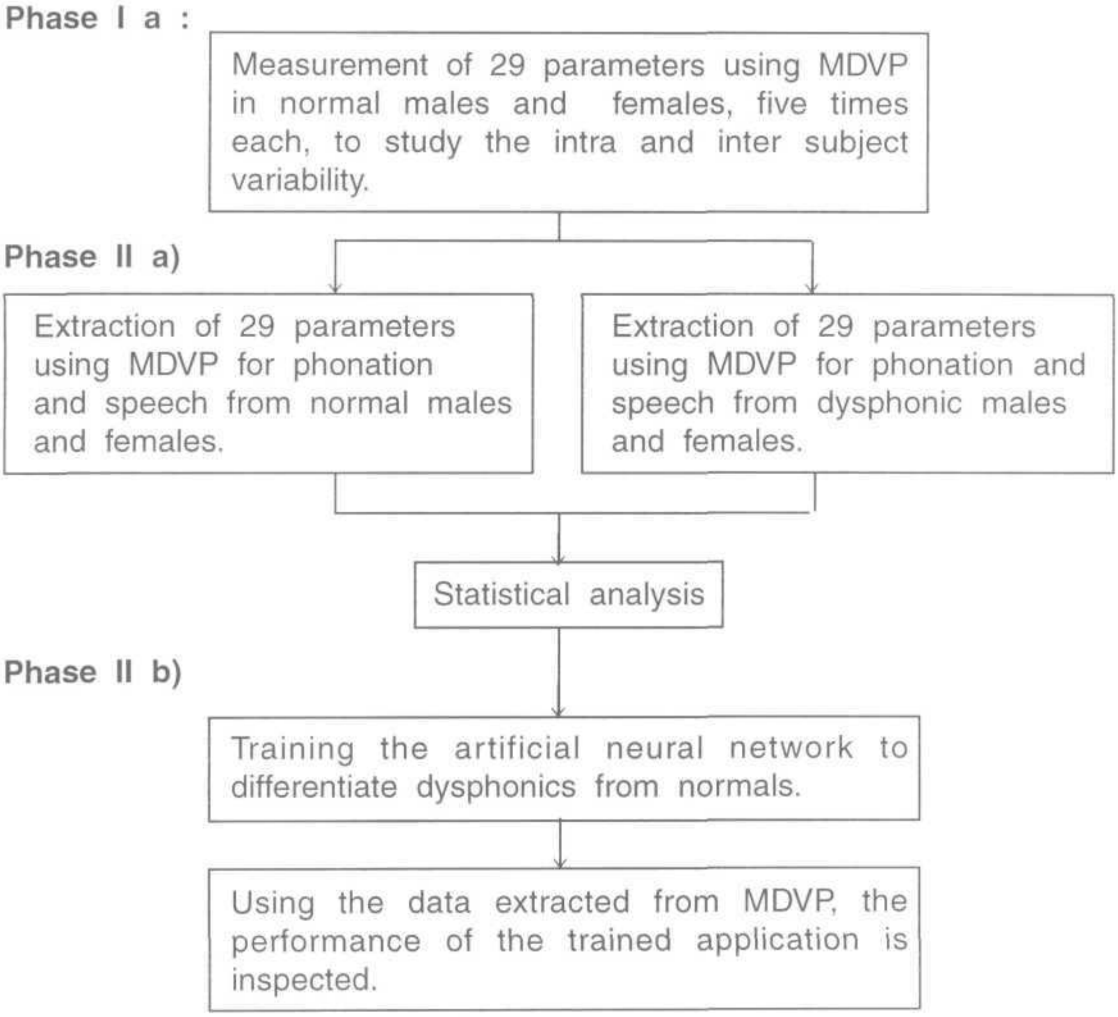
Binu (1998) - using artificial neural network, has tried to objectively classify voice disorders in terms of the acoustic and aerodynamic parameters. The aerodynamic parameters included mean air flow rate and phonation duration. The acoustic parameters included in the study were fundamental frequency in phonation, extent of fluctuations in frequency, speed of fluctuations in frequency, extent of fluctuations in intensity, speed of fluctuations in intensity, fundamental frequency range in phonation and fundamental frequency range in speech. The voice pathologies included were mass on the vocal folds, glottic chink and congestion. He was able to train the artificial neural network using ten aerodynamic and acoustic parameters to classify normals and dysphonics. Using the same parameters ANN was also able to differentiate four types of dysphonics and differentiate dysphonics from normals at 88.1% accuracy.

Thus from the review of literature it can be seen that ANN can be effectively used in paramedical fields in identification and differential diagnosis. The present study was aimed to train the ANN to differentiate dysphonics from normals using the acoustic parameters extracted using "multi dimensional voice program" software program.

METHODOLOGY

The purpose of the present investigation is to differentiate dysphonics from normals using the acoustic parameters extracted using multi dimensional voice program (MDVP) developed by Kay elemetrics Inc., N.J.

To achieve this main objective, it was planned to carry out the study as follows :



Phase I a :

Measurement of 29 parameters using MDVP in normal males and females, five times each, to study the intra and inter subject variability.

Phase II a)

Extraction of 29 parameters using MDVP for phonation and speech from normal males and females.

Extraction of 29 parameters using MDVP for phonation and speech from dysphonic males and females.

Statistical analysis

Phase II b)

Training the artificial neural network to differentiate dysphonics from normals.

Using the data extracted from MDVP, the performance of the trained application is inspected.

Phase Ia :

1. Subjects : 5 normal males and five normal females were taken and assessed in five sessions. Each trial consisted of phonation of vowels /a/, /iI, Iu/ and the sentence /idu pa:pu/ /idu ko:ṭi/ /idu kempu baṇṇa/, three times each.

2. Parameters measured :

The acoustic paratemeters extracted from MDVP are :

1. Average fundamental frequency (Fo)
2. Highest fundamental frquency (Fhi)
3. Lowest fundamental frequency (Flo)
4. Standard deviation of Fo (STD)
5. Phonatory Fo-range in semitones (PFR)
6. Fo Tremor frequency (Fftr)
7. Amplitude Tremor frequency (Fatr)
8. Absolute Jitter (Jita)
9. Jitter percentage (Jitt)
10. Relative average perturbation (RAP)
11. Pitch perturbation quotient (PPQ)
12. Smoothened pitch perturbation quotient (SPPQ)
13. Fundamental frequency variation (vFo)
14. Shimmer in dB (ShdB)
15. Shimmer in percentage (Shim)
16. Amplitude perturbation quotient (APQ)
17. Smoothened amplitude perturbation quotient (SAPQ)
18. Peak - amplitude variation (vAM)
19. Noise to harmonic ratio (NHR)

20. Voice turbulence index (VTI)
21. Soft phonation index (SPI)
22. Fo-tremor intensity index (FTRI)
23. Amplitude - tremor intensity index (ATRI)
24. Degree of voice breaks (DVB)
25. Degree of sub-harmonics (DSH)
26. Degree of vioceless (DUV)
27. No. of voice breaks (NVB)
28. Number of sub-harmonic segments (NSH)
29. Number of unvoiced segments (NUV)

3. Instrumentation :

The following instruments were used for acoustical analysis of voice.

1. C.S.L. Speech interface unit-model 4300B.
2. 486 Sx with C.S.L 50 Hardware card.
3. MDVP Software.

These were arranged as shown in Fig. 1a. (All the above three instruments are supplied by kay elemetrics Inc., New Jersey).

4. Digital audio tape deck (Sony DTC- 59ES)
5. Audio Mixer (Sony MU X051)
6. Digital Audio Tape - 3M, 4mm data tape, DDS-90.
7. Microphone (Unidirectional, sony)
8. Connecting Jack.

To connect the output of the digital audio tape deck to the input of the speech interface unit.

Instrumental setup :

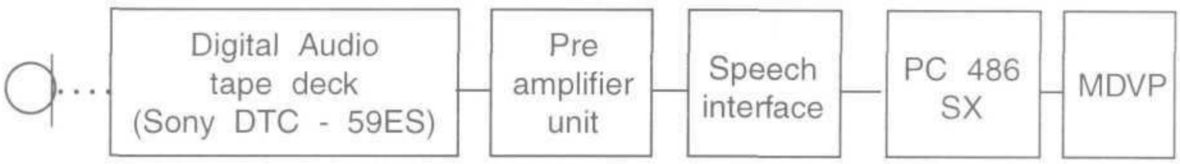


FIGURE 1.a.

4. Instructions : The subjects were made to sit comfortably and after taking a deep breath, to phonate /a/, /i/ and /u/ into the microphone. The subjects were then asked to say [/idu pa:pu/ /idu ko:ṭi/ /idu kempu baṇṇa/] in their normal speaking rate. Three trails of each of these are considered.

5. Recording : Voice samples were recorded using a digital audio tape deck (Sony DTC - 59ES). The microphone was kept at a distance of 4-6 inches during recording. All the recordings were carried out in a sound treated room of the speech science laboratory of the Department of speech sciences, AIISH.

6. Analysis :Three trails for phonation of each vowels and sentence was considered.

DATA ACQUISITION AND ANALYSIS FLOW CHART

Select Program.
MS-DOS prompt

↓

Enter CD\CSL 50

↓

Enter MDVP

↓

Connect output of digital
Audio tape deck to the input of
CSL speech interface unit

↓

Set for sustained phonation or set for
running speech in the drop down menu
from the main menu capture

↓

The subject is asked to phonate/speak
as the need may be

↓

To capture the signal select "capture channel F1",
in the drop down menu (or by pressing F1)

↓

Subjects phonates as long as he
can or utters the sentence

↓

After the desired signal is captured
(recorded), press the space bar to
stop recording

↓

Press the (F1) key for Analysis of the
captured signal.

↓

After analysis is over
press the (F8) key for the display of the
values of parameters and (F9) key for
graphical display (these functions can also
be selected from the drop down menu)

Thus the data values for ech vowel and utterance of sentence were obtained for each subjects of both groups. The data was subjected to statistical analysis.

Phase II a :

1. Subjects :

30 males and 30 females in the age range of 17-25 years. The subjects of this group had no apparent speech, hearing or E.N.T problem and were considered normal.

The second group consisted of dysphonics in the age range of 20-45 years who visited the All India Institute of speech and hearing, Mysore, with the complaint of voice problem and formed the experimental group. These subjects were diagnosed as cases of voice disorders after the routine speech, ENT, psychological and audiological evaluation.

2. Parameter extracted : Same as in phase Ia.

3. Instructions : Same as in phase Ia.

4. Recording : Same as in phase Ia.

5. Analysis : Same as in phase Ia.

For the control group, three trails were taken for all three vowels /a/, /i/ and /u/ and for the sentence /idu pa:pu/ /idu ko:t̪i/ /idu kempu baṇṇa/.

For the experimental group, one trial of the phonation of vowels /a/, /i/ and /u/ and the sentence /idu pa:pu//idu ko:t̪i idu//kempu baṇṇa / is considered.

The data was then subjected to statistical analysis to compare the subgroups of normal males and females and dysphonic males and females, with each other and between the groups.

Thus 29 parameters were obtained for each phonation of each subject of both the groups.

Phase IIb :

Phase IIb of the study was carried out in the "Neural connection" software.

Instruments :

The software was loaded on :

1. An IBM compatible PC with a pentium 200 MHz processor
2. Microsoft Windows 95
3. (16 MB) RAM
4. A hard disc with 2GB space
5. A SVGA Monitor, with appropriate graphics card
6. A mouse.

Procedure :

The following procedure as described in the mannual was followed.

The graphic user interface in the software provided the workspace, the program window where problem solving applications could be built.

The neural connection consisted of the following tool menus :

1. Data input - Imports data from files or cutting and pasting from their windows applications.

2. Outputting tools.

a. Data output-exports data to files, displays results as text, and shows the success rate.

b. Text output- as the name suggests, it displays results as text and shows the success rate.

c. Graphics output - displays a graph showing how the output from an application varies when two of the inputs vary.

d. What if - Displays a plot of how the output varies when two of the inputs vary and gives a description of how the change in one input affects the other.

e. Time series plot - Displays the results of time series prediction against time.

3. Tools for modelling and forecasting

a. Multilayer perception - A neural network modelling tool that was optimised for prediction applications.

b. Radial Basis function - A neural network tool that can be optimized for prediction and classification applications.

c. Bayesian Network - A neural network connection that can be optimized for prediction and classification applications.

d. Koheren Network - A neural network connection that can be optimized for clustering applications.

e. Closest class means classifier - A statistical modelling tool that could be used for classification application.

f. Regression - A statistical modelling tool that could be used for prediction problems.

g. Principal component analysis - A statistical modelling tool that could be used to reduce complexity of an application.

4. Tools for filtering the data

a. Filter - A tool that allows to filter the data, apply mathematical functions, and examine the statistical distribution of the data.
b. Network combiner - A tool that allows to join two or more other tools together.
c. Simulator - A tool that produces the data format needed by the "Graphics output" and "What if" tools.
d. Time series windows - A tool that allows to manipulate time series data in order to create windows before modelling takes place.

The neural network option choosen for the present study was "Multi layer perception" (MLP). The following tools were used in the following order:

1. Data input
2. Filter
3. Multi layer perception
4. Text output
5. Data output

All the tools could be moved into the work space by using the mouse. Initially the data input tool was brought to the work space and in order to input the training data into the tool the following steps were carried out :

1. Click the data input tool once. The tool menu appears.
2. From the data input toolchoose "view". A blank spread sheet appears. This box allows you to specify the developmental data file and run data file, which are loaded separately. The developmental data file will be used to train and test the application. The run data file will be used when the model is implemented.
3. In the developmental input group, click "flat file".

4. Click "Configure". A file open window appears.
5. The file name is typed, the present data is been stored in the SPSS data format. The training data used for the study was recorded in the file format "neural net. sav"
6. After selecting the file, click "Ok"
7. After this the spread sheet appears.

In the spread sheet each columns which contain different input variables are called fields. There are 30 input fields which includes sex and the acoustic parameters extracted using multidimensional voice profile. The 31st filed is the target field which contain the diagnosis of different voice disorders.

For the sake of convenience the output data which consists of the diagnosis had been given symbolic representations as given below :

a = normal
b = dysphonics

When the data is entered inside the data input tool a major portion of it is allocated for training the data and a smaller percentage of the data is used for validating and testing the data. Hence we have

* Training data - used to train the application.
* Validation data - used to monitor the network performance during training.
* Test data : used to measure the performance of a trained application.

For training the limits are :

* A maximum of 750 fields
* A maximum of 15,000 records.

Out of a total of data used in the study, datas were used for training, for validation and for testing.

When the spread sheet occurs different codes appear in the second column from the left, they are :

T = training data
X = Test data
V = Validation data
R = Run data

Also various codes appear on top rows of the screen:

I = input fields
T = Target fields
R = Reference fields, they are pased thorugh the application to be displayed on the output but not used or changed.
* ***Unused fields.*** These are not passed through the application.

The filter is important to preprocess the data when using a modelling and forecasting tool. Data fields can be switched on or off by the filter. When they are selected the filter passes them to successor or objects when requested. When they are not selected, the data are unavailable to successors objects.

The filter also helps in the weighting the data. The data can be weighed in several ways to maximize their usefulness to the application. These weightings cannot be used with symbolic data, nor can logarithms or natural logarithms be used with integer data.

Also each field can be analysed by the filter. The filter shows a histogram and several statistical measures such as mean, median, mode, variance, standard deviation, skewness, kurtosis and quantiles. If a particular field is badly skewed appropriate weighting function can be applied in the filter to process the data.

The data for the study was analysed in the filter to see the characteristics of the input acoustic variables. The histograms for individual data fields was also obtained.

"Multi layer perception" is the neural network model which is being used for the present study. To train the multi layer perception :

- From the multi layer perception menu, choose train.

The multi layer perception (MLP) is trained by an incremental learning technique where by it is trained in stages. In the first stage a sample of examples from the training set is used to train the multi layer perception. The best network in this stage is then passed to the 2nd stage and is used as a starting point for training. In 2nd stage, a larger sample of data is used to train the network, and best network is passed on to the next stage. This procedure continues for four stages.

The number of samples trained in each stages are

Stage I - 100
Stage II - 100
Stage III - 100
Stage VI - 10,000

The training automatically stops when the MLP finds a global minimal error which is equivalent to 0.001 for the training and validation data or it stops when the number of sample in the 4th stage comes to 10,000. Training can also be stopped by clicking "STOP" in the MLP performance window.

To improve the results of the MLP performance the following changes can be carried out.

- The number of hidden layers can be changed - presently MLP in the "Neural connection" software has 2 hidden layers'.
- The number of nodes in each hidden layer can be changed. Increasing the number of nodes in the hidden layers beyond a certain extent decreases the ability of the neural network to generalize based on the model created.
- The transfer or activation function used by the nodes. The activation function are of 3 types
  - Tanh function
  - Segmoid function
  - Linear function
- The learning algorithm used by the MLP-These are of 2 types
  - Conjugate gradient
  - Steepest descent

Data output Data :

Helps us to view the trained data and examine the results. The data sets are identified by coloured codes in the left hand column on the screen.

T = Training data
V = Validation data, in bright green
X = Test data, in yellow
R = Run data, in pale blue.

To view the results in the data output tool, the data output window is opened. Select "view" and from the view menu chose the training, validation, test or run data to be viewed.

3. Select "flat file" from the run data
4. Press the "Configure" button

After the run data is set up in the data output tool:

1. From the data out put menu, choose "view"
2. Select "output run data" Neural connection asks to confirm whether the specific run data file has to be outputted.

The results appeared in the spread sheet in the last column named target "T.

RESULTS AND DISCUSSION

The purpose of the present study was to differentiate between normal voice and abnormal voice using the acoustic parameters extracted from multi dimensional voice program developed by kay Elemetrics Inc., N.J. (MDVP)

The following acoustic parameters were extracted using MDVP.

I Frequency parameters

1. Average fundamental frequency
2. Highest fundamental frequency
3. Lowest fundamental frequency
4. Standard deviation of Fo
5. Phonatory Fo range in semitones
6. Fo tremor frequency
7. Absolute Jitter
8. Jitter percent
9. Relative perturbation quotient
10. Pitch perturbation quotient
11. Smoothened pitch perturbation quotient
12. Fundamental frequency variation
13. Fo- Tremor intensity index (FTRI)

II INTENSITY PARAMETERS

14. Amplitude tremor frequency
15. Shimmer in dB
16. Shimmer percent

17. Amplitude perturbation quotient
18. Smoothened Amplitude perturbation quotient
19. Peak amplitude variation
20. Amplitude tremor intensity index.

III OTHER PARAMETERS

21. Noise to harmonic ratio
22. Voice turbulence index
23. Soft phonation index
24. Degree of voice breaks
25. Degree of sub harmonics
26. Degree of voiceless
27. Number of voice breaks
28. Number of sub-harmonic segments
29. Number of unvoiced segments.

The study also intended to check whether these acoustic parameters varied when they were measured in five different sessions.

The results of different parameters are discussed here after analysing them using appropriate statistical tests. Average of threee vowels and sentence has been considered.

Average fundamental frequency (Fo)

Average fundamental frequency was measured for phonation and spontaneous speech. The mean, standard deviation (SD) and range for average F_o for normal males (NM), normal females (NF), dysphonic males (DM), and dysphonic females (DF) are presented in table 1 and the mean is shown in graph 1.

Table 1 : Table Showing mean, Standard Deviation (SD) and range for average fundamental frequency measured.

Groups	Mean	S.D.	Range
NM	130	17.7	94-178
DM	162	48.54	105 - 351
NF	230	14.67	199 - 282
DF	211	34.90	110 - 332

The comparison of normal Vs Dysphonics, both males and females and also comparison between sexes, in terms of average fundamental frequency are shown in Table 1.1.

Table 1.1 : Table showing significance across groups for F_o.

Group	Significance
NM Vs NF	+
NM Vs DM	+
DM Vs DF	+
DF Vs NF	+

(+) indicates significant difference at 0.05 significance level.

Table 1 shows that female subjects have significantly higher fundamental frequency than males. This is consistent with other studies in the Indian population for adult males and females in phonation of vowels.

Text output tool

Helps the results to be viewed in a simple format that can be easily understood. It shows the results in percentage as to how will the neural network has performed during the training - validation and test phases.

After the model was implemented, the "Run data file" was used to further check the performance of the neural network. Additional 20 subjects, 10 normals and 10 dysphonics were included for running the ANN for validation purpose.

In order to access the data into the "Run data file".

1. Click the data input tool once. The tool menu appears.
2. From the data input tool choose 'view' a spread sheet appears.
3. From the file menu choose "OPEN" the data input dialog box appears.
4. In the run data input group, click "fLat file"
5. Click "Configure" A file window appears.
6. The file name is typed. The run data was recorded in the format "ANN.Sav".
7. After selecting the file, click "OK".
8. After this the spread sheet of the run data appears.

The run data consisted of a total of 20 subjects, 10 each for the both groups. In the run data only the acoustic parameters were given without the diagnosis and the network model has to predict the diagnosis of the voice disorder.

In order to run the run data into the application model created:

1. From the data output menu choose "view"
2. From the file name choose "Setup"

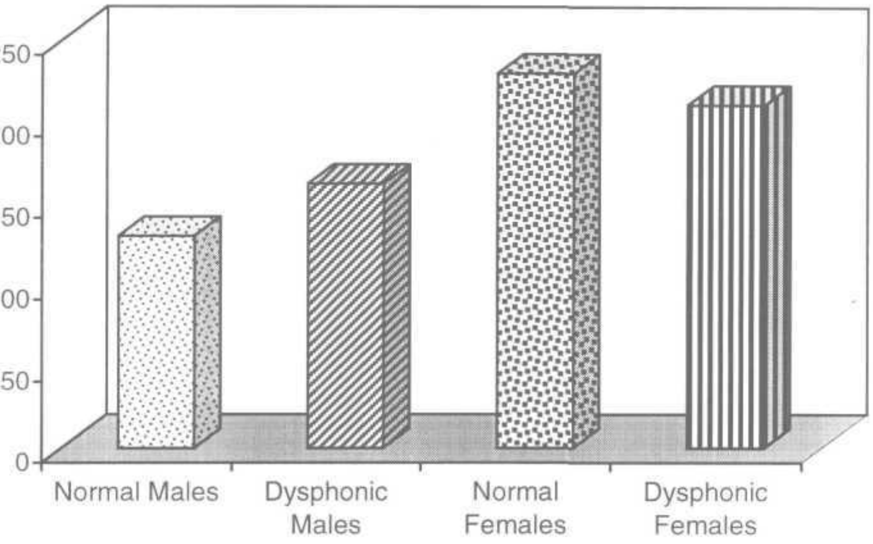
Table 1b : Showing the mean Fo in phonation reported by different investigator in normals males and females.

Investigators	Males	Females
Sheela (1974)	126	217
Jayaram (1975)	123	225
Nataraj and Jagadeesh (1984)	141	237
Vanaja (1986)	127	234
Nataraja (1986)	119	223
Anitha (1994)	129	240
Present study	130	230

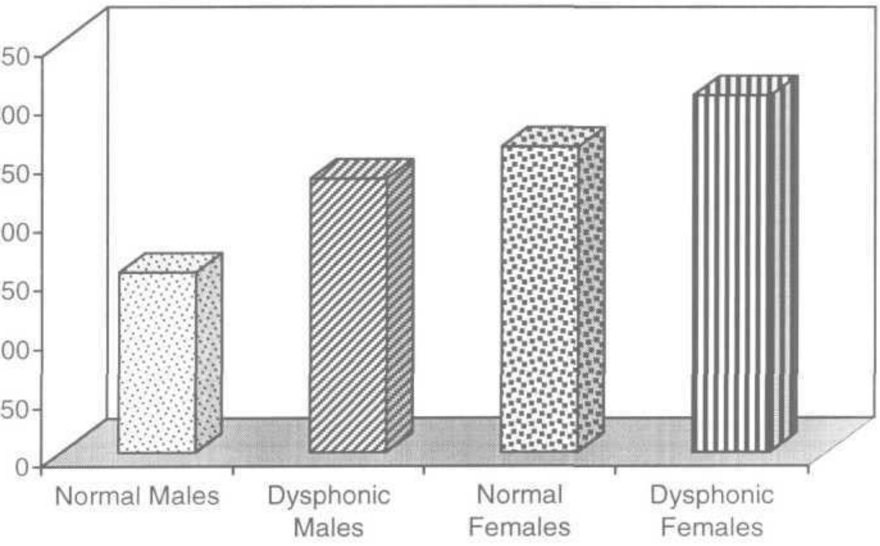
Table 1 shows that higher mean values for dysphonic males than normal males and lower mean values for dysphonic females than normal females. The standard deviation and range are also higher in the dysphonic groups. This coincides with the findings of Anitha (1994) and Das (1996). Greater standard deviation and range in the dysphonic group could be due to the hetrogeneity of the disordered group.

Thus the hypothesis stating that there is no significant difference between normals and dysphonics, both males and females is rejected. The hypothesis stating that there is no significant difference between normal males and normal females and dysphonic males and dysphonic females is also rejected. Indira (1982) found fundamental frequency as an important factor in differentiating abnormal cries from normal cries.

This parameter has been considered useful in differentiation of dysphonics from normals by other investigators also (Jayaram, 1975). Hence therapies have been concentrated to change the fundamental frequency and



GRAPH 1: MEANS OF AVERAGE FUNDAMENTAL FREQUENCY (F_o)



GRAPH 2 : MEANS OF HIGHEST FUNDAMENTAL FREQUENCY (Fhi) use his optimum pitch (Nataraja and Jayaram, 1982 ; Cowan. 1936 ; West et al., 1957 ; Anderson, 1961 ; Van Riper and Irwin, 1966).

Moreover, there was no significant difference in fundamental frequency across the five sessions in normal males and females indicating that fundamental frequency is a stable measurement atleast in normals.

Thus the hypothesis stating that their no significant difference in fundamental frequency accross the five trails is accepted.

2. Highest Fundamental Frequency (FHi)

The mean, SD and range for FHi in normal males, normal females, dysphonic males and dysphonic females are shown in Table 2 and the mean is shown in Graph 2.

Table 2 : Table showing mean, SD and range for FHi in different groups

Groups	Mean	S.D.	Range
NM	153	48.81	97 - 559
DM	233	145.60	79 - 672
NF	260	39.39	204 - 407
DF	304	203.00	136 - 2509

It is seen that mean for dysphonics were more than normals for both males and females. The standard deviation and range were also higher for dysphonics than normals. This finding is supported by the findings of Aparna (1996) who found Fhi to be greater in hearing impaired than normals, for both males and females.

The comparison of normals and dysphonics, both males and females and the comparison between normal males and normal females and dysphonic males and dysphonic females is shown in Table 2.1.

Table 2.1 : Table showing presence (+) or absence (-) of significance.

Group	Significance
NM Vs NF	+
NM Vs DM	+
DM Vs DF	-
DF Vs NF	-

It is seen in Table 2.1 that there is significant difference in HFi between normal males and normal females. Normal females show higher HFi than normal males. This is consistent with the findings of Biran (1995) who found significant difference between males and females across age groups in children. Gopal krishna (1995) also found greater Fhi in females than males which supports the findings of the present study.

Significant difference in terms of Fhi seen between normal males and normal females was not seen between dysphonic males and dysphonic females although dysphonic females showed higher mean HFi than dysphonic males.

There was no significant difference between dysphonic females and normal females but significant difference was noted between dysphonic males and normal males. This difference between males and females could be because the dysphonic males and dysphonic females included in the study was not matched for type and severity of the disorder. Significant difference between normal males and disphonic males is also supported by the findings of Das (1996).

It was also noted that HFi showed significant difference across sessions in males and not in females. Thus the hypothesis stating that there is no significant difference across five sessions is rejected for males and accepted for females.

The hypothesis stating that there is no significant difference between normals and dysphonics is accepted for females and rejected in males. The hypothesis stating that there is no signficant difference between normal males and normal females is rejected. The hypothesis stating that there is no significant difference between dysphonic males and dysphonic females is accepted.

General conclusion about the diagnostic value of Fhi is difficult to make because such measurements are helpful in certain pathological conditions and not in others (Kent, 1976).

3. Lowest fundamental frequency (FLO) :

When this parameter was subjected to statistical analysis the following results were obtained and the mean values are also shown in Graph 3.

Table 3 : Table showing mean, SD and range for FLO

Groups	Mean	S.D.	Range
NM	117	22.5	66 - 171
DM	123	37.8	64 - 254
NF	206	33.59	88 - 274
DF	167	52.89	0 - 245

Table 3.1 : Table showing comparison between the groups, for the parameter FLO

Group	Significance
NM Vs NF	+
NM Vs DM	-
DM Vs DF	+
DF Vs NF	+

From Table 3, it is seen that mean FLO is greater for females than males, for both in normal and dysphonic groups. It is also seen that the range and standard deviation were greater in the dysphonic group than in the normal group. Greater Flo in females than in males has been also reported by Anitha (1994), Gopal Krishna (1995), Biran (1995). The greater range seen in the dysphonic group could be due to the heterogeneity of disorders included in the dysphonic group. The range and S.D. are also greater in females than in males, for both the disordered and normal population.

Table 3.1 indicates that there is significant difference between males and females, in both the disordered and normal population.

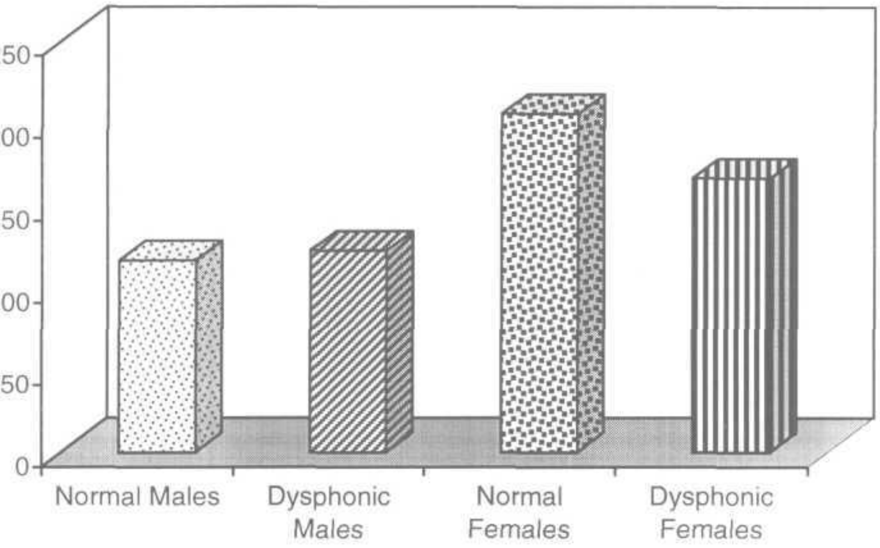
It was also seen that there is significant difference between dysphonic females and normal females but there is no significant difference between normal males and dysphonic males. This differences could be again because the male dysphonic group and female dysphonic group are not matched for severity and type of the disorder. As quoted by Kent (1976), these parameters may be helpful in differentiating certain pathological conditions but not in others.

It is also seen that across the sessions there is no significant difference in normals which implies that FLO is a relatively stable measure and does not vary across sessions.

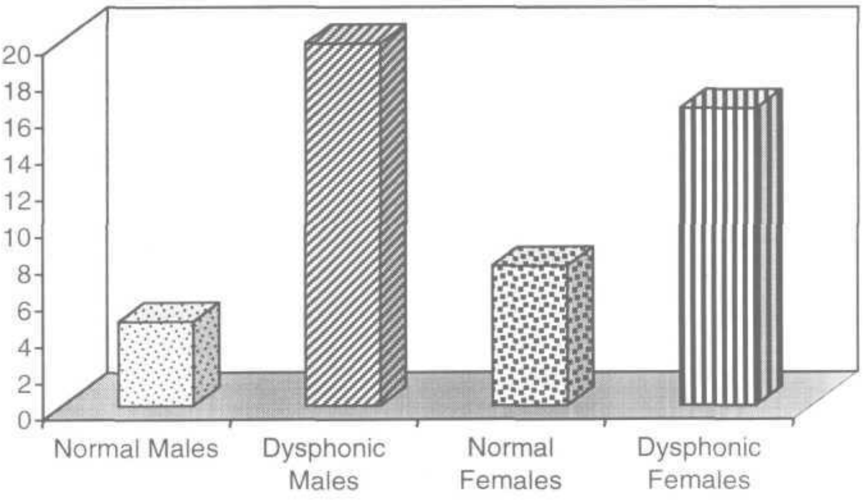
Thus the hypothesis stating that the there is no significant difference across the five sessions is accepted. The hypothesis stating that the there is no significant difference between normals and dysphonics is accepted in males and rejected in females. The hypothesis stating that there is no difference between males and females is rejected.

Hence, relatively extensive data has to be collected before it can be applied to the clinical population.

GRAPH 3 : MEANS OF LOWEST FUNDAMENTAL FREQUENCY (Flo)



GRAPH 4 : MEANS OF STANDARD DEVIATION OF FUNDAMENTAL FREQUENCY (STD)



4. Standard Deviation of F_o (STD) :

Table 4 : Showing mean, S.D. and range for STD

Groups	Mean	S.D.	Range
NM	4.61	7.21	0.58 - 54.16
DM	19.84	33.34	1.26 - 198
NF	7.63	9.71	0.41 - 56
DF	16.29	19.45	1.18 - 77.20

Table 4.1 : Table showing comparison between the groups, for the parameter STD; (+) indicates presence of significant difference between the groups

Group	Significance
NM Vs NF	+
NM Vs DM	+
DM Vs DF	-
DF Vs NF	+

Graph 4 shows mean STD.

Table 4 shows that mean STD are greater in normal females than normal males. This coincides with the findings of Gopal Krishna (1995). In the dysphonic group, both males and females, show greater mean STD than normals. The SD and range are also greater for the disorder population than for normal subjects.

There is significant difference between males and females in normals no significant difference is found between dysphonic males and dysphonic females.

The hypothesis stating that there is no signficant difference between normals and dysphonics is rejected. The hypothesis stating that there is no significance between males and females is rejected in normals and accepted in dysphonics.

It is also seen that across the sessions there is no significant difference in normals. Hence the hypothesis stating that there is no significant difference across the five sessions is expected.

Significant difference is seen between normal males and dysphonic males and between normal females and dysphonic females. Moreover there is no significant differences in STD across sessions in the normal population for both males and females. This indicates that STD is a reliable and stable measure and can be also used to differentiate between normals and dysphonics. The results of experiment by Anitha (1994) and Das (1996) support the results of the present study.

5. Phonatory F_o range semitones (PFR) :

The mean PFR is shown in Table 5 and Graph 5. Table 5 also shows S.D. and range of PFR.

Table 5 : Table showing mean, S.D. and range for PFR.

Groups	Mean	S.D.	Range
NM	5.35	5.57	1 - 34
DM	11.00	10.44	2 - 41
NF	5.24	5.11	1 - 27
DF	9.42	7.53	2 - 26

Table 5.1 : Table showing significance between groups for PFR.

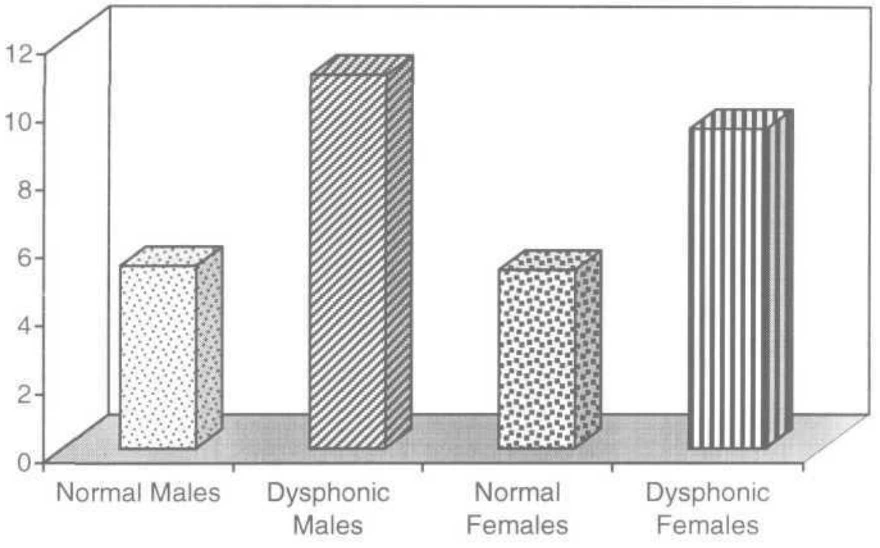
Group	Significance
NM Vs NF	-
NM Vs DM	+
DM Vs DF	-
DF Vs NF	+

Table 5 shows that mean PFR is greater for the dysphonic group than normals, in both males and females. Greater mean values in the dysphonic group was also noted by Das (1996). Table 5 also shows that mean PFR is approximately the same for males and females in the normal group. This is also seen in the disordered group.

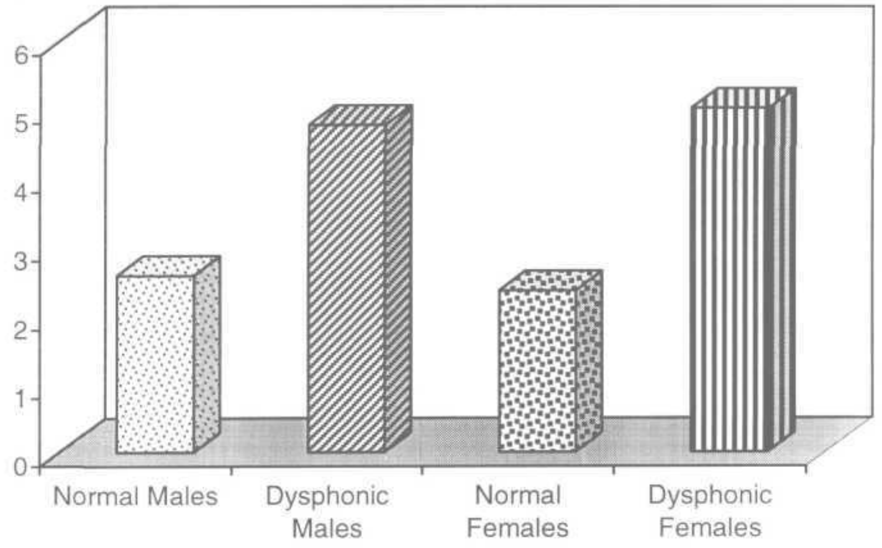
Inspection of Table 5.1 shows significant difference between normal males and dysphonic males. Significant difference is also seen between normal females and dysphonic females.

There is no significant difference between normal males and normal females and between dysphonic males and dysphonic females. All the above findings are in agreement with the findings reported by Anitha (1994).

GRAPH 5 : MEANS OF PHONATORY F_o RANGE IN SEMITONES (PFR)



GRAPH 6 : MEANS OF Fo TREMOR FREQUENCY (Fftr)



It is also seen that there are no significant differences across the sessions for this parameter. Hence, it is concluded that PFR is a relatively stable parameter and does vary across sessions. The hypothesis stating that there is no significant difference across the five sessions is accepted.

The higher values of PFR in dysphonics can be attributed to the inability of the dysphonics to maintain a constant pitch during sustained phonation.

Hence, the hypothesis that there is no significant difference been normals and dysphonics, in both males and females is rejected. Hence this study implies that PFR can not be used as a diagnostic measure to differentiate normals from dysphonics. Thus the hypothesis stated that there is no significant difference between males and females is accepted for both normals and dysphonics. The hypothesis that there is no significant difference between males and females, in both normal and disordered group is accepted.

6. F_o Tremor frequency (Fftr)

The results of this parameter are summarised in Table 6, Table 6.1 and Graph 6.

Table 6 : Table showing mean, S.D. and range for Fftr in normals and dysphonics

Groups	Mean	S.D.	Range
NM	2.58	2.62	0 - 21
DM	4.77	5.81	0 - 23
NF	2.36	2.24	0 - 14
DF	5.02	5.68	0 - 22

Table 6.1 : Table showing significance between groups for PFR.

Group	Significance
NM Vs NF	-
NM Vs DM	+
DM Vs DF	-
DF Vs NF	+

Table 6 shows that mean, S.D. and range are greater for the dysphonic groups than normals, in both males and females. The mean values for males and females are approximately the same for both dysphonics and normals.

Table 6.1 shows that there is significant difference between normal males and dysphonic males and between normal females and dysphonic females. No significant difference was found between males and females, in both normal and dysphonic groups.

The above findings are in consance with the findings of Anitha (1994). Das (1996) also found no significant difference between normals and dysphonics for this parameter.

No significant difference was observed across the sessions for both males and females. Thus hypothesis stating that there is no significant difference across sessions is accepted.

The hypothesis stating that there is no significant difference between normals and dysphonics is rejected. The hypothesis stating that there is no significant difference between males and females, for both normals and dysphonics is accepted. Hence Fftr can be used to differentiate between normals and dysphonics.

7. Amplitude tremor frequency (FATR)

The mean, standard deviation and range of FATR for the four groups i.e. normal males, normal females, dysphonic males and dysphonic females are presented in Table 7. Graph 7 shows the mean values of the four groups. Table 7.1 shows the results of T test for this parameter.

Table 7 : Showing mean, SD and range in the four groups for FATR

Groups	Mean	S.D.	Range
NM	2.14	1.55	0-10.52
DM	4.97	5.96	0-22.2
NF	1.77	1.71	0-10.81
DF	25.1	112.47	0-727

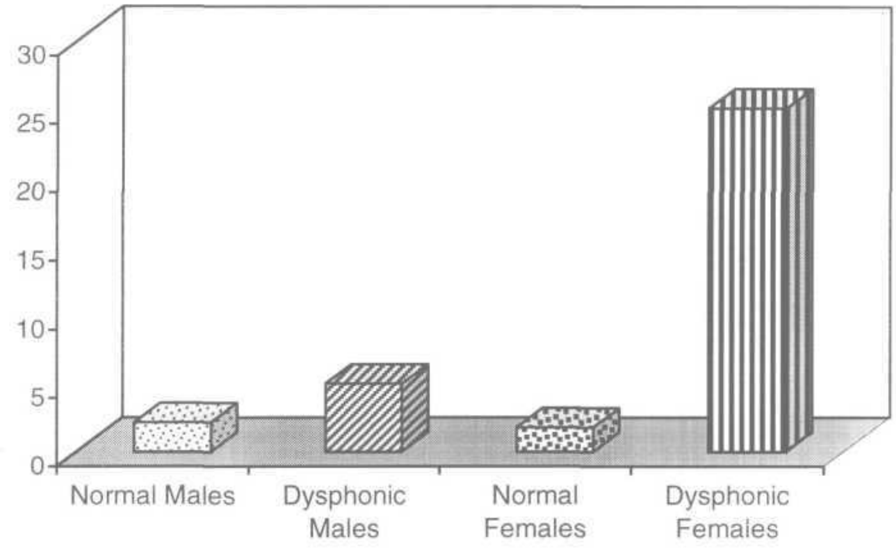
Table 7.1 : Comparison among the four groups for FATR

Group	Significance
MM Vs NF	-
ISM Vs DM	+
DM Vs DF	-
DF Vs NF	-

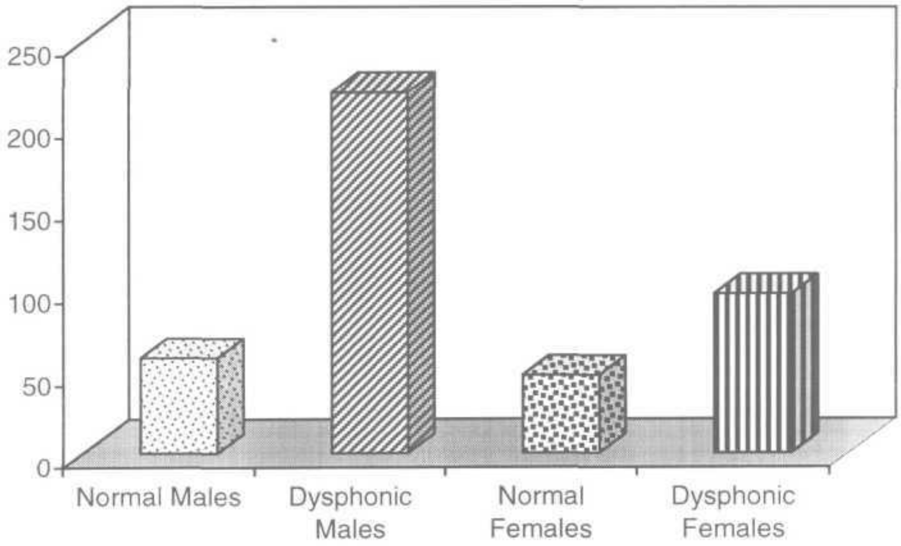
'+' indicates presence of significant difference.

Table 7 indicates that the mean FATR scores are greater for Dysphonics than normals. The dysphonics also show greater SD and range.

GRAPH 7 : MEANS OF AMPLITUDE TREMOR FREQUENCY (Fatr)



GRAPH 8 : MEANS OF ABSOLUTE JITTER (Jita)



Greater mean values are observed in normal males than normal females. This is supported by the findings of Gopal Krishna (1995).

From table 7.1, it can be seen that there is significant difference between normal males and Dysphonic males but no significant difference between normal males and normal females; Dysphonic males and Dysphonic females and also Dysphonic females and normal females. FATR was found to be relatively stable across sessions for normals.

Thus the hypothesis stating that there is no significant difference across the sessions is accepted for this parameter. The hypothesis stating that there is no significant difference between males and females is accepted for both normals and dysphonics. The hypothesis stating that there is no significant differnce between normals and dysphonics is accepted for females and rejected for males.

It is also seen that mean values are greater for the Dysphonic group than normals. This could be attributed to the inability of the dysphonics to maintain a constant pitch and intensity. This findings also correlates with the findings reported by Anitha (1994). Das (1996) also found that mean FATR for vowels was greater in the dysphonic group than normal group.

Although significant difference was found between normal males and dysphonic males, it was not true between normal females and dysphonic females. This could be because the dysphonic males and dysphonic females were not matched in terms of type and severity.

8. Absolute Jitter (Jita)

The results are briefed down in Table 8, Graph 8 and Table 8.1.

Table 8 : Showing mean, SD and range for Jita

Groups	Mean	S.D.	Range
NM	58.07	50.44	0.29-286
DM	219.2	221.2	31-1192
NF	47.42	30.04	14-150
DF	96.77	84.51	2.25-501

Table 8.1 : Showing comparison between groups. '+' indicates presence of significant difference.

Group	Significance
NM Vs NF	-
NM Vs DM	+
DM Vs DF	+
DF Vs NF	+

Table 8 shows that mean Jita is greater in the dysphonic groups than in normals, for both males and females. The S.D and range are also greater among the dysphonics. Jita values are also found to be greater in males than in females for both normal and dysphonic groups. Greater means in normal males when compared to normal females was observed by Gopal Krishna (1995).

Table 8.1 shows that there is no significant difference between normal males and normal females. However significant difference is observed between normal males and dysphonic males; normal females and dysphonic females; and dysphonic males and dysphonic females.

No significant difference was noticed across the five sessions for both males and females. This shows that Jita is a relatively stable measure and does not vary across sessions atleast in normals.

Thus the hypothesis stating that there is no significant difference across sessions is accepted. The hypothesis stating that there is no significant difference between normals and dysphonics is rejected for both males and females. The hypothesis stating that there is no significant difference between males and females is accepted for normals and rejected in dysphonics.

The greater Jita values in dysphonics may be due to the inability of the dysphonics to maintain a constant pitch in phonation. This is in agreement with the results of study done by Vonleder et al (1966), Chandrashekar (1987) and Anitha (1994).

Greater mean values of Jita in males may be because Jita is an absolute measure and shows the result in microsecond which makes it dependent on the average fundamental frequency of voice. For this reason, higher pitch results in a lower Jita and lower pitch in a higher Jita. This finding supports the reports of earlier studies by Higgins and Saxman (1989); Robert and Baker (1984).

Since significant difference is observed between normals and dysphonics, both males and females this can be used to differentiate between normals and dysphonics.

9. Jitter Percent (Jitt) :

The results of this parameter are shown in Tables 9, 9.1 and Graph 9.

Table 9 : Table showing mean, S.D and range of Jitt :

Groups	Mean	S.D.	Range
NM	1.19	3.79	0.16-35.97
DM	9.22	36.29	0.50-246
NF	1.59	3.91	0.12-37
DF	2.43	2.05	0.48-10.90

Table 9.1 : Table showing comparison between groups.

Group	Significance
NM Vs NF	-
NM Vs DM	-
DM Vs DF	-
DF Vs NF	-

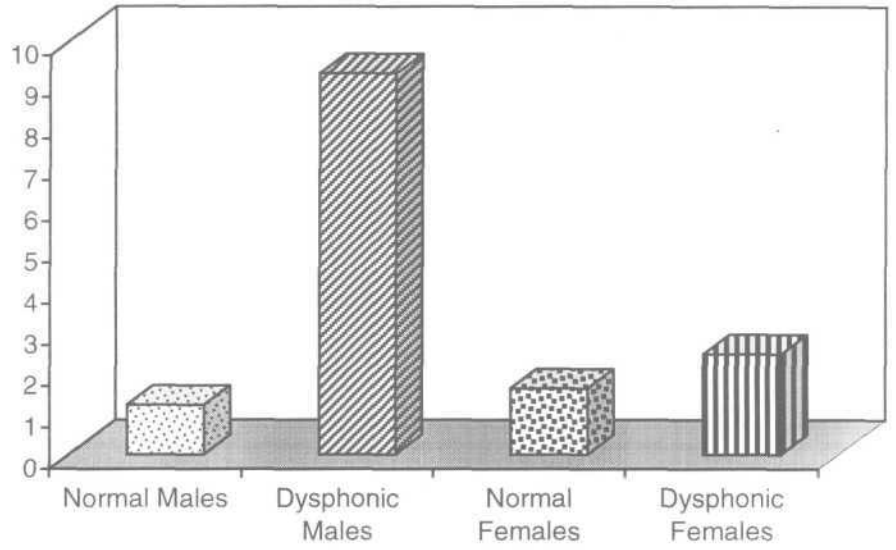
Table 9 shows that the mean values are higher in dysphonics than normals. The mean Jitt in normal is approximately the same for males and females.

Table 9.1 shows that there is no signficant difference between normal males and normal females; normal males and dysphonic males; normals females and dyshonic females and dysphonic females and dysphonic males.

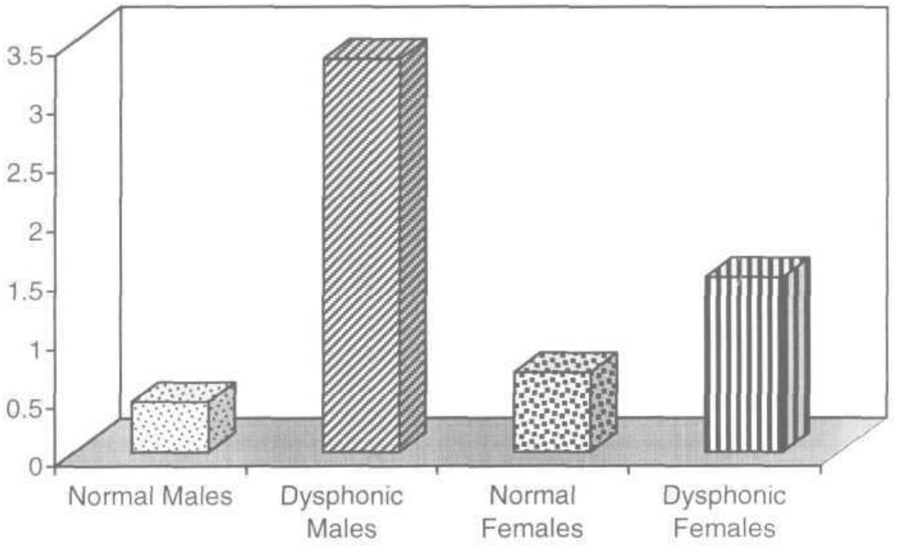
No significant diffrence is noted across sessions for this parameters.

Thus the hypothesis stating that there is no significant difference across sessions is accepted. The hypothesis stating that there is no significant difference between normals and dysphonics in both males and females in

GRAPH 9 : MEANS OF JITTER PERCENTAGE (Jitt)



GRAPH 10 : MEANS OF RELATIVE AVERAGE PERTUBRATION (RAP)



accepted. The hypothesis stating that there is no significant difference between males and females is also accepted.

Absolute Jitter and Jitter percentage represent evaluations of the same type of pitch. But Jitt is a relative measure and the influence of the average fundamental frequency of the subject is significantly reduced. Hence no significant difference is seen between males and females unlike the Jita measures. This is in agreement with the findings of Anitha (1994) and Gopal Krishna (1995)

10. Relative average perturbation (RAP)

The results of statistical analysis for this parameter are summarsied in Table 10, Table 10.1 and Graph 10.

Table 10 : Showing mean, S.D and range for RAP

Groups	Mean	S.D.	Range
NM	0.43	0.34	0.08-2.25
DM	3.35	7.93	0.29-53.3
NF	0.68	0.45	0.20-2.29
DF	1.49	1.26	0.27-6.43

Table 10.1 : Showing comparison between groups "+" indicates presence of significant difference.

Group	Significance
NM Vs NF	+
NM Vs DM	+
DM Vs DF	-
DF Vs NF	+

Table 10 shows that mean, S.D and range is greater in dysphonics than normals. Normal females have a greater mean values than normal males . This is supported by the findings of Anitha (1994) and Gopal Krishna (1995).

Inspection of Table 10.1 shows that there is significant difference between normal males and normal females; normal males and dysphonic males and normal females and dysphonic females.

No significant difference was seen across sesions indicating that RAP is relatively stable across sessions.

Thus the hypothesis stating that there is no significant difference across sessions is accepted. The hypothesis stating that there is no significant difference between normals and dysphonics is rejected. The hypothesis stating that there is no significant difference between males and females is rejected. The hypothesis stating there is no significant difference between dysphonics males and females is accepted.

Thus RAP can be used to differentiate between normals and dysphonics. This is consistent with the findings of Venkatesh (1992), Anitha (1994) and Das (1996).

Greater mean in dysphonics seems to be associated with the inability of the vocal cords to support a periodic vibration with a defined period.

11. Pitch perturbation quotient (PPQ) :

The results regarding this parameter are summarised in Table 11, Table 11.1 and Graph 11.

Table 11 : Table showing mean, S.D and range for PPQ

Groups	Mean	S.D.	Range
NM	0.43	0.33	0.10-2.17
DM	2.33	2.40	0.28-11.18
NF	0.64	0.40	0.2-2.14
DF	1.5	1.4	0.27-7.44

Table 11.1 : Showing comparison between groups. "+" indicates presence of significant difference

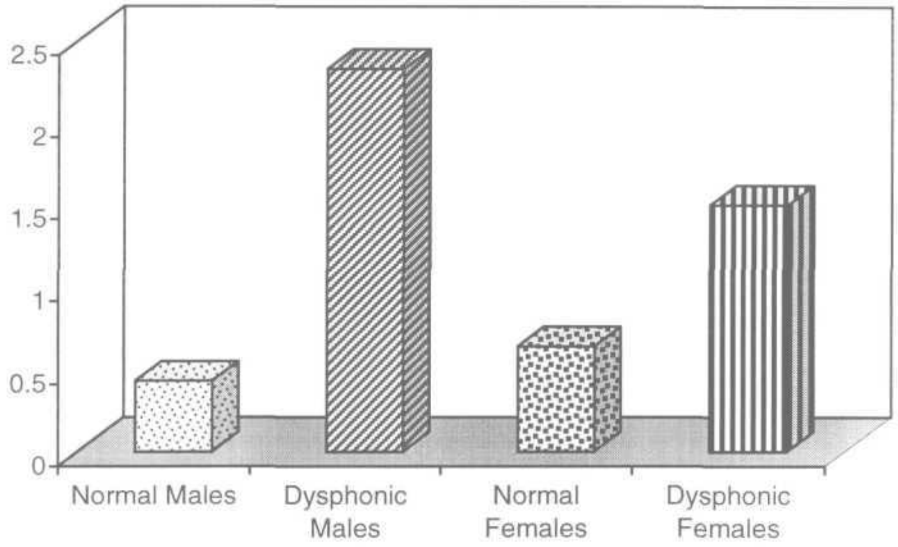
Group	Significance
NM Vs NF	+
NM Vs DM	+
DM Vs DF	+
DF Vs NF	+

Table 11 shows that the mean values of PPQ are greater for the dysphonic group, for both males and females than in case of normals. Mean RAP is also greater for normal females than normal males.

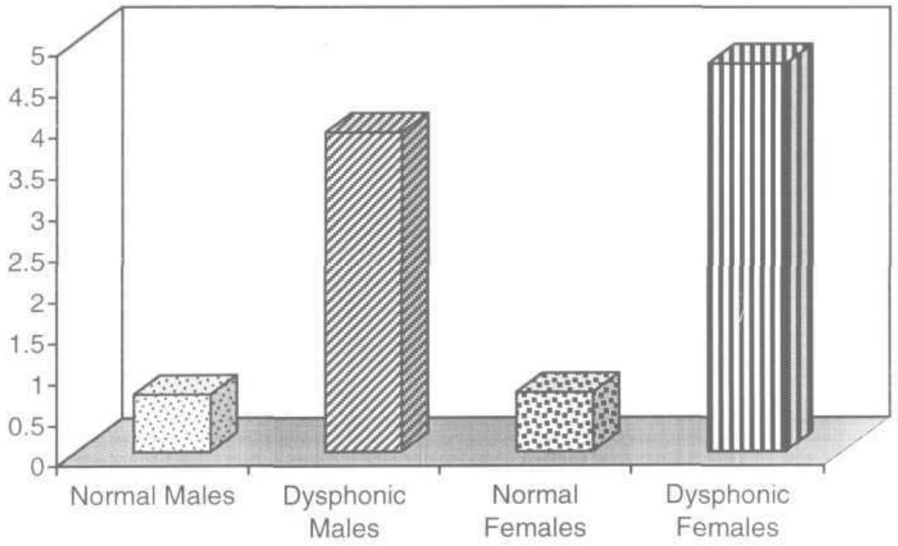
From the study of Table 11.1, it can be stated that there is significant difference between normal males and normal females; Dysphonic males and normal males; dysphonic males and dysphonic females and also dysphonic females and normal females.

It is also observed that there is no significant difference across sessions in both males and females. Hence RAP is relatively stable across sessions atleast in normals.

GRAPH 11 : MEANS OF PITCH PERTUBRATION QUOTIENT (PPQ)



GRAPH 12 : MEANS OF SMOOTHED PITCH PERTUBRATION QUOTIENT (SPPQ)



Thus the hypothesis stating that there is no significant difference across sessions is accepted. The hypothesis stating that there is no significant difference between normals and dysphonics is rejected, with respect to both males and females. The hypothesis stating that there is no significant difference between males and females is rejected.

The results of this study goes along with the findings of the study by Davis (1976) Anitha (1994) and Das (1996). The increase in means in dysphonics may be attributed to inability to maintain constant pitch during phonation by the dysphonics due to various vocal pathology in their vocal folds.

12. Smoothned pitch perturbation quotient (SPPQ)

The results of this parameter are shown in Table 12, Table 12.1 and Graph 12.

Table 12. Showing mean, S.D and range for SPPQ

Groups	Mean	S.D.	Range
NM	0.7	0.51	0.25-4.18
DM	3.89	5.54	0.4-23.9
NF	0.72	0.38	0.26-2.16
DF	4.73	17.21	0.42-116

Table 12.1: Showing comparison between groups. "+" indicates presence of significant difference

Group	Significance
NM Vs NF	-
NM Vs DM	+
DM Vs DF	-
DF Vs NF	-

Table 12 shows that mean SPPQ values are greater in Dysphonics than in normals, for both males and females. S.D and range are also found to be greater in the dysphonic group. This agrees with the findings of Anitha (1994) and Das (1996). sPPQ remained stable across the five sessions for males and females.

On inspection of Table 12.1, it can be stated that there is no significant diffrence between males and females, for both normals and dysphonics. There is significant diference between normal males and dysphonic males but no significant difference between normal females and Dysphonic females.

This difference could be because the dysphonic male and dysphonic female gorups are not matched for type and severity. SPPQ may be increased in certain voice disorders (Eg: Spasmodic dysphonia) but not in others.

Thus the hypothesis stating that there is no significant difference across sessions is accepted. The hypothesis stating that there is no significant difference between normals and dysphonics is rejected in case of males and accepted for females. The hypothesis stating that there is no significant difference between males and females is accepted for both normals and dysphonics.

13. Co-efficient of Fo variation (vFo)

The results of statistical analysis for this parameter are summarised h Table 13, Table 13.1 and Graph 13.

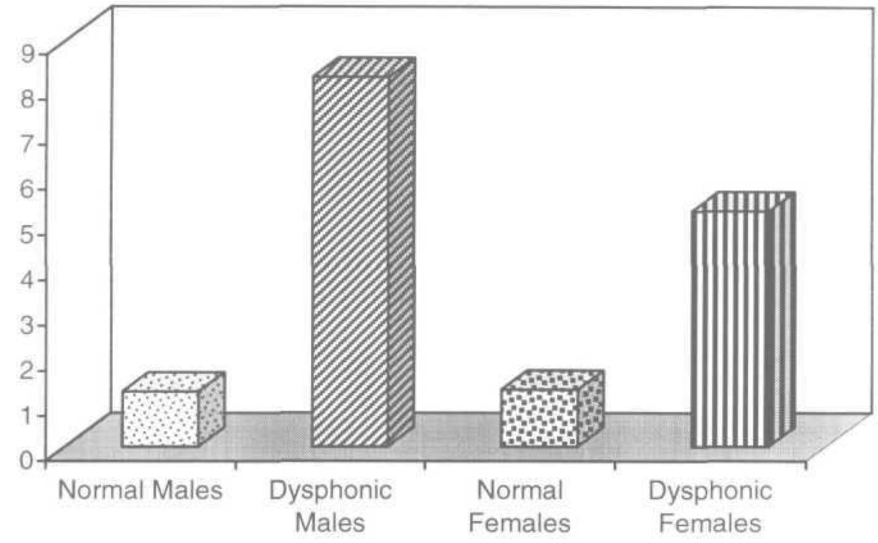
Table 13. Showing mean, S.D and range for vFo

Groups	Mean	S.D.	Range
NM	1.23	1.05	0.43-8.24
DM	8.19	11.45	0.84-40.9
NF	1.27	1.24	0.01-10.62
DF	5.22	8.45	0.75-40.82

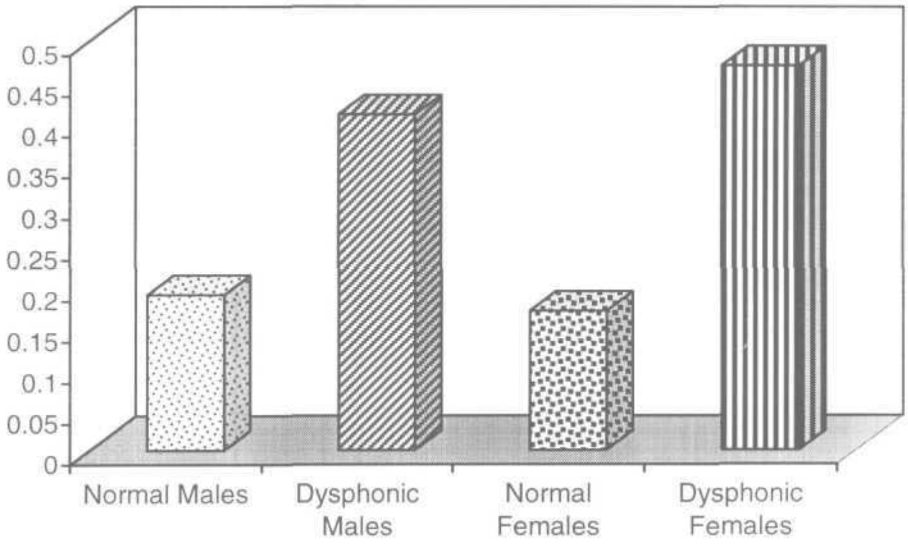
Table 13.1 : Showing comparison between groups. "+" indicates presence of significant difference

Group	Significance
NM Vs NF	-
NM Vs DM	+
DM Vs DF	-
DF Vs NF	+

From Table 13, we can see that mean values for dysphonics is greater than for normals, in both males and females. S.D. and range is also greater for the dysphonics, than for normals. Mean values for normal males is lower than normal females which is supported by the findings of Gopal Krishna (1995).



GRAPH 13 : MEANS OF CO-EFFICIENT OF F_o VARIATION (vFo)



GRAPH 14 : MEANS OF SHIMMER IN dB (Shdb)

Table 13.1 shows that there is no significant difference between males and females, in both normals and dysphonics. However, significant difference is noted between normal males and dysphonic males and also normal females and dysphonic females.

No significant difference is noted across the sessions indicating that VFO is a relatively stable measure atleast in normals.

Thus the hypothesis stating that there is no significant difference across sessions is accepted. The hypothesis stating that there is no significant difference between normals and dysphonics is rejected in case of males and females. The hypothesis stating that there is no significant difference between males and females is accepted for both normals and dysphonics.

Greater means are seen in the dysphonic group and it can be used to differentiate dysphonics from normals. This is similr the findings of Yanagihara (1967); Iwata and Vonleden (1970); Nataraja (1981); Bhat (1981); Anitha (1994) and Das (1996).

The means of VFO might have increased in dysphonics because of the inability of the dysphonics to maintain a constant pitch during phonation and speaking.

14. Shimmer in dB (ShdB)

The results of statistical analysis for ShdB are shown in Table 14, Table 14.1 and Graph 14.

Table 14 : Showing mean, S.D. and range for ShdB.

Groups	Mean	S.D.	Range
NM	0.19	0.21	0.1-1.18
DM	0.41	0.62	0.02-3.77
NF	0.17	0.19	0.01-1.37
DF	0.47	1.08	0.02-7.2

Table 14.1 : Showing comparison between groups. "+" indicates presence of significant difference:

Group	Significance
NM Vs NF	-
NM Vs DM	+
DM Vs DF	-
DF Vs NF	-

Table 14, reveals that mean values of ShdB are higher in Dysphonics than in normals, for both males and females. S.D and range are also greater in the Dysphonic group.

Table 14.1 shows that there is no significant difference between males and females, in both normals and dysphonics. Significant difference is seen between normal males and dysphonic males. No significant difference is noted between dysphonic females and normal females although dysphonic females show higher mean values than normal female.

Results of T test reveals that ShdB varied across sessions, in both males and females and thus it is not a stable measure even in normals.

Thus the hypothesis stating that there is no significant difference across sessions is rejected. The hypothesis stating that there is no significant difference between normals and dysphonics is accepted in case of females and rejected in case of males. The hypothesis stating that there is no significant difference between males and females is accepted for both normals and dysphonics.

The mean values of dysphonics is greater than in normals. This is in concurrence with the findings of Vonleden et al (1960); Venkatesh et al (1992); Kitajima and Gould (1976); Anitha (1994) and Das (1996). This may be attributed to the inability of the dysphonics to maintain a constant intensity of voice.

The results of the present study coincides with the findings of Chandrashekar (1987) who found shimmer values to be greater for vocal nodule cases than normals with respect to both male and female groups, but the values were significant only for males. This may be because in normals, fluctuations in intensity increase with age more markedly than in females. (Vanaja, 1986; Tharman, 1991 and Suresh, 1991).

15. Shimmer percentage (shim) :

The mean, S.D and range and results of "T" test are shown in Table 15, Table 15.1 and Graph 15.

Table 15 : Showing mean, S.p and range for shim.

Groups	Mean	S.D. .	Range
NM	2.16	2.37	0.02-13.33
DM	4.24	6.24	0-37.24
NF	1.61	1.34	0.10-7.51
DF	3.39	3.69	0.19-19.21

Table 15.1 : Showing comparison between groups. '+' indicates presence of significant difference at 0.05 level

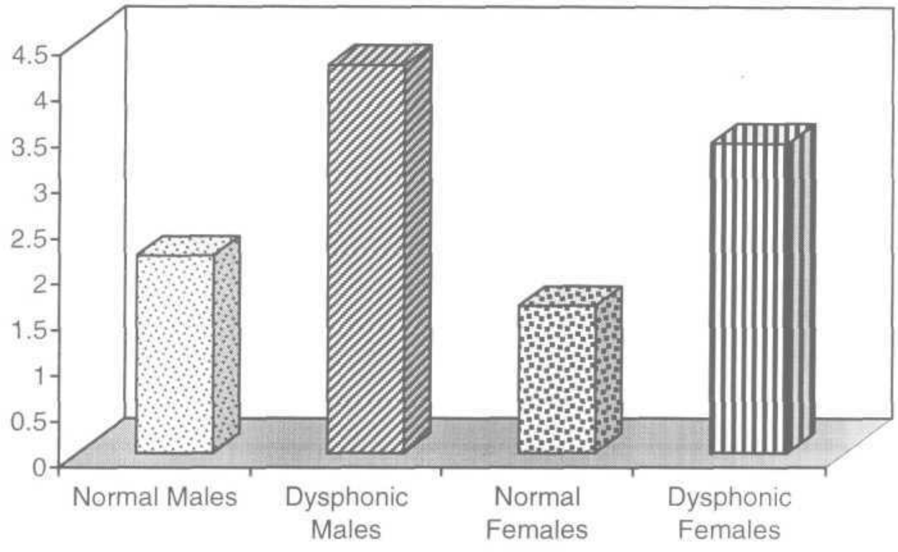
Group	Significance
NM Vs NF	-
NM Vs DM	+
DM Vs DF	-
DF Vs NF	+

On inspection of Table 15, it can be stated that mean values for the dysphonics are greater than normals, for both males and females. S.D and range are also greater in case of dysphonic group. However females show lower mean values than males.

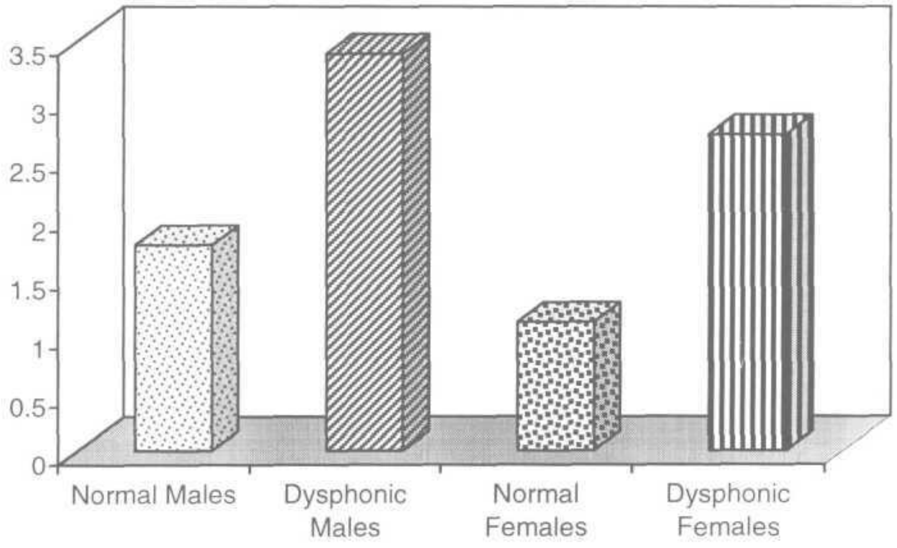
Table 15.1 reveals that there is significant difference between normals and dysphonics, in both males and females. No significant differences are observed between males and females both in normals and dysphonics. Significant difference is observed across trials.

Thus the hypothesis stating that there is no significant difference across sessions is rejected. The hypothesis stating that there is no significant

GRAPH 15 : MEANS OF SHIMMER IN PERCENTAGE (Shim)



GRAPH 16 : MEANS OF AMPLITUDE PERTUBRATION QUOTIENT (APQ)



difference between normals and dysphonics is rejected in case of males and females . The hypothesis stating that there is no significant difference between males and females is accepted for both normals and dysphonics.

Greater variations in dysphonics may indicate the inability of the vocal cords to support a periodic vibration for a defined period and also due to the presence of turbulent noise in the voice signal. This finding is similar to the findings of Anita (1994) and Das (1996) for the phonation of vowels.

It is also seen that females have lower values of shimmer in percentage when compared to males, for both normals and dysphonics. This finding is in agreement with the findings and the study done by Sorensen and Horii (1983), Anitha (1994) and Gopal Krishna (1995).

16. **Amplitude perturbation quotient (APQ)** :

The results of statistical analysis for this parameter are shown in Table 16, Table 16.1 and Graph 16.

Table 16 : **Showing mean, S.D and range for APQ**

Groups	Mean	S.D.	Range
NM	1.76	1.97	0.09-11.9
DM	3.39	5.15	0.15-30.9
NF	1.1	0.8	0.09-4.56
DF	2.7	3.04	0.14-13.33

Table 16.1 : Showing comparison between groups. '+' indicates presence of significant difference

Group	Significance
NM Vs NF	-
NM Vs DM	-
DM Vs DF	-
DF Vs NF	-

From table 16, mean values for dysphonics is greater than normals for this parameter S.D and range are also greater in the dysphonic groups. Table 16.1 shows that there is no significant difference between any groups i.e between normal males and normal females; dysphonic males and normal males; dysphonic males and dysphonic females; and dysphonic females and normal females. APQ was found to be stable across trials.

Higher mean values may be associated with the inability of the cords to support a periodic vibration with in a specified period and with the presence of turbulent noise in the voice signal. Greater mean values in dysphonics coinsides with the findings of Anitha (1994) and Das (1996).

Thus the hypothesis stating that there is no significant difference across sessions is accepted. The hypothesis stating that there is no significant difference between normals and dysphonics is accepted for both males and females. The hypothesis stating that there is no significant difference between males and females is accepted for both normals and dysphonics.

17. Smoothed Amplitude perturbation quotient (sAPQ) :

The values of this parameter are summarised in Table 17, Table 17, and Graph 17.

Table 17 : Showing mean, S.D and range for SAPQ

Groups	Mean	S.D.	Range
NM	3.47	3.53	0.13-19.23
DM	4.29	6.46	0.18-38.17
NF	1.87	1.11	0.11-5.6
DF	4.63	7.4	0.18-45.31

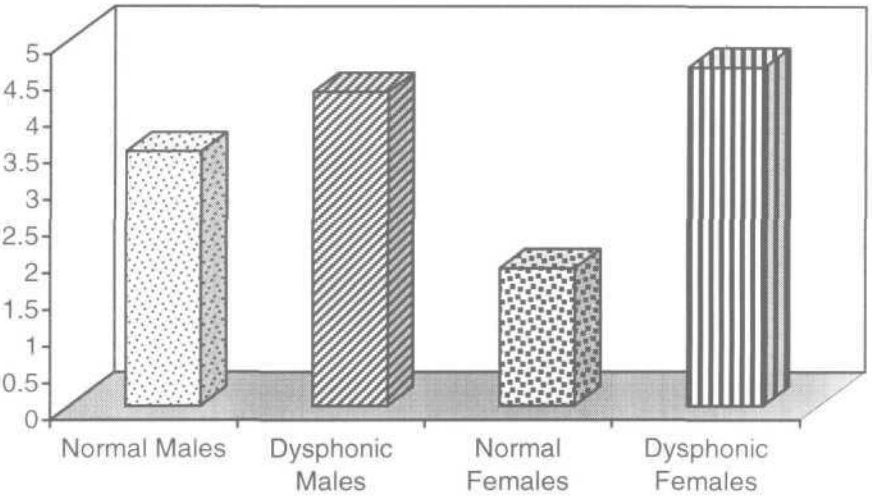
Table 17.1 : Showing comparison between groups. '+' indicates presence of significant difference at 0.05 level

Group	Significance
MM Vs NF	+
NM Vs DM	-
DM Vs DF	-
DF Vs NF	+

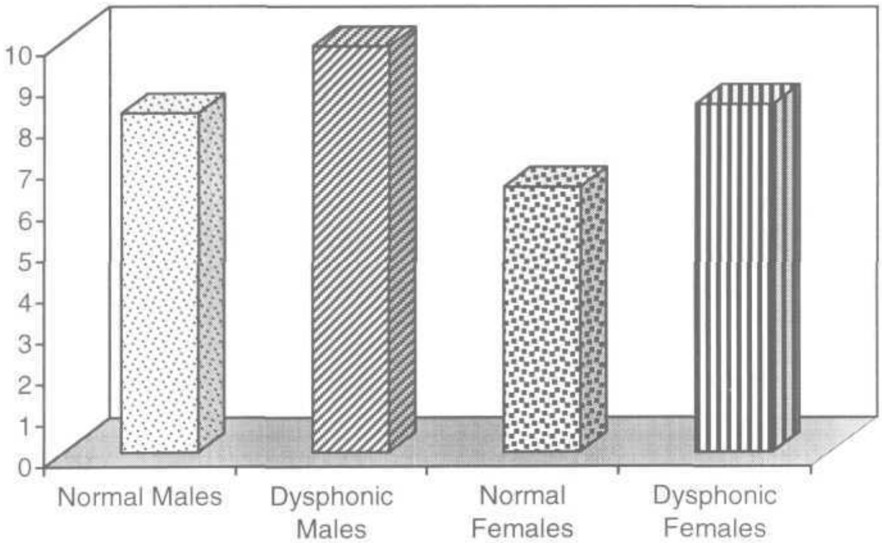
Table 17, shows that mean value are greater for the dysphonics than normals, with respect to both males and females. S.D and range is also greater for the dysphonic group.

From Table 17.1, there is significant difference between normal males and normal females and between normal females and dysphonic females. No significant difference is noted between normal males and dysphonic males and also between dysphonic males and dysphonic females.

GRAPH 17 : MEANS OF SMOOTHED AMPLITUDE PERTUBRATION QUOTIENT (sAPQ)



GRAPH 18 : MEANS OF CO-EFFICIENT OF AMPLITUDE VARIATION (vAM)



No significant difference was observed across trails for both males and females, for this parameter.

Thus the hypothesis stating that there is no significant difference across sessions is accepted. The hypothesis stating that there is no significant difference between normals and dysphonics is accepted in case of males and rejected for females. The hypothesis stating that there is no significant difference between males and females is accepted in case of dysphonics and rejected in case of normals.

The means of dysphonics were higher than means of normals for this parameter. This result is in agreement with the findings of Diliyski, Orlikoff and Kahane (1991); Anitha (1994) and Das (1996). Greater mean values in normal males than normal females is supported by the findings of Gopal Krishna (1995).

18. Co-efficient of amplitude variation (VAM)

The results of statistical analysis for this parameter are summarised h Table 18, Table 18.1 and Graph 18.

Table 18 : Showing mean, SD and range of vAM

Groups	Mean	S.D.	Range
NM	8.24	7.85	0.26-48.49
DM	9.87	10.85	0.21-40.61
NF	6.45	4.25	0.17-17.20
DF	8.47	9.19	0.12-35.55

Table 18.1 : Showing comparison between groups. '+' indicates presence of significant difference at 0.05 level.

Group	Significance
NM Vs NF	-
NM Vs DM	-
DM Vs DF	-
DF Vs NF	-

From Table 18, mean values for females are lesser than males, in both normals and dysphonics. S.D is also greater for the dysphonics, both males and females. Greater mean values in males agrees with the findings of Gopal Krishna (1995).

Table 18.1 shows that there is no significant difference between the groups, that is normal males and normal females; normal males and dysphonic males; dysphonic males and dysphonic females; and dysphonic females and normal females. No significant difference was found across trials for this parameter.

Thus the hypothesis stating that there is no significant difference across sessions is accepted. The hypothesis stating that there is no significant difference between normals and dysphonics is accepted in both males and females. The hypothesis stating that there is no significant difference between males and females is accepted with respect to both normals and dysphonics. Greater mean in dysphonics is supported by the findings of Anitha (1994) and Das (1996).

19. Noise to Harmonic Ratio (NHR)

The values for this parameter are given in Table 19, Table 19.1 , and Graph 19.

Table 19 showing mean, SD and range for NHR

Groups	Mean	S.D.	Range
NM	0.14	0.09	0.05-0.93
DM	1.7	9.65	0.06-65
NF	0.21	0.99	0.04-9.47
DF	0.44	1.72	0.05-11.71

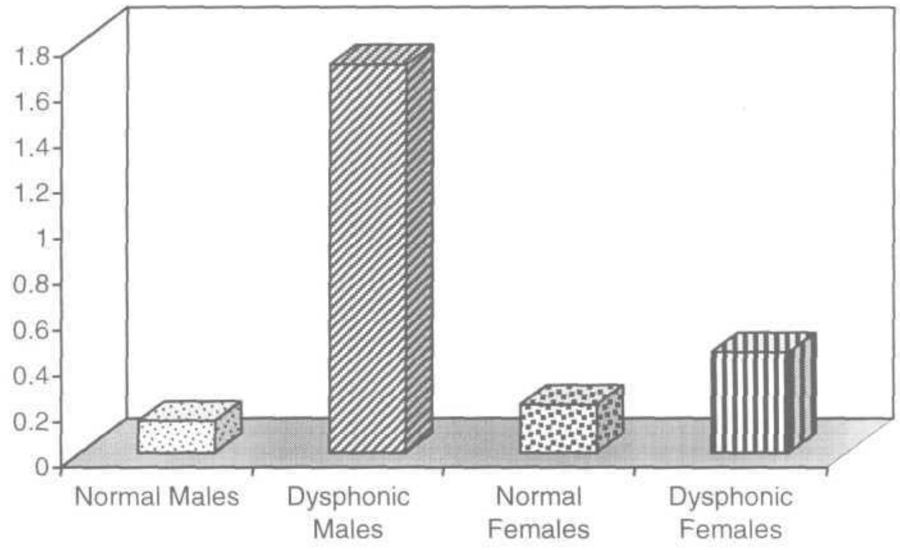
Table 19.1 : Showing comparison between groups. '+' indicates presence of significant difference at 0.05 level

Group	Significance
NM Vs NF	-
NM Vs DM	-
DM Vs DF	-
DF Vs NF	-

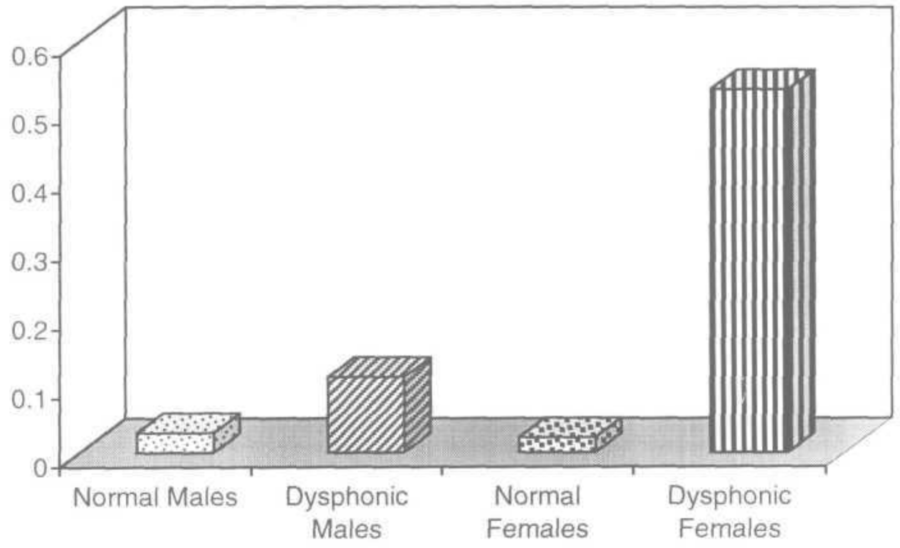
From Table 19, mean values are greater in dysphonics than normals.

However from Table 19.1, we can say that there is no significant difference between normal males and normal females, normal males and dysphonic males, dysphonic males and dysphonic females and dysphonic females and normal females. No significant difference was seen across session, for both males and females in normals.

GRAPH 19 : MEANS OF NOISE TO HARMONIC RATIO (NHR)



GRAPH 20 : MEANS OF VOICE TURBULENCE INDEX (VTI)



Thus the hypothesis stating that there is no significant difference across sessions is accepted. The hypothesis stating that there is no significant difference between normals and dysphonics is accepted in case of males and females. The hypothesis stating that there is no significant difference between males and females is also accepted.

The results of the present that NHR increases in dysphonics coincides with the findings of Kitojima (1981), Anitha (1994) and Das (1996).

Studies have shown a positive correlation between perceived degree of hoarsness or roughness and HNR (Emanuel, 1979; Yumoto Nataraja 1981; Bhat 1981; Gould and Baer, 1982; and Yumoto, 1983).

20. Voice Turbulance Index (VTI)

The results of statistical analysis for this parameter are shown in Table 20, Table 20.1 and Graph 20.

Table 20 : Showing mean, SD and range for VTI

Groups	Mean	S.D.	Range
NM	0.029	0.01	0.01-0.14
DM	0.11	0.17	0.01-1.04
NF	0.022	0.01	0.0-0.05
DF	0.53	3.14	0.01-21.15

Table 20.1 : Showing comparison between groups.'+' indicates presence of significant difference

Group	Significance
NM Vs NF	+
NM Vs DM	+
DM Vs DF	-
DF Vs NF	+

Fran Table 20, mean values of VTI is greater in dysphonics than normals SD and range is also greater in dysphonics.

From Table 20.1, we can say that there is significant difference between normals and dysphonics, in both males and females. Significant difference is also noted between normal males and normal females. No significant difference is noted between dysphonic males and dysphonic females.

There was no significant difference across trials in both males and females. This indicates that VTI is a relatively stable measure.

Thus the hypothesis stating that there is no significant difference across sessions is accepted. The hypothesis stating that there is no significant difference between normals and dysphonics is rejected for both males and females. The hypothesis stating that there is no significant difference between males and females is accepted in case of dysphonics and rejected for normals.

Findings of this study is consistent with the findings of Anitha (1994). The mean values are greater in dysphonics because VTI mostly correlates with the turbulence caused by incomplete or loose adduction of the vocal folds. VTI analyses high frequency components to extract an acoustic correlate to "breathiness".

21. Soft Phonation Index (SPI)

The values of SPI are briefed down in Table 21, Table 21.1 and Graph 21.

Groups	Mean	S.D.	Range
NM	37.42	32.07	0.01-210
DM	17.31	19.88	0.03-79.68
NF	51.72	54.53	3.34-264
DF	15.61	16.34	0.52-74.49

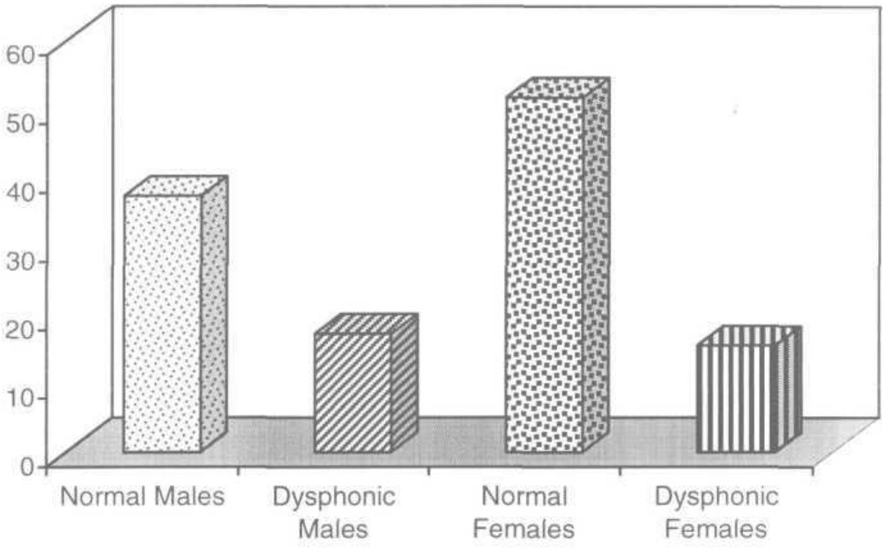
Table 21.1 : Showing comparison between groups. '+' indicates presence of significant difference

Group	Significance
NM Vs NF	+
NM Vs DM	+
DM Vs DF	-
DF Vs NF	+

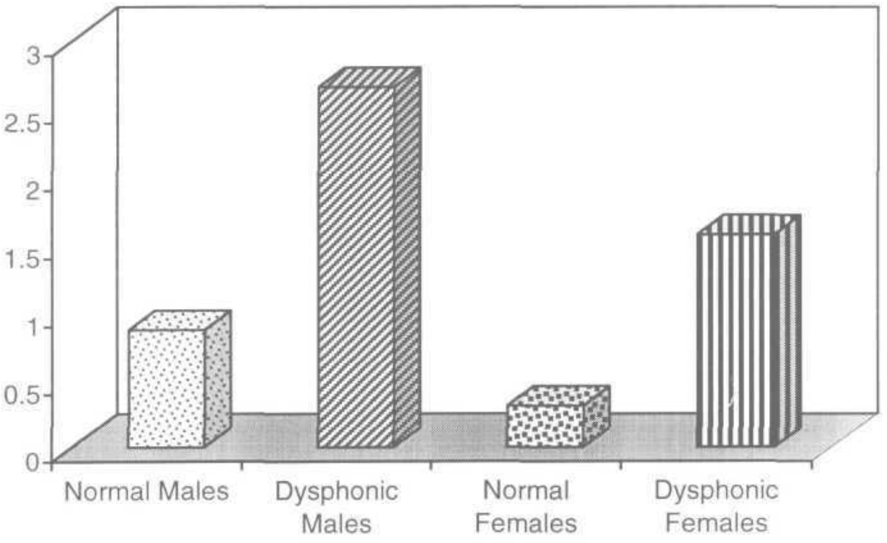
On inspection of Table 21, we see that mean SPI is lower in dysphonics than normals, with respect to both males and females.

From Table 21.1, we see that there is significant difference between normals and dysphonics, in both males and females. Significant difference was observed between normal males and normal females. No significant difference was noticed between dysphonics males and dysphonic females.

GRAPH 21 : MEANS OF SOFT PHONATION INDEX (SPI)



GRAPH 22 : MEANS OF F_o TREMOR INTENSITY INDEX (FTRI)



There was significant differences observed in this parameter across sessions indicating SPI values varied across sittings for a same individual.

Thus the hypothesis stating that there is no significant difference across sessions is rejected. The hypothesis stating that there is no significant difference between normals and dysphonics is rejected for both males and females. The hypothesis stating that there is no significant difference between males and females is rejected for normals and accepted in case of dysphonics.

The above results can be discussed as follows :

SPI indicates 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. A person may naturally speak with a softer "attack" and hence have an allevated SPI. Similarly, patients with "pressed" phonation may likely have a "normal" or "lower" 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.

SPI values increases with psychological stress and also the amplitude of the sustained phonation. If the subject phonates softly, SPI may be high. These two factors could have contributed for the variability of SPI across sessions.

22. Frequency tremor intensity index (FTRI)

The results of statistical analysis for this parameter are shown in Table 22, Table 22.1 and Graph 22.

Table 22 : Showing mean, SD and range for FTRI

Groups	Mean	S.D.	Range
NM	0.87	3.63	0-33.8
DM	2.66	7.92	0-51.33
NF	0.31	0.16	0-0.89
DF	1.57	2.11	0.06-9.53

Table 22.1 : Showing comparison between groups. '+' indicates significant difference at 0.05 level

Group	Significance
NM Vs NF	-
NM Vs DM	-
DM Vs DF	-
DF Vs NF	+

On inspection of Table 22, we can see the mean values are greater in dysphonics than normals. SD and range are also greater in dysphonic males and dysphonic females than normal males and normal females respectively. Mean values are greater in males than in females. This coincides with the findings reported by Gopal Krishna (1995) for normals.

Table 22.1 shows that there is no significant difference between males and females, in both normals and dysphonics. No significant difference was seen between normal males and dysphonic males. Significant difference was noted between normal females and dysphonic females.

No significant difference was also noted across sessions for FTRI. Thus the hypothesis stating that there is no significant difference across sessions is accepted. The hypothesis stating that there is no significant difference between normals and dysphonics is accepted in case of males and rejected in case of females. The hypothesis stating that there is no significant difference between males and females is also accepted.

Mean values are greater in dysphonics than normals. This is an agreement with the findings of Anitha (1994) and Das (1996).

23. Amplitude tremor intensity index (ATRI)

The results of statistical analysis for this parameter are summarised in Table 23, Table 23.1 and Graph 23.

Table 23 : Showing mean, SD and range for ATRI

Groups	Mean	S.D.	Range
NM	2.69	3.02	0-18.41
DM	2.96	4.09	0-16.48
NF	1.66	1.63	0-6.77
DF	3.39	4.85	0-15.14

Table 23.1 : Showing comparison between groups. +' indicates significant difference at 0.05 level

Group	Significance
NM Vs NF	+
NM Vs DM	-
DM Vs DF	-
DF Vs NF	+

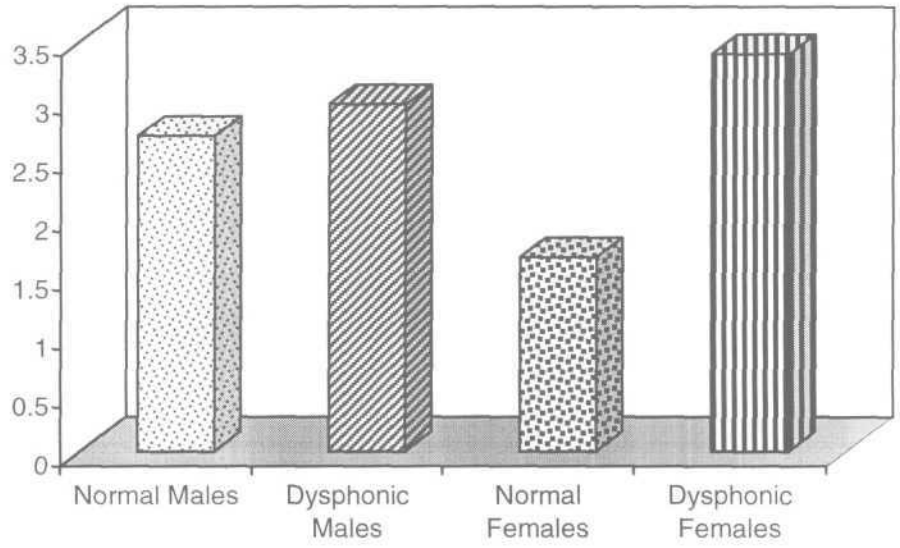
Table 23 shows that dysphonics have greater mean values than normals. Greater mean values in normal males as compared to normal females as seen in table 23 was also observed in the experiment conducted by Gopal Krishna (1995).

Table 23.1 reveals that there is significant difference between normal males and normal females. No significant difference is noted between dysphonic males and dysphonic females. No significant difference was apparent between dysphonic males and normal males. Significant difference was noted between normal females and dysphonic females.

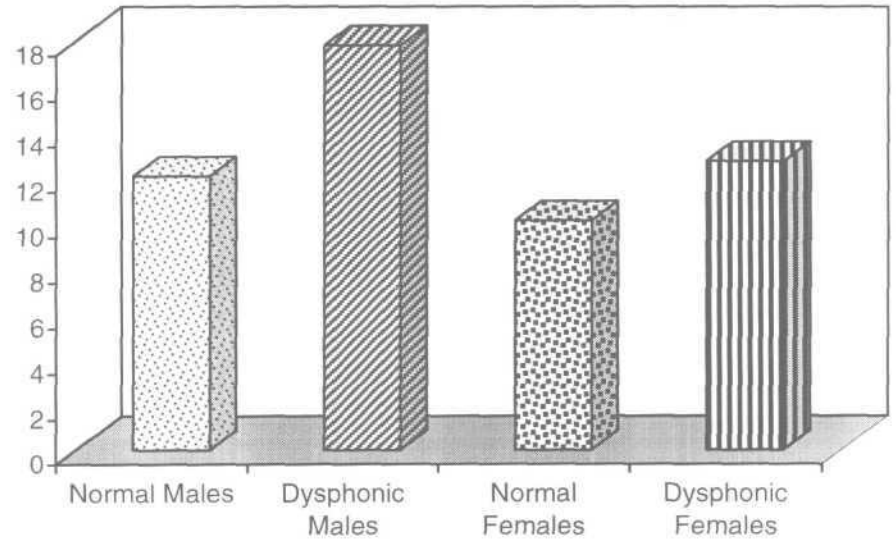
This parameter was found to be stable across sessions in both males and females.

Thus the hypothesis stating that there is no significant difference across sessions is accepted. The hypothesis stating that there is no significant difference between normals and dysphonics is accepted in case of males and rejected in case of females. The hypothesis stating that there is no significant difference between males and females is rejected for normals and accepted in case of dysphonics.

GRAPH 23 : MEANS OF AMPLITUDE TREMOR INTENSITY INDEX (ATRI)



GRAPH 24 : MEANS OF DEGREE OF VOICE BREAKS (DVB)



Dysphonics showed greater mean values than normals. This findings is in agreement with the findings of Anitha (1994) for phonation of vowels and Das (1996).

24. Degree of voice breaks (DVB)

The results are shown in Table 24, Table 24.1 and Graph 24.

Table 24 : Showing mean, SD and range for DVB

Groups	Mean	S.D.	Range
NM	12.05	21.0	0-66.9
DM	17.8	26.9	0-82.7
NF	10.09	17.8	0-56
DF	12.7	20.8	0-81.4

Table 24.1 : Showing comparison between groups. '+' indicates significant difference at 0.05 level

Group	Significance
NM Vs NF	-
NM Vs DM	-
DM Vs DF	-
DF Vs NF	-

From the Table 24, we can say that mean values are greater in dysphonics than normals.

However Table 24.1 shows that there is no significant difference between the groups compared.

No significant difference was observed across trails in normal males and normal females.

Thus the hypothesis stating that there is no significant difference across sessions is accepted. The hypothesis stating that there is no significant difference between normals and dysphonics is accepted for both males and females. The hypothesis stating that there is no significant difference between males and females is also accepted.

Greater means was observed in dysphonics than normals. This is in agreement with the findings of Anitha (1994) and Das (1996).

Greater mean values in dysphonics is because of their inability to sustain uninterrupted voicing during sustained phonation.

25. Degree of sub-harmonic components (DSH)

The results are shown in Table 25, and Graph 25 for this parameter.

Table 25 : Showing mean, SD and rane of DSH

Groups	Mean	S.D.	Range
NM	0.15	1.04	0-8.14
DM	3.24	8.16	0-45.9
NF	1.03	4.14	0-28.7
DF	2.56	6.15	0-31.8

Table 25.1 : Showing comparison between groups. '+' indicates significant difference at 0.05 level

Group	Significance
NM Vs NF	-
NM Vs DM	+
DM Vs DF	-
DF Vs NF	-

Table 25 shows that mean values for this parameter is greater in dysphonics than normals. SD and range is also greater for the dysphonic group than normals.

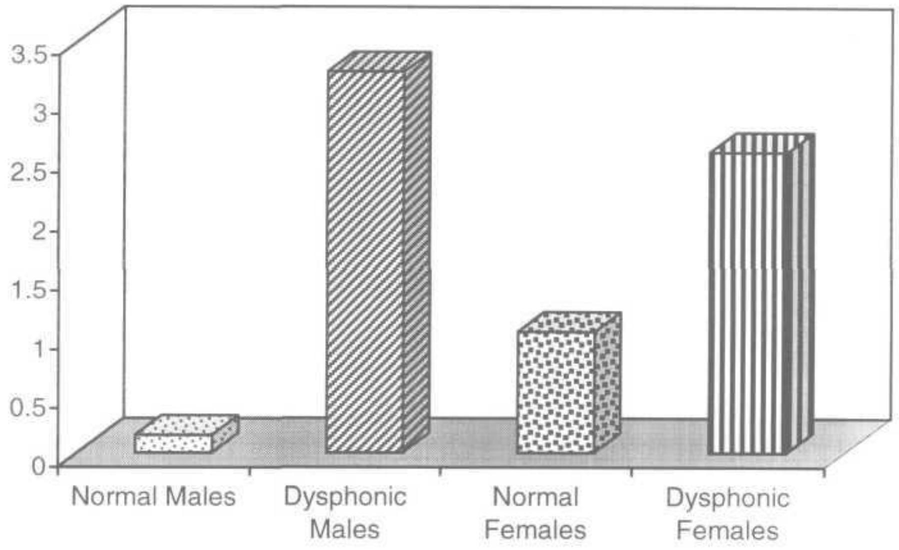
From Table 25.1, we can say that there is no significant difference between males and females in both normals and dysphonics.

No significant difference was noted across sessions for this parameter in both males and females.

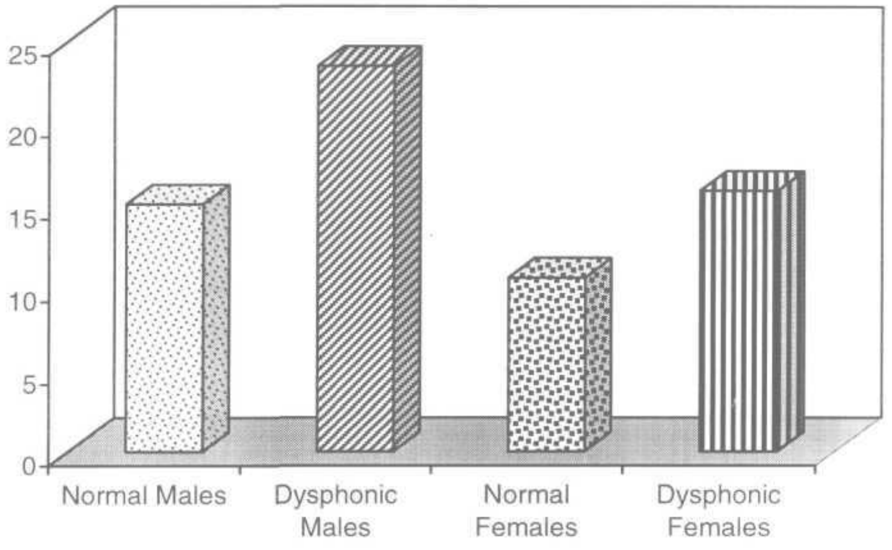
Thus the hypothesis stating that there is no significant difference across sessions is accepted. The hypothesis stating that there is no significant difference between normals and dysphonics is accepted in case of females and rejected for males. The hypothesis stating that there is no significant difference between males and females is accepted in case of dysphonics and normals.

The greater mean values in dysphonics agrees with the findings of Anitha (1994) and Das (1996).

GRAPH 25 : MEANS OF DEGREE OF SUB-HARMONICS (DSH)



GRAPH 26 : MEANS OF DEGREE OF VOICELESS (DUV)



26. Degree of voiceless (DUV)

The results are summarised in Table 26, Table 26.1 and Graph 26 for this parameter.

Table 26 : Showing mean, SD and range for DUV

Groups	Mean	S.D.	Range
NM	15.03	24.66	0-72.5
DM	23.47	30.47	0-82.2
NF	10.57	18.44	0-57.5
DF	15.86	23.81	0-81.9

Table 26.1 : Showing comparison between groups. '+' indicates significant difference at 0.05 level

Group	Significance
NM Vs NF	-
NM Vs DM	-
DM Vs DF	-
DF Vs NF	-

Table 26 shows that the mean values are greater in dysphonics than normals.

Table 26.1 reveals that there is no significant difference between any of the groups compared, that is normal males and normal females; normal males and dysphonic males, dysphonic males and dysphonic females and dysphonic females and normal females.

No significant difference was noted across sessions for this parameter in both males and females.

Thus the hypothesis stating that there is no significant difference across sessions is accepted. The hypothesis stating that there is no significant difference between normals and dysphonics is accepted. The hypothesis stating that there is no significant difference between males and females is also accepted.

Greater mean values in dysphonics correlates with the findings of Anitha (1994) and Das (1996).

The above results can be discussed as follows :

DUV measures the ability of the voice to sustain uninterrupted voicing. Dysphonics showed increased DUV because of their inability to maintain a constant pitch and uninterrupted voicing due to different vocal pathologies.

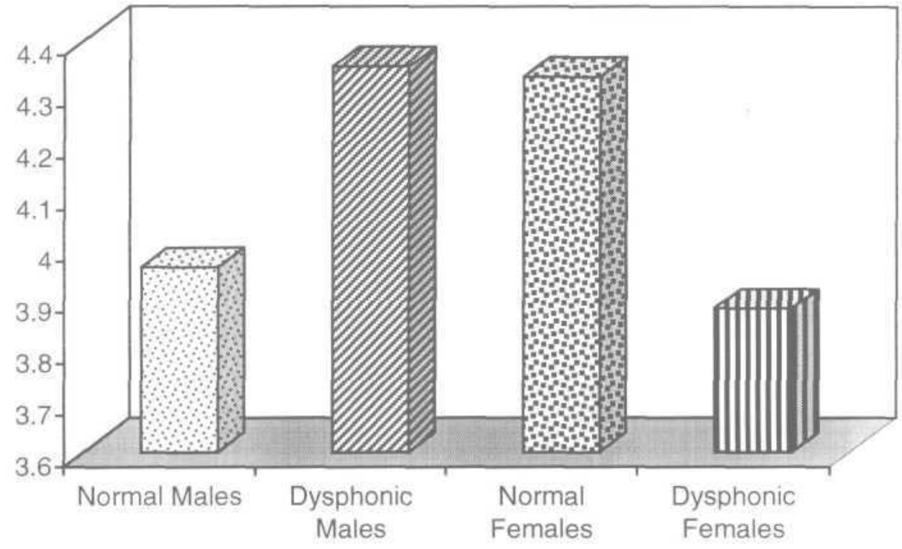
27. Number of Voice Breaks (NVB)

The results of statistical analysis of this parameter are shown in Table 27, Table 27.1 and Graph 27.

Table 27 : Showing mean, SD and range for NVB

Groups	Mean	S.D.	Range
NM	3.96	6.97	0-22
DM	4.35	6.78	0-25
NF	4.33	7.84	0-27
DF	3.88	6.25	0-23

GRAPH 27 : MEANS OF NUMBER OF VOICE BREAKS (NVB)



GRAPH 28 : MEANS OF NUMBER OF SUB-HARMONIC SEGMENTS (NSH)

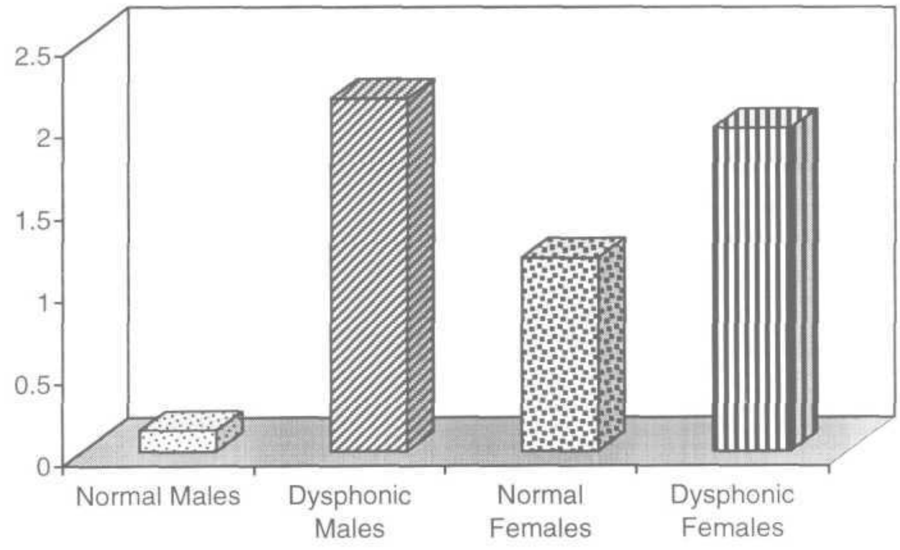


Table 27.1 : Showing comparison between groups. '+' indicates significant difference at 0.05 level

Group	Significance
NM Vs NF	-
NM Vs DM	-
DM Vs DF	-
DF Vs NF	-

From Table 27 and 27.1, we see that NVB does not differentiate normals from dysphonics in both males and females. It does not differentiate males and females, in both normals and dysphonics.

There was no significant difference observed across session for this parameter.

Thus the hypothesis stating that there is no significant difference across sessions is accepted. The hypothesis stating that there is no significant difference between normals and dysphonics is accepted in case of both males and females. The hypothesis stating that there is no significant difference between males and females is also accepted.

28. Number of sub-harmonic segments (NSH)

The results of statistical analysis for this parameter is shown in Table 28, Table 28.1 and Graph 28.

Table 28 : Showing mean, SD and range for NSH

Groups	Mean	S.D.	Range
NM	0.13	0.9	0-7
DM	2.15	6.5	0-40
NF	1.18	5.6	0-48
DF	1.97	4.8	0-27

Table 28.1 : Showing comparison between groups. '+' indicates presence of significant difference

Group	Significance
NM Vs NF	-
NM Vs DM	+
DM Vs DF	-
DF Vs NF	-

Table 28 shows that the mean values are greater in dysphonics than normals for this parameter.

Table 28.1 shows that there is significant difference between normal males and dysphonic males. No significant difference was observed between males and females in both normals and dysphonics. No significant difference was observed between normal females and dysphonic females. No significant difference was observed across trials for this parameter in both males and females.

Thus the hypothesis stating that there is no significant difference across sessions is accepted. The hypothesis stating that there is no significant difference between normals and dysphonics is accepted in case of females and rejected for males. The hypothesis stating that there is no significant difference between males and females is also accepted.

The greater mean values for this parameter is supported by the findings of Anitha (1994) and Das (1996).

The above findings can be discussed as follows : The mean values of NSH for dysphonic group were higher than in 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 mean value of number of subharmonic segments.

29. Number of unvoiced segments (NUV):

The mean, S.D and range are shown in Table 29 and Graph 29 shows the mean. The results of 'T' test are shown in Table 29.1.

Table 29 : Showing mean, SD and range for NUV

Groups	Mean	S.D.	Range
NM	0.72	2.7	0-19
DM	3.04	10.5	0-57
NF	0.28	1.09	0-8
DF	6.88	14.03	0-55

GRAPH 29 : MEANS OF NUMBER OF UNVOICED SEGMENTS (NUV)

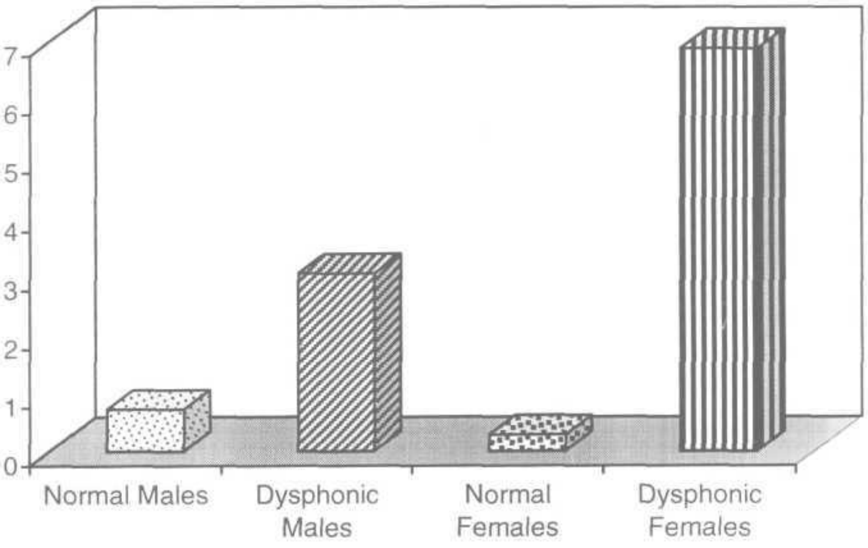


Table 29.1 : Showing comparison between groups. '+' indicates presence of significant difference

Group	Significance
NM Vs NF	-
NM Vs DM	+
DM Vs DF	-
DF Vs NF	-

Table 29 shows that the mean values are greater in dysphonics than normals.

Table 29.1 shows that there is no significant difference between males and females, in both normals and dysphonics. T test showed significant difference across trials for this parameter.

Thus the hypothesis stating that there is no significant difference across sessions is rejected. The hypothesis stating that there is no significant difference between normals and dysphonics is accepted in case of females and rejected in case of males. The hypothesis stating that there is no significant difference between males and females is also accepted.

The mean values for NUV are greater in dysphonics. This agrees with the findings of Anitha (1994) and Das (1996). This could be because of the inability of the dysphonics to sustain uninterrupted voicing during sustained phonation.

Hence, the parameters used for this study could efficiently differentiate normal and dysphonic group.

These data was fed and preprocessed using a filter, to analyse the data in terms of their even distribution, skewness and kurtosis.

After the data was pre-processed using the filter the following fields were taken to be fed in to the nueral network, for processing and classification.

1. Sex of the subject
2. Average fundamental frequency (Fo)
3. Highest fundamental frquency (Fhi)
4. Lowest fundamental frequency (Flo)
5. Standard deviation of Fo (STD)
6. Phonatory Fo-range in semitones (PFR)
7. Fo Tremor frequency (Fftr)
8. Amplitude Tremor frequency (Fatr)
9. Absolute Jitter (Jita)
10. Jitter percentage (Jitt)
11. Relative average perturbation (RAP)
12. Pitch perturbation quotient (PPQ)
13. Smoothened pitch perturbation quotient (SPPQ)
14. Fundamental frequency variation (vFo)
15. Shimmer in dB (ShdB)
16. Shimmer in percentage (Shim)
17. Amplitude perturbation quotient (APQ)
18. Smoothened amplitude perturbation quotient (SAPQ)
19. Peak - amplitude variation (vAM)
20. Noise to harmonic ratio (NHR)
21. Voice turbulence index (VTI)

22. Soft phonation index (SPI)
23. Fo-tremor intensity index (FTRI)
24. Amplitude - tremor intensity index (ATRI)
25. Degree of voice breaks (DVB)
26. Degree of sub-harmonics (DSH)
27. Degree of vioceless (DUV)
28. No. of voice breaks (NVB)
29. Number of sub-harmonic segments (NSH)
30. Number of unvoiced segments (NUV)

The data was trained using the Multi layer perceptin (MLP) Neural network model by using the data of 90 subjects. In order to obtain the best possible model to identify voice disorders the neural network was individually trained by adjusting the following parameters of the neural network (MLP).

1. The learning algorithms functions of the neural network-The two learning algorithmic - conjugate gradient and steepest descent were taken up and the neural network was trained in each of these algorithms to find which of the learning algorithm gave the best results.
2. The number of hidden layers - The number of hidden layers was varied to calculate as to which option could give the best results.
3. The number of nodes in each hidden layer - The nodes in each hidden layer was adjusted sequentially to findout which combination gave the best results.

The results indicated the following :

1. The learning algorithm 'steepest descent' gave better results than the learning algorithm 'Conjugate gradient'.

2. By increasing the hidden layers to four, the neural network could classify the data samples more effectively.
3. The best performance was obtained by using two nodes in the 1st hidden layer and one node in the 2nd, 3rd and 4th hidden layer in case of 'steepest descent' learning algorithm function.

By using the best combination in terms of hidden layers and number of nodes in each layer in the learning algorithm 'steepest descent' the network model could correctly predict the voice disorder by 90%.

During the training phase a total of 76 datas were trained. Out of the total datas, 70 of them were identified correctly. In terms of percentage the network could identify voice disorders accurately 92% of times.

Out of the 50 normal subjects in training data the network could identify 45 of them correctly.

In the validation data total of 26 samples of dysphonics were taken. Among them 24 of them was identify correctly.

Out of the 10 normals taken in the validation data, the network was able to identify 9 of the subject correctly.

The test data consisted of 4 samples. It was seen that the network identify pathologies 100% of the times.

Finally the run data consisted of 20 subjects, which included both dysphonics and normals.

It was seen that the network could identify 90% of normal subjects correctly and 90% of dysphonics correctly.

It is seen that the percentage of correct identification vary during the training phase, the validation phase, in the test phase and during the run data. This variation was basically because the values of parameters obtained by different subjects involved in different stages varied.

The network thus proved to be capable of classifying the normals and dysphonics based on the acoustic paratmers that were used.

Better results can be obtained by increasing the number and varieties of cases.

Thus the objective of the study to classify the normals and dysphonics based on acoustic parameters using artificial neural networks has been achieved. The study has opened up new possibilities of computer application in the field of voice disorders.

SUMMARY AND CONCLUSION

"Voice has been defined as the laryngeal modulation of the pulmonary air steam, which is further modified by the configuration of the vocal tract" (Micheal and Wendahl, 1971).

The present study was designed to objectively classify dysphonics from normals with the help of Artificial Neural Network using the 29 acoustic parameters extracted using Multi Dimensional Voice program.

The Parameters that were studied include :

I FREQUENCY PARAMETERS

1. Average fundamental frequency
2. Highest fundamental frequency
3. Lowest fundamental frequency
4. Standard deviation of Fo
5. Phonatory Fo range in semitones
6. Fo tremor frequency
7. Absolute Jitter
8. Jitter percent
9. Relative perturbation quotient
10. Pitch perturbation quotient
11. Smoothened pitch perturbation quotient
12. Fundamental frequency variation
13. Fo Tremor intensity index (FTRI)

II INTENSITY PARAMETERS

14. Amplitude tremor frequency
15. Shimmer in dB
16. Shimmer percent
17. Amplitude perturbation quotient
18. Smoothened Amplitude perturbation quotient
19. Peak amplitude variation
20. Amplitude tremor intensity index.

III OTHER PARAMETERS

21. Noise to harmonic ratio
22. Voice turbulence index
23. Soft phonation index
24. Degree of voice breaks
25. Degree of sub harmonics
26. Degree of voiceless
27. Number of voice breaks
28. Number of sub-harmonic segments
29. Number of unvoiced segments.

The normal group consisted of thirty males and thirity females in the age range of 18-25 years. Among them five males and five females were assessed on five sessions to find the stability of the parameters measured. The dysphonic group consisted of fifteen males and fifteen females in the age range of 20 - 45 years.

Most of the parameters were stable across sessions except,

1. Highest fundamental frequency.
2. Shimmer in dB.
3. Shimmer in percent.
4. Number of unvoiced segments.

The following parameters showed significant difference between normals and dysphonics:

1. Average fundamental frequency.
2. Highest fundamental frequency.
3. Standard deviation of F_o.
4. Phonatory F_o - range in semitones.
5. F_o tremor frequency.
6. Amplitude tremor frequency.
7. Absolute Jitter.
8. Relative average perturbation.
9. Pitch perturbation quotient.
10. Smoothed pitch perturbation quotient.
11. Fundamental frequency variation.
12. Shimmer in dB.
13. Shimmer in percent.
14. Voice turbulence index.
15. Soft phonation index.
16. Degree of sub harmonics.
17. Number of sub harmonics.
18. Number of unvoiced segments.

Parameters which showed significant difference between normals and dysphonics, in females were:

1. Average fundamental frequency.
2. Lowest fundamental frequency.
3. Standard deviation of F_o.
4. Phonatory F_o range in semi tones.
5. F_o tremor frequency.
6. Absolute Jitter.
7. Relative average perturbation.
8. Pitch perturbation quotient.
9. Fundamental frequency variation.
10. Shimmer in percent.
11. Amplitude perturbation quotient.
12. Smoothed amplitude perturbation quotient.
13. Voice turbulence index.
14. Soft phonation index.
15. F_o tremor intensity index.
16. Amplitude tremor intensity index.

The input data consiting of the acoustic parameters and the output data consisting of the diagnosis were fed to the Neural network. The Neural network used for this study was Multi Layer Perception (MLP). After training the data, the neural network model was tested on the test data and it was found that the neural model could classify normal and abnormal voice correctly by 100%.

CONCLUSIONS

1. Most of the acoustic parameters extracted from MDVP are stable and reliable.
2. These acoustic parameters are useful in differentiating normals from dysphonics.
3. The Neural network can be trained using the acoustic parameters of voice to differentiate normals from dysphonics.
4. The reliability and validity of the classification could be improved by altering the hidden layers and nodes of the neural network.

Thus the aim of the present study to objectively classify normals and dysphonics was achieved.

REFERENCES

Adams, L. (1981). Cited in Anitha. (1994). "Multi Dimensional Analysis of Voice Diorders". Unpublised Masters Dissertation, University of Mysore, Mysore.

Anderson, V. (1961). "Training the speaking voice", Oxford Univ., N.Y.

Anderson, V. (1961). Cited in Nataraja .N.R (1986). Differential Diagnosis of Dysphonias. Unpublished Doctral thesis, University of Mysore, Mysore.

Anitha, V. (1994). "Multi dimensional Analysis of Voice Disorders", Unpublished Master's Dissertation, University of Mysore.

Aparna, M. (1996). "Multidimensional Analysis of Voice in the hearing impaired", Unpublished Master Dissertation,Univeristy of Mysore.

Aronson. (1980).cited in Anitha (1994) 'Multi Dimensional Analysis of Voice Diorders". Unpublised Masters Dissertation, University of Mysore, Mysore.

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

Baer, T. (1980). "Vocal jitter - A neuromuscular explanation". Transcripts of the Eighth symposium of the care of the professional voice, voice foundation, New York.

Berry, M.F., Eisenson, J. (1956). "Speech disorders - Principles, Practice of therapy, Appleton - Century Crofts, N.Y.

Binu, R. (1998). "An Objective classification of Voice Disorders using Artificial Neural Networks", Unpublished Master Dissertation, University of Myosre.

Biran, A. (1995). "Multi Dimensional Voice Analysis in Children", Unpublished Master Dissertation, Univeristy of Mysore.

Bolt. (1992). Cited in Binu, R. (1998). "An objective classification of voice disorders using Artifical Neural Network". Unpublished Masters Dissertation, University of Mysore, Mysore.

Boone, D. (1977). 'The Voice and Voice Therapy", 2nd edn. Prentice Hall, Inc., Englewood Cliffs, N.J.

Boone, D.E. (1971). "The voice and voice therapy (3rd ed.), Prentice - Hall Inc., Englewood Cliffs, N.J.

Boone, D.E. (1983). "The voice and Voice therapy", (3rd ed), Prentice Hall Inc., Engle Wood Cliffs, N.J.

Bowler, N.W. (1964). "A fundamental frequency analysis of harsh vocal quality". Speech Monograph, 31, 128-134.

Brackett, I.P. (1971). "Parameters of voice quality". In L.E. Travis (Ed.), Handbook of speech pathology and Audiology, Appleton Century - Crofts, N.Y.

Broad, D. (1979). "The new theories of vocal fold vibration". In N.J. Lass (Ed.), Adwances in basic research and practice, vol 2, NewYork Academic press, Newyork.

Carahart, R. (1941). "The spectra of model larynx tones". Speech Monograph, 8, 76-84.

Carhart, R. (1938). "Infra glottal resonance and a cushion pipe". Speech Monograph, 5, 65-90.

Chandrashekhar, K.R. (1987). "Electroglottography in dysphonics". Unpublished master's Dissertation, University of Mysore.

Cooper, M. (1974). "Spectrographic analysis of fundamental frequency and hoarsensess before and after vocal rehabilitation". Journal of Speech and Hearing Disorders 39, 286-296.

Cowan, J. (1936). "Pitch and Intensity characterstics of stage speech". Arch. Speech. Supl.l, 7-85.

Crystal and Jackson. (1970). "Methodology and Results on Laryngeal Disorder Detection through Spectral Analysis". Final report, contract PH-86-68-192, Singatron. Inc., Lexington, Massachusetts.

Curry, E.T., (1940). The pitch characterstics of the adolescent male voice". Speech Monograph, 7, 48-52.

Curtis, J. (1968). "Acoustics of speech production and Nasalization" in spriestersbach, D.C., Sherman, D (ed.), Cleft Palate and communciation, Academic press, New York.

Davis, H. (1935). cited in Stevens, S.S., Davis, H. (1938) "Hearing its psychology and physiology", Chapman Hall, N.Y.

Davis, S.B. (1976). "Computer evaluation of laryngeal pathology based on inverse filtering of speech". SCRL Monograph, 13, California.

Deal, R.E., Emanuel, F.W. (1978). "Some waveform and spectral features of vowel Roughness". Journal of Speech and hearing research, 21(2), 250-264.

Delgutte. (1984). Cited in Binu, R. (1998). "An objective classification of voice disorders using Artifical Neural Network". Unpublished Masters Dissertation, University of Mysore, Mysore.

Donald, H. (1949). Cited in Binu, R. (1998). "An objective classification of voice disorders using Artifical Neural Network". Unpublished Masters Dissertation, University of Mysore, Mysore.

Emanuel, F.W., White Head, R.L. (1979). "Harmonic levels and vowel roughness", Journal of Speech and Hearing Research, 22(4), 829-840.

Emerick, L, Hatten, J. (1974). Cited in Nataraja .N.P. (1986). "Differential Diagnosis of Dysphonias". Unpublished Doctral thesis, University of Mysore, Mysore.

Emrickson, C.I. (1959). "The basic factors in the human voice". Psy. Monographs, Univ., IOWA Studies in Psychology, 10, 86-112.

Enanuel, F.W., Sansone, FE. (1969). Cited in Anitha (1994) "Multi Dimensional Analysis of Voice Diorders". Unpublised Masters Dissertation, University of Mysore, Mysore.

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

Fairbanks, G. (1940). "Recent Experimental Investigations of Vocal Pitch in Speech". Journal of Acoustical Society of America, 11, 457-466.

Fant, G. (1960). "Acoustic theory of speech production". Mournton and Co., S. Gravehage.

Farely, R.G. (1994). "Control of Voice of Fo by an artificial neural network". Journal of Acoustic Society and America, 105, 1411-1422.

Fitch, J.L., Holbrook, A. (1970). "Modal vocal fundamental frequency of Young adults". Archives of Otolaryngology, 92, 379-382.

Fletcher, W. (1959). Cited in Nataraja .N.P. (1986). "Differential Diagnosis of Dysphonias". Unpublished Doctral thesis, University of Mysore, Mysore.

Freeman (1985). Cited in Rashmi. "Acoustic Aspects of the Speech of Children". Unpublished Masters Dissertation, University of Mysore, Mysore.

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.

Gilbert, H.R. and Cambell, M.I. (1980). "Speaking fundamental frequency in three groups of hearing impaired individuals". Journal of Communication Disorders, 13, 195-205.

Gopal Krishna, S. (1995). "Susceptibility Criteria for Vocal Fatigue", Unpublished Master Dissertation, Univeristy of Mysore.

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.

Gould, W.J., Okamura, H. (1974). "Static lung volumes in singers". Annals of otorhinolaryngology, 82, 89-94.

Gray, G.W, Wise, CM. (1959). "The bases of speech", Harper and Row Publishers, New York.

Greene, M.C.L. (1964). 'The Voice and its Disorders". Mitman Medical, London.

Greene, M.C.L. (1972)."Multi Dimensional Analysis of Voice Diorders". Unpublised Masters Dissertation, University of Mysore, Mysore.

Hanley (1972). Cited in Hollien, H., Shipp, T. (1972)."Speaking fundamental frequency and chronologic age in Males". Journal of Speech and Hearing Research, 15(1), 155-159.

Hanson, D.G., Gerrat, B.R., Ward, H.R (1983). "Glottographic measurement of vocal dysfunction -A preliminary report". Annals of Otorhinolaryngology, 92(5), 413-4190.

Haykins, S. (1994). "Neural Network - A comprehensive Foundation", New York, Macmillan.

Haykins, S. (1995). Cited in Binu, R. (1998). "An objective classification of voice disorders using Artifical Neural Network". Unpublished Masters Dissertation, University of Mysore, Mysore.

Heiberger, V.L., Horii, Y. (1982). "Jitter and Shimmer in sustained phonation". cited in N.J.Lass,(ed.),Speech and language advance in basic research and practice. Academic Press, New York.

Higgins, B.M., Saxman, H.J. (1989). "A comparision of intra subject variation across sessions of three vowel frequency perturbation indices". Journal of Acoustic Society of America. 86(3), 911-916.

Hilhour. (1972). Cited in Hollien, H., Shipp, T. (1972). "Speaking fundamental frequency and chronologic age in Males". Journal of Speech and Hearing Research, 15(1), 155-159.

Hirano, M. (1981). "Clinical examination of voice". Disorders of human communication, 5, Springer, Wien.

Hirano, M. (1981). Cited in Nataraja .N.R (1986). "Differential Diagnosis of Dysphonias". Unpublished Doctral thesis, University of Mysore, Mysore.

Hirano. M. (1989). "Multi Dimensional Analysis of Voice Diorders". Unpublised Masters Dissertation, University of Mysore, Mysore.

Hollen, Jackson. (1972). Cited in Hollien, H., Shipp, T. (1972). "Speaking fundamental frequency and chronologic age in Males". Journal of Speech and Hearing Research, 15(1), 155-159.

Hollien, H., Shipp, T. (1972). "Speaking fundamental frequency and chronologic age in males". Journal of Speech and Hearing Research, 15 (1), 155-159.

Horii, Y. (1980). "Vocal Shimmer in sustained phonation". Journal of Speech and Hearing Research, 23(1), 202-209.

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

Husson, R. (1950). Cited in Nataraja .N.R (1986). "Differential Diagnosis of Dysphonias". Unpublished Doctral thesis, University of Mysore, Mysore.

Indira, N. (1982). "Analysis of infant cries", Unpublished Masters Dissertation, University of Mysore.

Irwin, J.V. (1966). "Multi Dimensional Analysis of Voice Diorders". Unpublished Masters Dissertation, University of Mysore, Mysore.

Ishizake, K., Matsudaira, M. (1972). Third Mechanical considerations of vocal fold vibration, SCRL monograph, 8, Santhabarbara, California.

Iwata, S., Von Leden, H. (1970). "Phonation quotient in patients with laryngeal diseases". Folia Phoniatrica, 22, 117-128.

Jayaram, K. (1975). "An attempt at differential diagnosis of dysphonic". Master's Dissertation, University of Mysore.

Johnson, W., Michel, J.F. (1969). "The effect of selected vowels on laryngeal jitter". A.S.H.A., 11, 96-109.

Juson. L.S.V, Weaver. AT. (1965). "Voice science" (2nd Ed.), Meredith publishing company, U.K.

Kasuya. (1985). Cited in Rashmi. "Acoustic Aspects of the Speech of Children". Unpublished Masters Dissertation, University of Mysore, Mysore.

Kasuya, Ogawa, Mashima and Ebihara. (1986). "Multi Dimensional Analysis of Voice Diorders". Unpublised Masters Dissertation, University of Mysore, Mysore.

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, Hirano, M. (1982). "Sound spectrographic analysis of the voice of patients with recurrent laryngeal nerve paralysis", Folia Phoniatrica, 34, 124-133.

Kitajima, K. (1981). "Quantitative evaluation of the noise level in the pathologic voice". Folia Phoniatrica, 115-124.

Kitajima, K., Gould, WJ. (1976). "Vocal Shimmer in sustained phonation of normal and patologic voice". Annals of Otology Rhinology Laryngology, 85, 377-381.

Koike, Y. (1969). "Vowel amplitude modulations in patients with laryngeal diseases". Journal of Acoustical Society of America, 45, 839-844.

Koike, Y (1973). "Application of some acoustic measures for the evaluation of laryngeal dysfunction". Stud. phonal, 7, 19-23.

Kushal Raj, P. (1983). "Acoustic Analysis of Speech of Children". Unpublished Master's Dissertation, University of Mysore.

Leinonen, L, Mujunen, R., Kangas, J., Torkkola, K. (1993). "Acoustic pattern recognition of fricative - vowel coartiuclation by the self-organizing map". Folia Phoniatrica, 45, 173-181.

Liberman, P. (1961). "Perturbations in vocal pitch". Journal of Acoustical Society of America, 33, 597-603.

Liberman, P. (1963). "Some measures of the fundamental periodicity of normal and pathological larynges". Journal of Acoustical Society of America, 35, 344-353.

Linville, S.E., Korabic, E.W. (1987). "Fundamental frequency stability characterstics of elderly women's voices", Journal of Acoustical Society of America, 81, 1196-1199.

Lippmann, R.P. (1987). "Pattern classification using Neural networks". IEEE Communications Magazine, April, 4-20.

Luchsinger and Arnold. (1965). Cited in Boone, D.R. (1971). "The voice and voice therapy".

Makhoul. (1985). Cited in Binu, R. (1998). "An objective classification of voice disorders using Artifical Neural Network". Unpublished Masters Dissertation, University of Mysore, Mysore.

Martin. (1970). Cited in Binu, R. (1998). "An objective classification of voice disorders using Artifical Neural Network". Unpublished Masters Dissertation, University of Mysore, Mysore.

Mavrian, H. (1960). Cited in Binu, R. (1998). "An objective classification of voice disorders using Artifical Neural Network". Unpublished Masters Dissertation, University of Mysore, Mysore.

Mc Cullock, E., Pitts, W. (1943). "A logical calculus of the ideas immanent in the nervous activity". Bulletin of Mathematical Biophysics, 5, 115-133.

McGlone, R.E., Hollien, H. (1963). "Vocal Pitch Characteristics of Aged Women", Journal of Speech and Hearing Research, 6, 164-170.

Metz, D. E., Schiavetti, N., Knight, S.D. (1992). "The use of Artificial Neural Networks to Estimate Speech Intelligibility from Acoustic Variables - A preliminary analysis". Journal of Communication Disorders. 25, 43-53.

Michel, J.F, Hollien, H., Moors, P. (1965). "Speaking fundamental frequency characteristics of 15, 16 and 17 year old girls". Lang, spech, 9, 46-51.

Michel, J.F., Wendahl, R. (1971). "Correlates of voice production" in L.E.Travis, (ed.), Hand Book of Speech Pathology and Audiology, Prentice Hall, Inc., Englewood Cliffs, N.J. 465-480.

Michel, J.F., Wendahl.R. (1971). "Correlates of voice production" in L.E. Travis (ed.), Handbook of speech pathology and Audiology, Prentice Hall, Inc., Englewood Cliffs, N.J.

Moore, G.R (1971). "Organic Voice Disorders", Prentice Hall. Inc., Englewood Cliffs, New Jersy.

Moore, P., Von Ledun, H. (1958). "Dynamic variations of the Vibratory pattern in the normal larynx. Folia Phoniatirica, 10, 205-238.

Mujunen, R., Leinonen, L, Kangas, J., Tookkola, K. (1993). "Acoustic Pattern Recognition of /s/ Misatriculation by elf-organizing map". Folia Phonatrica, 45, 135-144.

Muller, J. (1943). Cited in Nataraja .N.P. (1986). "Differential Diagnosis of Dysphonias". Unpublished Doctral thesis, University of Mysore, Mysore.

Murry, T. (1978). "Speaking fundamental frequency characterstics associated with voice pathologies". Journal of Speech and Hearing Disorders, 43(3), 374-379.

Murry, T., Doherty, E.T. (1980). "Selected Acoustic characterstics of pathologic and normal speakers", Journal of Speech and Hearing Research, 23(2), 361-369.

Nataraja, N.P. (1972). "Objective method of locating optimum pitch". Unpublished Master's Dissertation, University of Mysore.

Nataraja N.R, Jayaram, M. (1982). "A new approach to the classification of voice disorders". Journal of All India Institute of Speech and Hearing., 8, 21-28.

Nataraja, N.P. (1986). "Differential diagnosis of dysphonias". Doctoral thesis, University of Mysore.

Nataraja, N.R, Jagadish, A. (1984). "Vowel duration and fundamental frequency". Journal of All India Institute of Speech and Hearing., 15, 72-81.

Nataraja, N.R, Savitri, S.R. (1990)."Multi Dimensional Analysis of Voice Diorders". Unpublised Masters Dissertation, University of Mysore, Mysore.

Plomp, R. (1967). "Pitch of Complex tones". Journal of Acoustical Society of America, 41, 1526-1534.

Provrouest. (1972). Cited in Hollien, H., Shipp, T. (1972). "Speaking fundamental frequency and chronologic age in Males". Journal of Speech and Hearing Research, 15(1), 155-159.

Ramig, L. (1980). "Acoustic characterstics of voice and selected measures of body Physiology". Unpublished Doctrol thesis, Purdue University.

Raol, R.J., Mankame, S.S. (1996). "Artificial Neural Networks - A brief Introduction", Resonance, 47-54.

Rashmi, M. (1985). "Acoustic aspects of the speech of children". Unpublished Master's Dissertation, University of Mysore.

Robert, R., Baken, R.J. (1984). Multi Dimensional Analysis of Voice Diorders". Unpublised Masters Dissertation, University of Mysore, Mysore.

Robinson. (1992). Cited in Binu, R. (1998). "An objective classification of voice disorders using Artifical Neural Network". Unpublished Masters Dissertation, University of Mysore, Mysore.

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 Master's Dissertation, University of Mysore.

Sansone, RE., Emanuel, F.W. (1970). "Spectral noise level and roughness severity ratings for normal and stimulated rangh vowels prdouced by Adult males". Journal of Speech and Hearing Research, 13 (3), 489-502.

Sawashima, M. (1968). "Movements of larynx in articulation of Japanese consonants". Annual Bulletin -Research Ins., of logopedics and Phoniatrics, University of Tokyo, 2, 11-20.

Shanta, Y.S. (1973). "Establishing and validating isochronal tone stimultion technique", unpublished Masters Dessertation, Mysore

Shantha Y.S. (1973). "Establishing and Validiating Isochronal tone stimulation technique". Master's Dissertation, University of Mysore.

Sheela, E.V. (1974). "A comparitive study of vocal parameters of trained and untrained singers". Master's Dissertation, University of Mysore.

Shipp, T. and Huntington, D. (1965). "some acoustic and perceptual factors in acute-laryngitic hoarseness". Journal of Speech and Hearing Disorders, 14, 761-768.

Shridhara, R. (1986). "Glottal wave forms in normals". Unpublished Master's Dissertation, University of Mysore.

Skinner, E. (1935). "A calibrated recording and analysis of pitch, force and quality of vocal tones expressing happiness and sadness". Speech Monograph - II, 81-137.

Smith, S. (1954). "Remarks on the physiology of vibrations of the vocal cords". Folia Phoniatrica, 6, 81-137.

Sorensen, D., Hori, Y, Leonard, R. (1980). "Effects of laryngeal topical anesthesia on voice fundamental frequency perturbation". Journal of Speech and Hearing Research, 23, 274-284.

Sorenson, D. (1970)."Multi Dimensional Analysis of Voice Diorders". Unpublised Masters Dissertation, University of Mysore, Mysore.

Sorenson, D., Horii, Y. (1983). "Directional perturbation factors for Jitter and for Shimmer". Journal of Communication Disorders, 17, 143-151.

Stevens, K.N. (1967). Cited in Nataraja .N.R (1986). "Differential Diagnosis of Dysphonias". Unpublished Doctral thesis, University of Mysore, Mysore.

Stevens, K.N. (1977). Cited in Nataraja .N.R (1986). "Differential Diagnosis of Dysphonias". Unpublished Doctral thesis, University of Mysore, Mysore.

Stoicheff, M.L. (1981). "Speaking fundamental characteristics of non-smoking female adults", Journal of Speech and Hearing Research, 24(3), 437-441.

Suresh, T. (1991). "Acoustic Analysis of Voice in getiatries population". Unpublished Master's Dissertation, University of Mysore.

Tharmar, S. (1991). "Acoustic Analysis of voice in children and adults". Unpublished Master's Dissertation, University of Mysore.

Thondorf, W. (1955). Cited in Nataraja .N.R (1986). "Differential Diagnosis of Dysphonias". Unpublished Doctral thesis, University of Mysore, Mysore.

Titze, I.R. (1980), "Comments on the Myoclasitc Acrodynamic theory of Phonation". Journal of Speech and Hearing Research, 23, 495-510.

Titze, I.R. (1994). Cited in Stemple. J.C. Glaze., L.E., Gerdeman, B.K. (1994). "Clinical voice pathology - theory and management", 2nd edn, 39-41.

Usha, A.A. (1978). "A study of fundamental frequency in Indian population". Unpublished Master's Dissertation, University of Mysore.

Vandenberg. (1958). Cited in Nataraja .N.R (1986). "Differential Diagnosis of Dysphonias". Unpublished Doctral thesis, University of Mysore, Mysore.

Vanriper, C, Irwin, J.V. (1966)."Multi Dimensional Analysis of Voice Diorders". Unpublised Masters Dissertation, University of Mysore, Mysore.

Venkatesh. C.S. (1992). "Multi Dimensional Analysis of Voice Diorders". Unpublised Masters Dissertation, University of Mysore, Mysore.

Von Leden, H., Koike, Y. (1970). "Detection of Laryngeal Disease by Computer Technique". Archives of Otolaryngology, 91, 3-10.

Von Leden, H., Moore, P, Timicke, R. (1960). "Laryngeal vibrations. Measurements of the glottal wave. Part III. The pathologic larynx". Archives of otorhino laryngology, 71, 16-35.

Wendahl, R., Page, L. (1967). "Glottal wave periods in V.C.V environments". Journal of Acoustical Society of America, 42, 1208A-1210A.

Wendahl. R.W. (1971). 'Multi Dimensional Analysis of Voice Diorders". Unpublised Masters Dissertation, University of Mysore, Mysore.

West, R., Ansberry, M., Carr, A. (1957). "The rehabilitation of speech", (III ed.), Harper and Row, N.Y.

White head, R.L., Emanuel, F.W. (1974). "Some spectrographic and perceptual features of vocal fry, abnormally rough and modal register vowel phonations", Journal of Communication Disorders, 7, 305-319.

Wilcox, K. (1978). "Age and vowel differences in vocal jitter". Unpublished Master's Thesis, purdue University.

Wilcox, K., Horii, Y. (1980). "Age and changes in vocal Jitter". Journal of Gerentology, 35, 194-198.

Wilson, D.K. (1979). "Voice problems in Children", 2nd edn. The Williams and Wilkins Company, Baltimore.

Wyke, B. (1967). "Recent advances in the neurology of phonation; Phonatory reflex mechanisms in the larynx". The British Journal of Disorders of Communication, 2, 2-14.

Wyke, B. (1969). "Deus ex mechine vocis - An analysis of the laryngeal reflex mechanisms of speech". British Journal of Disorders of Communication, 4,(3-23).

Yanagihara, N. (1967). "Significance of harmonic change and noise componenets in hoarseness". Journal of speech and Hearing Research, 10, 531-541.

Yoon, M.K., Kakita, Y, Hirano, M. (1982). "Multi Dimensional Analysis of Voice Diorders". Unpublised Masters Dissertation, University of Mysore, Mysore.

Yoon, M.K., Kakita, Y, Hirano, M. (1984). "Sound Spectrographic Analysis of the Voice Patients with Glothic Carcinomas". Folia Phoniatrica, 36, 24-30.

Yumoto, E. (1982). "Harmonics-to-Noise ratio as an index of the degree of hoarseness". Journal of Acoustical Society of America, 71, 1544-1550.

Yumoto, E., Gould, W.J., Baer, T. (1982). "Multi Dimensional Analysis of Voice Diorders". Unpublised Masters Dissertation, University of Mysore, Mysore.

Yumoto, E., Sasaki, Y, Okamura, H. (1983). "The quantitative evaluation of hoarseness. A new harmonics to noise ratio method". Archives of otorhinolaryngology.

Zemlin, W.R. (1962). "Speech and Hearing Science" Prentice Hall Inc., Englewood Cliffs, N.J.

Zemlin, W.R. (1968). Speech and hearing science, Englewood Cliffs, Prentice-Hall, Inc., New Jersy.

Zurada, J.M. (1992). "Introduction to Artificial Neural Systems". PWS Publishing Company, Boston.

Appendix I

Comparison of mean values of the 29 parameters for normal males, dysphonic males, normal females and dysphonic females as shown by various studies and the results obtained in the present study.

	Anitha (1994)				Das (1996)	Gopal Krishna (1995)		Present study			
Speech stimuli	(for phonation of /a/ /i/ /u/ and sentence)					(for phonation of /a/ /i/ /u/		(for phonation of /a/ /i/ /u/ and sentence			
Parameters	NM	DM	NF	DF	DM	NM	NF	NM	DM	NF	DF
1. Fo	135	149.9	246	221	170	114	207	130	162	230	211
2. Fhi	135.25	187.7	296	254	174.5	120	217	153	233	260	304
3. Flo	123.00	116.5	224	186	137	108	193	117	123	206	167
4. STD	1.71	14.73	7.35	10.25	2.67	1.4	2.04	4.61	19.84	7.63	16.29
5. PFR	4.073	8.96	4.38	6.43	6.43	-	-	5.35	11.0	5.24	9.42
6. Fftr	3.832	5.21	3.86	4.05	6.23	4.73	1.34	2.58	4.77	2.36	5.02
7. Fatr	2.85	4.08	2.84	3.36	3.34	2.79	2.36	2.14	4.97	1.77	25.1
8. Jita	89.79	237	61.15	104.7	104.8	71.7	23.09	58.07	219.2	47.42	96.77
9. Jitt	1.39	3.53	1.49	2.17	2.17	0.82	0.47	1.19	9.22	1.59	2.43
10. RAP	0.78	2.05	0.86	1.27	1.26	0.47	0.27	0.43	3.35	0.68	1.49
11. PPQ	0.84	2.25	0.91	1.33	1.34	0.42	0.27	0.43	2.33	0.64	1.5
12. sPPQ	1.76	4.45	1.61	2.57	2.59	0.7	0.39	0.7	3.89	0.72	4.73

13. vFo	2.94	9.3	3.13	5.44	2.32	1.23	0.97	1.23	8.19	1.27	5.22
14. ShdB	0.44	0.87	2.46	1.19	0.95	7.43	0.17	0.19	0.41	0.17	0.47
15. Shim	4.49	8.85	3.89	5.55	6.95	3.04	1.93	2.16	4.24	1.61	3.39
16. APQ	4.83	7.4	3.73	4.78	6.17	2.43	1.3	1.76	3.39	1.1	2.7
17. sAPQ	10.70	14.59	9.77	11.02	12.8	4.01	2.07	3.47	4.29	1.87	4.63
18. vAm	15.94	22.75	16.46	21.72	21.95	9.43	8.29	8.24	9.87	6.45	8.47
19. NHR	1.75	14.15	1.88	6.98	6.79	0.13	0.11	0.14	1.7	0.21	0.44
20. VTI	0.083	0.11	0.06	0.061	0.41	5.3	3.74	0.029	0.11	0.022	0.53
21. SPI	15.29	20.76	15.66	22.72	25.79	16.17	41.37	37.43	17.31	51.72	15.61
22. FTRI	1.23	2.5	1.32	1.6	3.03	0.41	0.44	0.87	2.66	0.31	1.57
23. ATRI	7.3	9.22	7.09	9.21	2.7	3.77	1.59	2.69	2.96	1.66	3.39
24. DVB	1.62	3.11	2.35	3.11	4.26	0	0	12.05	17.8	10.09	12.7
25. DSH	0.077	2.81	0.275	3.00	4.3	0	0	0.15	3.24	1.03	2.56
26. Duv	16.2	24.13	16.65	15.5	15.35	0	0	15.03	23.47	10.57	15.86
27. NVB	0.033	0.64	0.43	0.45	0.892	0	0	3.96	4.35	4.33	3.88
28. NSH	0.058	1.56	0.17	2.45	2.92	5.03	0	0.13	2.15	1.18	1.97
29. Nuv	13.76	17.58	10.98	11.31	12.38	0	0	0.72	3.04	0.28	6.88

NM - Normal males, DM - Dysphonic males, NF - Normal Females, DF - Dysphonic males

Appendix - II

MDVP option acquires, analyzes and displays 29 voice parameters from a single vocalization. This program uses the computerized speech lab hardware system (CSL) for signal acquisition, analysis and play back. The 29 parameters are available as a numerical file or can be displayed graphically in comparison, to a data base. This data base of extracted voicing parameters is used for comparison and the user can add his own finding to this database. The advantage of MDVP extraction is that different parameters are important for the analysis of different vocal pathologies. The MDVP extracts 29 parameters in about 16 seconds. The definition of the 29 parameters as given in the MDVP manual are as follows :

1. Amplitude perturbation quotient (APQ)/%/

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

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

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

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. The smoothing reduces the sensitivity of APQ 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. Breathy and hoarse voice usually have an increased APQ. APQ should be regarded as the preferred measurement for shimmer in the MDVP.

2. Amplitude Termor Intensity Index (ATRI)/%/

Average ratio of the amplitude of the most intense low-frequency amplitude modulating component (amplitude tremor) to the total amplitude of the analyzed 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.

3. Degree of subharmonic components (DSH)/%/

Estimated relative evaluation of sub-harmonic to Fo components in the voice sample.

4. Degree of Voice Less (DUV)/%/

Estimated relative evaluation of nonharmonic areas (where Fo cannot be detected) in the voice samples.

5. Degree of Voice Breaks (DVB)/%/

Ratio of the total length of the areas representing voice breaks to the time of the complete voice sample. DVB is computed as the ratio of the sum of all voice break length to the length of the complete voice same T sam (Voice command) as :

$$DVB = \frac{t_1 + t_2 + t_3 \ldots + t_n}{T_{sam}}$$

where $t_{1,}$ t_2 ....t_n - lengths of the 1st, 2nd ... voice break.
T_{sam} - length of analyzed voice data samples.

DVB does not reflect the pauses before the 1st and after the last voiced areas of the recording. It measure the ability of the voice to sustain uninterrupted voicing. The normative threshold is "O" because a normal voice, during the task of sustaining voice, should not have any voice break areas. In cases of phonation with pauses (such as running speech, voice, breaks, delayed start or earlier and of sustained phonation) DVB evaluates only the pauses between the voiced areas.

6. Amplitude tremor frequency (FATR)/Hz/

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 threeshold, the Fatr value is zero.

7. Fx - Tremor frequency (FFTR)/Hz/

The number of semi-tones frequencies found in the range F/o to Fhi consists of PFR.

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

8. Highest Fundamental Frequency (FHI)/Hz/

The greatest of all extracted period to period fundamental frequency values. Voice break areas are excluded. Fhi is the highest fundamental frequency from the extracted period to period pitch data (Voice command). It is computed as

$$Fhi = \max\ [Fo^{(i)}] \quad i = 1,2,......N.$$

where $Fo^{(i)} = \dfrac{1}{To^{(i)}}$ - period to period fundamental frequency values

To (i), i = 1,2,......N - extracted pitch period data.

N - Number of extracted pitch periods

9. Lowest Fundamental Frequency (FLO)/Hz/

The lowest of all extracted period to period fundamental frequency values. Voice break areas are excluded. F/o is the lowest fundamental frequency from the extracted period - to - period pitch data (Voice command).

It is computed as

$FIO = \min Fo^{(i)}, \; i = 1,2 \ldots\ldots N,$

Where $Fo^{(i)} = \dfrac{1}{To^{(i)}}$ period-to-period fundamental frequency values

$To\,(i) = i = 1,2, \ldots\ldots N$- extracted pitch period data.N- Number of extracted pitch period.

10. 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 :

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

Where $Fo\,(i) = \dfrac{1}{TO^{(i)}}$ = Period to period fundamental frequency.

$To\,(i), \; i = 1,2, \ldots\ldots N$ - extracted pitch period data, N = PER - number of extracted pitch periods.

11. 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 analyzed voice signal.

12. Absolute Jitter (JITA)/usec/

An evaluation of the period to period variability of the pitch period within the analyzed voice sample. Voice break areas are excluded.

Jita is computed from the extracted period-to-period pitch data (voice command) as :

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

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

Absolute Jitter measures the very of the pitch 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. MDVP also provides the jitter the research literature contains normative data for three four parameter. The MDVP customer is generally advised to use RAP or PPQ instead of Jita and Jitt for determining jitter present in the voice.

Both Jita and Jitt represent evaluations of the same type of pitch pertubation. Jita 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 of Jita for men and women differ significantly. Higher pitch results into lower Jita. That's why, the Jita values of two subjects with different pitch are difficult to compare. Jitt is a relative measure and the influence of the average fundamental frequency of the subject is significantly reduced.

13. Jitter Percent (JITT)/%/

Relative evaluation of the period -to-period (every short term) variability of the pitch within the analyzed voice sample voice break areas are excluded. Jiit is computed from the extracted period to period pitch data (voice commana) as :

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

where $To^{(i)}$, $i = 1,2 \ldots\ldots\ldots N$-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.

14. Noise to Harmonic Ration (NHR)

Average ratio of the inharmonic spectral energy in the frequency range 1500-4500 Hz to the harmonic spectral energy in the frequency range 70-4500 Hz. This is general evaluation of noise present in the analyzed signal.

15. Number of subharmonic segments (NSH)

Number of autocorrelation segments where the pitch was found to be a sub-harmonic of Fo.

16. Number of unvoiced segments (NUV)

Number of unvoiced segments detected during the autocorrelation analysis.

17. 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).

18. Phonotory fundamental frequency range in semitones (PFR)

It is the range between he highest Fo (HFi) and lowest F_o (FLo) expressed in the number of semin tones. PFR is computed by accounting the number of semi-tones in the frequency range FLO to Fhi. The ratio of the two consequentive semi tones (ki) is equal to 12th root of 2.

First are the frequency of semitones $Fst^{(k)} = f_1 a^k$, $K = 1, 2, 3 \ldots\ldots. n$ are computed within the frequency range 55 Hz - 1055 Hz where a = 12... V2.

$f_1 = 55$ Hz, $f_2 = 1055$ Hz,

$f_1 \leq fst(K) \leq f_2$

19. Pitch period perturbation quotient (PPQ)/%/

Relative evaluation of the period-to-period variability of the pitch within the analyzed voice sample with a smoothing factor of 5 periods. Voice break areas are excluded. PPQ is computed as,

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

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

PPQ measures the short term (cycle-to-cycle with a smoothing factors of 5 periods) irregulatrity 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 perturbation of the voice very well. Hoarse and/or breathy voices may have an increased PPQ.

20. Relative average perturbation (RAP)/%/

Relative evaluation of the period-to-period variability of the pitch within the analyzed voice sample with smoothing factor of 3 periods. Voice breaks areas are excluded. It is computed as :

$$RAP = \frac{\frac{1}{N-2}\sum_{i=2}^{N-1}\left|\frac{To^{(i-1)} + To^{(i)} + To^{(i+1)}}{3} - To^{(i)}\right|}{\frac{1}{N}\sum_{i=1}^{N} To^{(i)}}$$

where $To^{(i)} = i = 1, 2,....N$ - 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 3 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-to-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. MDVP also provides the jitter parameters PPQ, Jitta and Jita because the research literature contains normative data for these parameters the MDVP customer is advised to use RAP or PPQ instead of Jita and Jitt as an indication of jitter in the voice.

21. Smoothed amplitude perturbation quotient (SAPQ)/%/

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.

$$\mathrm{SAPQ} = \frac{\dfrac{1}{N-sf+1} \displaystyle\sum_{i=1}^{N-sf+1} \dfrac{1}{sf} \sum_{r=0}^{sf-1} A^{(i+r)} - A^{(i+m)}}{\dfrac{1}{N} \displaystyle\sum_{i=1}^{N} A^{(i)}}$$

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

Sf = smoothing factor defined by voice. Smooth SAPQ allows the experimenter to define his own amplitude perturbation measure by changing the smoothing factor from 1 to 199 periods. This flexibility is desirable because in the scientific literature researchers use amplitude perturbation measures with different smoothing factors or without smoothing.

With a small smoothing factors, SAPQ is sensitive mostly to the short-term amplitude variation of the voice impulses.

With a smoothing factor of 1 (no smoothing), SAPQ is identitical to Jitter variations occurring between consecutive pitch periods. Usually this type of variation is random. It is typical for hoarse voices. However,pitch extraction errors may object Jitter percent significantly.

22. Shimmer in dB (shdb)/db/:

Evaluation is dB of the period-to-period (very short term) variability of the peak-to-peak amplitude within the analyzed voice sample voice break areas are excluded.

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

where $A^{(i)}$, i = 1,2……….N - extracted peak to peak amplitude data.
N - number of extracted impulses.

Shimmer in dB measure the very short term (cycle-to-cycle) irregulatiry to peak-peak amplitude of the voice. This measure is widely used in the research literature on voice perturbation(Iwata & vonLeden 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 can vary for a number of reasons. 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. MDVP also provides the shimmer parameters APQ and Shim. APQ because the research literature contains normative data for these three parameters 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 and dB.

23. Shimmer Percent (SHIM)/%/

Relative evaluation of the period-to-period (very short term) variation of the peak-to-peak amplitude within the analyzed voice sample voice break means are excluded. Shim is computed from the extracted peak-to-peak amplitude data (voice command) as :

$$\text{shim} = \frac{\dfrac{1}{N-1} \sum_{i=1}^{N-1} \left| A^{(i)} - A^{(i+1)} \right|}{\dfrac{1}{N} \sum_{i=1}^{N} A^{(i)}}$$

where $A^{(i)}$, $i = 1,2\ldots\ldots N$ - 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.

24. Soft Phonation Index (SPI)

Average ration of the lower frequency harmonic energy in the range 70-1600 Hz to the higher frequency harmonic energy in the range 1600-4500 Hz.

25. Smoothed Pitch Period Perturbation Quotient (sPPQ)/%/

Relative evaluation of the short or long-term variability of the pitch period within the analyzed voice sampl9e at smoothing factor defined by the user. The factory setup for the smoothing factor is 55 periods. Voice break areas are excluded.

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

$$sPPQ = \frac{\frac{1}{N-sf+1} \sum_{i=1}^{N-sf+1} \frac{1}{sf} \sum_{r=0}^{sf-1} To^{(i+r)} - To^{(i+m)}}{\frac{1}{N} \sum_{i=1}^{N} To^{(i)}}$$

where : $To^{(i)}$, $i = 1,2\ldots N$ - extracted pitch period data, N=PER-number of extracted pitch periods. sf-smoothing factor defined by VOICE.SMOOTH.SPPQ.

26. Standard deviation of fundamental frequency (STD)/Hz/

Standard deviation of all extracted using the voice command period to period fundamental frequency values. Voice break areas are excluded.

STD is computed as standard deviation of the extracted period-to-period fundamental frequency data (Voice command) as

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

$$\text{where } Fo = \frac{1}{N} \sum_{i=1}^{N} \quad Fo^{(i)} = \frac{1}{To^{(i)}} \text{ - period to perio}$$

fundamental frequency values.

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

N = Number of extracted pitch period data.

27. 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 areas of the signal where the influence of the frequency and amplitude variations, voice breaks and sub-harmonic components are minimal. VTI measures the relative energy levels of a high frequency noise .

28. Co-efficient of Fo variation (VFO)/%/

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

$$vFO = \frac{\sigma}{Fo} = \frac{\frac{1}{N} \sum_{i=1}^{N} \left(\frac{1}{N} - \sum_{j=1}^{N} Fo^{(i)} - Fo^{(i)} \right)^{2}}{\frac{1}{N} \sum_{i=1}^{N} Fo^{(i)}}$$

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

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

To (1), i = 1,2. . . . . .N, extracted pitch period data.

N = 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 increase the value of VFO. Because the sustained phonation normative thresholds assume that the Fundamental frequency should not change, any variations in the fundamental frequency are reflected in vFO. These changes could be frequency tremors (i.e., periodic modulation of the voice) or non periodic changes, very high jitter or simply rising or falling pitch over the analysis length.

29. Co-efficient of amplitude variation (vAM)/%/

Relative standard deviation of peak-to-peak amplitude. It reflects in general to peak-to-peak amplitude variations (short term to long term) within the analyzed voice sample. Voice break areas are excluded.

vAM is computed as ration of the standard deviation to the average value of the extracted peak-to-peak amplitude data (voice command) as :

$$\mathrm{VAM} = \frac{\frac{1}{N} \sum_{i=1}^{N} \left(\frac{1}{N} \sum_{i=1}^{N} A^{(i)} - A^{(i)} \right)^{2}}{\frac{1}{N} \sum_{i=1}^{N} A^{(i)}}$$

where $A^{(i)}$, i-1,2.........N - extracted peak to peak amplitude data.
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 increase the value of vAM.