ACOUSTIC CORRELATES
OF
HOARSE VOICE

David Thangapulvar

REG. NO. M9602

A DISSERTATION SUBMITTED AS PART FULFILMENT OF FINAL YEAR M.Sc. (SPEECH AND HEARING) TO THE UNIVERSITY OF MYSORE

ALL INDIA INSTITUTE OF SPEECH AND HEARING
MYSORE - 570 006

MAY 1998

DEDICATED

To

My **Lord Jesus Christ** *who is the way, the truth and the life.*

My **loving parents** *who guided me to the Eternal life.*

My **loving wife** *who encourages me to walk in the Eternal life.*

CERTIFICATE

This is to certify that the dissertation entitled, **"ACOUSTIC** CORRELATES OF HOARSE VOICE", *is the bonafide work in part fulfillment for the degree of Master of Science (Speech of Hearing), of the student with* ***Reg. No. M9602.***

Mysore
May 1998

Dr. S.NIKAM
Director
All India Institute of Speech and Hearing
Manasagangothri
Mysore - 570 006
INDIA

CERTIFICATE

This is to certify that this disseration entitled , "ACOUSTIC CORRELATES OF HOARSE VOICE", *has been prepared under my supervision and guidance.*

Mysore
May 1998

Dr. N.P. Nataraja,
Professor and Head
Department of Speech Sciences
All India Institute of Speech and Hearing
Manasagangothri
Mysore 570 006.
INDIA

DECLARATION

This dissertation entitled , **"ACOUSTIC CORRELATES OF HOARSE VOICE",** *is the result of my own study under the guidance of* **Dr. N.P. Nataraja,** *Professor and Head, Department of Speech Sciences, All India Institute of Speech of Hearing, Mysore, and has not been submitted earlier at any University for any other Diploma or Degree.*

Mysore
May, 1998

Reg. No. 9602

ACKNOWLEDGEMENT

First and foremost I extend my heartfelt gratitude to my lecturer and guide **Dr. N.P. Nataraja,** Professor and Head, Department of Speech Sciences, All India Institute of Speech and Hearing, Mysore for this valuable guidance, support and encouragement to complete this work.

My thanks to **Dr. S**. **Nikam,** Director, All India Institute of Speech and Hearing for her kind permission to carry out this study.

I am grateful to **Dr. S.R. Savithri M'am, Mrs. Sri Devi M'am, M/s Lalitha M'am** and other staff in the Speech Science Department for their timely help & cooperation.

My thanks to Library staff for their cooperation in my library work related to this study.

I am much obliged to all the subjects who were patient enough to lend their voice for this study.

My sincere thanks to all my classmates as well as First M.Sc. Students, particularly Muthu Krishnan.

I am grateful to my loving wife, **Mrs. Lavanya David** for her constant help & encouragement in completing this work.

Simply saying thanks is meaningless to **the Lord Almighty** who has provided me with good health, strength and knowledge.

CONTENTS

LIST OF TABLES

PAGE NO.

LIST OF GRAPHS

INTRODUCTION

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

The underlying basis of speech is voice. According to Green (1964) "voice plays the musical accompaniment to speech rendering it tuneful, pleasing, audible and coherent and is an essential feature of efficient communication by the spoken word". The usefulness of voice in human life is immeasurable. Damage to voice by means of either misuse or abuse of voice or any pathology in the laryngeal system can paralyse social interaction to a great extend, resulting in considerable psychological, social and economic imbalance. At this juncture, the role of speech and language therapist is drawn into picture. Some of the main functions of speech and language therapist are early indentification and prevention of voice disorders, planning appropriate intervention program and monitoring the prognosis during the course of treatment.

Traditionally, the clinicians use visual inspection of larynx and subjective perceptual evaluation of voice quality to diagnose the laryngeal pathology (Yanagihara, 1967). Subjective perceptual evaluation have had some degree of success in separating normal and pathological voice. However, it has its own limitation on test-retest and inter-rater reliability (Yanagihara 1967; Koike 1969).

Voice is considered as multidimenssional series of measurable events. Development of technology has permitted the analysis and measurement of various aspects of vocal function. There have been many attempts over the years to find different voice parameters and objective methods that aid in early detection, diagnosis and treatment of dysphonics

The various objective approaches are high speed cinematography, stroboscopy, electroglottography, sound spectrography, photoglottography, echoglottography and inverse fittering. Even though, these techniques have been promising, there have been problem with instrumentation, methodology and analysis.

Another approach, far and wide, clinically used is acoustic analysis which includes time domain analysis and frequency domain analysis. The frequently studied jitter and shimmer measures are time domain analysis. The frequency domain analysis is also known as spectral analysis which gives quantitative and objective information on voice.

Numerous studies have been reported on quantification of hoarse voice, some have studied jitter measures alone to quantify [Liberman 1963; Hecker and Kreul (1971); Michel and Doherty (1973)], while others have considered shimmer measures (Koike 1969; Kitajima and Gould 1976; Emanuel 1978). Some others have experimented with spectral measures of hoarse voice (Nessel 1960; Yanagihara 1967; Emanuel, Lively & MeCoy 1973).

Both jitter and shimmer measures were used by Wendahl (1966). Balaji (1988) had considered Long term average spectrum and electroglottography in dysphonics and Pathak (1995) studied the combination of spectral and perturbation measures. On the quantification of hoarseness. There is hardly any Indian study that combines several measures in a single experiment on hoarse voice.

Hence, the present study encompasses the combination of frequency and intensity measures for the purpose of:

1) differentiating horse voice from normal voice
2) classifying the hoarse voice
3) comparing perceptual estimation with acoustic estimation of hoarse voice.

The parameters considered for the present study are:

A) SPECTRAL MEASURES:

1. Harmonics to Noise Ratio (HNR)
2. First Harmonic Amplitude (H1A)
3. Number of Harmonics (NOH)
4. Alpha Ratio (AR)
5. Beta Ratio (BR)
6. Gama Ratio (GR)

B) PERTURBATION MEASURES:

I) Jitter Measures

7. Mean Fundamental Fequency (JFO)
8. Percent Jitter (PJ)
9. Period Variability Index (JPVI)
10. Relative Average Perturbation (JRAP - 3 point)
11. Directional Perturbation Quotient (JDPQ)
12. Deviation from Linear Trend (JDLT)

ii) Shimmer Measures:

13. dB Shimmer (SdB)
14. Amplitude Variability Index (SAVI)
15. Amplitude Perturbation Quotient (SAPQ)
16. Directional Perturbation Quotient (SDPQ)

C) FREQUENCY MEASURES:

17. Mean Fundamental Frequency (FFO)
18. Range of Frequency (FRAN)
19. Extend of Fluctuation in Frequency (EFF)
20. Speed of fluctuation in Intensity (SFI)

D) INTENSITY MEASURES:

21. Mean Intensity (IMAO)
22. Range of Intensity (IRAN)
23. Extend of Fluctuation in Intensity (EFI)
24. Speed of Fluctation in Intensity (SFI)

HYPOTHESIS:

1. There is no significant difference between normals and dysphonics in terms of parameters studied.

2. There is no significant difference between males and females in terms of parameters studied.

3. There is no significant difference among the four groups based on degree of hoarse voice in terms of parameters studied.

Limitations:

1. The study has been limited to 50 normal (each sex 25) and 30 dysphonic (each sex 15) subjects.

2. The age range of subjects was limited to 18 to 50 years.

3. Most of the instrumental analysis (except LTAS) were carried out on phonations.

REVIEW OF LITERATURE

Human communiation involves as a rich tapestry of information conveyed through elements of movements, emotional expression and vocalization. Human communication system can be classified into two broad divisions i.e. verbal and nonverbal systems. Spoken language is one form of verbal communication system that enables human to convey information with specificty and detail. Language whether spoken, written or signed involes a system of symbols tht conveys meaning. Language involves the interaction of many skills which combine for effective communication. According to Curtis (1978) "Every human society, no matter how primitive, has developed the ability to communicate through speech and our ability to communicate through spoken and written langauge has been cited as the single most important characteristic that 3ets human apart from other animals".

Speech is the way of life for man and it is the chief medium of social adaptation and control. According to Boone (1971), "the act of speaking is a very specialized way of using the vocal mechanism. The act of singing is even more so. Speaking and singing demand a combination or interaction of the mechanisms of respiration, phonation, resonation and speech articulation". Thus voice forms the basis of speech.

Voice, articulation and language are the major elements of human speech production. When a disorder is present related to any of these elements, the ability to communicate may be impaired. Voice is in elements of sp. that provides the speaker with the vibratory signal upon which speech is carried. Regarded as magical or mystical at ancient times, today, this production of voice is viewed as a powerful common and activity tool. It serves as the melody of speech and provides expression of feelings, intent and mood to thoughts.

"Voice plays the musical accompaniment to speech rendering it tuneful, pleasing, audible and coherent and is an essential feature of efficient communication by the spoken word". (Greene 1964). The main function of voice is for normal daily communication. It is also used for other professional purposes by individuals such as singers, actors, Radio/TV artists, lawyers, teachers, sales persons and others. These professionals are in need to use their voice efficiently. The inefficient or abuse of vocal system leads to organic changes in the system. This causes loss of voice or abnomral voice. Voice problem may severely disturb communiation with others, resulting in considerable economic, social, and psychological disturbances. The demoralising effect on communication is greater in the case of professional users of voice. In addition to this, the human voice serves as sublinguistic purpose of survival such as ventilating emotions such as anger, grief and affection whcih are essential to the maintenance of psychologic equilibrium. According to Perkins (1971) thee are atleast five kinds of nonlinguistic functions of voice. Voice can reveal speaker's identity, health, emotional state, personality and aesthetic orientation. Voice is also a carrier of cannotative communicative content.

Voice can reveal sex, age, inteligence, regional and socio-economic origin, education and occupation. The physiological factors of genetic endowment of physical structures, the health of the individual may affect the voice. The health of an individual may be indicated by qualities of voice that portray pain, resiratory diseases or by those that show fitness and well being. Voice gives psychological clues to a person's self image, perception of others and emotional health. Self image such as confidance, shyness, and aggressiveness can be identified by voice quality. Conclusively, it can be infered that voice is more than a means of communicating verbal massage, it serves as a powerful conveyer of personal identity, emotional state, education and social status.

A voice disorder exists when a person's voice quality, pitch and loudness differ from those of similar age, sex, background and geographical back ground. (Aronson 1980; Boone 1977, Greene 1972, Moore 1971). In other words, when the acoustic and aerodynamic properties of voice are so deviant that they draw attention to the speaker's voice, then disorder of voice is considered to be present.

Damage to voice by means of either misuse or abuse of voice or any pathology in laryngeal system can paralyse social interaction to a great extend, resulting in considerable psychological, social and economic imbalance. Therefore the voice problem must be treated immediately after it is identified.

The voice disorders are classified in terms of etiologic (cause), perceptual (acoustic) and Kinesiologic (vocal hypo function and vocal hyper function).

Voice disorders are grouped according to acoustic perceptual attributes as quality, pitch, loudness and flexibility.

Voice quality is the perception of physical complexity of laryngeal tone modified by cavity resonation. Fairbanks (1960) tried to distill voice quality defects into three cateogries-harseness, breathiness and hoarseness. Individual variation more often exists in perceptual judgement of voice quality.

Vocal pitchs is the perceptual correlate of fundamental voice frequency. Disorders of pitch refer to abnormally high and low pitch voices.

Loudness is the perception of vocal intensity. The voice may be too weak or too loud.

Flexibility is the perceptual correlate of frequency, intensity and complexity variations. The normal voice possessess adequate pitch, loudness and quality variability during spontaneous speech to convey more intellectual and emotional meaning. In voice disorder, these flectuations are either inappropriately flattened or excessive.

Rarely in clinical practice, abnormal voice vary along a single dimension of quality, loudness, pitch or flexibility. Most of the time, eventhough one may predominate, the others are usually present in different combinations and proportions.

DESCRIPTION OF HOARSENESS:

Defining hoarseness is a difficult task, because hoarseness is a psycho acoustic term used in broader sense to mean any abnormal voice quality due to laryngeal pathology. The term hoarseness is being understood differently by different groups. To a lay-man it implies a sudden change in voice quality or an unpleanent voice. Several have defined hoarseness of voice as follows:

According to Baynes (1966) hoarseness is a quality of voice that is rough, grating, harsh, more or less disocrdant and lower in pitch than normal for the individual.

Moore, Silverman and zimmer (1971) define hoarseness is characterised by noise of a relatively high frequency that is produced by transient or highly unstable variations.

Casper M. etal (1981) considers Hoarseness as a deviation in the tonal quality of the voice resulting when the vocal cords vibrate in an aperiodic or haphazard manner.

Van Riper and Irwin (1978) describes hoarseness in terms of breathiness and harsiness.

Seth and Gruthrie (1935) stated that the hoarseness is tonal quality produced when the vocal folds vibrate in an aperiodic, irregular or hapazard manner.

In spite of several meanings assigned to hoarseness, the common factor invariably noticed is that hoarseness is a phonatory phenomenon rather than a resonatory phemomenon. i.e. it is produced by the laryngeal sound generator. Therefore the assumption is that hoarseness is the result of some sort of abnormal vibration of vocal cords.

PATHOLOGIES ASSOCIATED WITH HOARSENESS:

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

CONGENITAL	ACQUIRED
I. Web of the larynx	I. Traumatic
II. Cysts	II. Inflammatory
a) Cystic hydroma	a) Acute - Non specific
b) Dermoid cyst	- Specific
c) Branchial cyst	b) chronic - Non specific
	- Specific
III. Tumors	III. Neoplastic
a) Lipoma	a) Benign
b) Fibroma	b) Malignant
c) Leiomyoma	IV. Paralytic
d) Chondroma	V. Miscellaneous.
e) Haemangioma	

Sederholm et al (1992) showed with the help of factor analysis that hyperfunction, breathiness and roughness are good predictors of hoarseness. Hoarseness and breathiness are two components of hoarseness. Harseness is perceived due to irregularity of vocal fold vibrations (Coleman, 1960; Wendahl 1963; 1966; Moore 1975) i.e. Variations or perturbations in both amplitude and time period from cycle to cycle give the impression of harseness. Breathiness is perceived by escape of air through partially closed glottis and the resultant turbulance noise reduces the harmonic to noise ratio (HNR). Excessive aperiodieity also generates noise and reduces the prominence of the harmonics, hence reducing the Harmonics to noise ratio. Thus, hoarseness is defined as a voice quality which clearly contains noise components and that can be labelled harsh and breathy (i.e. source noise elements plus friction noise); its perceived pitch tends to vary substantially; common description of this quality are 'noisy', 'harsh', 'wet' (Anders et al, 1956).

PATHOPHYSIOLOGY OF HOARSENESS:

The primary or common factor in hoarseness is noise of a relatively high frequency that is produced by transient vibrations. These sounds are combined with other phonatory sounds that are frequently at low pitch as the result of laryngeal disease or any other condition that would lower the frequency of vocal fold vibration. The trani3ent disturances seem to occur on the surface of the vocal folds, particularly along the glottis, but other laryngeal structures may also contribute to the total effect.

Sources of laryngeal transients can be grouped into four categories (1) Accumulation of sticky mucus secretion in the larynx. Excessive mucus tends to interfere with normal movements of vocal folds by weighting them unevenly and damping their excursion through causing them to adhere to each other. (2) Relative flacidity of one or both vocal folds. The flacidity causes independent vibration, resulting transient disturbances. (3) Additions to the mass of the folds. Mass causes pitch change, hoarseness by weighting.

stiffening and influencing vocal fold's compliance. (4) The destruction of all or part of the vocal folds cause random vibrations and transients resulting hoarseness (Moore 1971).

Objective measurements (Ultra-high speed photography, synchronstroboscopy, photoglottography, EGG) reveal that these are mainly three phases in a laryngeal wave form: a) Opening phase, b) Closing phase and c) Closed phase, with a definite temporal relationship. The relationship among these functions change as vocal output varies. For example, the phases of the cycle vary with different loudness levels (frequency being constant). The closed phase becomes shorter for louder sounds as compared to softer sounds. Other patterns can be produced by other intensity-frequency combinations. However, the following two conditions are present for any normal phonation, even though variation among the patterns may occur. (Moore and Thompson, 1965). (a) All three phases of vibratory cycle can be seen, (b) The motion of the two cords tends to be relatively synchronous and equal in amplitude.

During sustained normal phonation, the length and amplitude of adjacent cycles are generally similar. However, careful observation of a phonatory sequence may demonstrate small changes in the contour showing frequency and amplitude of the cycles. (i.e.) these parameters are rarely precisely the same among cycles within a series. This has been supported by Scripture (1906), Simon (1927) and Lieberman

(1961). Thus, the normal laryngeal vibration provides a basis for analyzing vocal fold motion in abnormal hoarse voice. In this respect, it would be theoretically possible for the vocal folds to move in a number of atypical ways in individual cycles or sequences of vibrations. The vocal folds could move within a single cycle in atleast five ways: (a) absence of glottic closure (b) different amplitudes of movement in each cord (d) lack of movement by one cord and (e) dissimilar movement patterns along the extent of one or both folds. (Moore and Thompson, 1965).

In addition to the above mentioned abnormalities, laryngeal vibrations possibily exist with sequences of vibratory cycles. These would include random and patterned changes in the amplitude or a period for successive glottal openings. These changes could occur simultaneously in both folds or independently in either fold. Thus, the potential complexity of vibratory patterns resulting from cyclic abnormalities and sequential irregularities are almost endless. Accordingly, if hoarseness can be assumed to result from abnormal vocal fold vibration, its origin should be found in one or more of the suggested vibratory patterns. Based on the above mentioned assumption, perceptual and acoustic studies either in isolation or together have been conducted to find out the correlation between these two measures of hoarseness.

QUANTIFICATION OF HOARSENESS:

A) PERCEPTUAL ANALYSIS OF HOARSENESS:

Human ears have the ability to identify and recognize speaker's voice. Well trained voice clinician are often able to determine the causative pathologies on the basis of psychoacoustic impression of voice (Hirano, 1975).

Voice quality is a term that subsumes a wide range of possible meanings, conveying both supra laryngeal and laryngeal aspects (Kratt and Klatt 1990). As to the concept of normal verses pathological voice qualities, some parameters such as diplophonia or aphonia must be regarded as pathological. For most voice quality parameters, however, there is no distinct border between normal and pathological. For instance, some studies have shown that breathy voice seems to be a common female voice (Henton and Bladon 1985; Bless, Biever and Campros 1989; Sodersten and Lindestad 1990), whereas creaky or vocal fry is a normal male voice characteristic (Henton and Bladon 1988).

Voice quality is also to some extent culturally conditioned and will likely be influenced by aspects specific to a certain language community. Thus vocal fry/creaky is more common in certain regional accents such as modified Northern English in the area of Leeds, Yorkshire (Henton and Bladon 1988). However significant correlation between frequency perturbation and perceptual qualities such as instability, roughness, flutter, diplophonia and creakiness/vocal fry were found. This is in agreement with the findings of Deal and Emanual (1978) and Askenfelt and Hammarberg (1986). Hammarberg and Gauffin (1986) concluded that perceptual evaluation by well-trained listeners is reliable and reproducible and cna be used for systematic evaluation purposes, if handled with precaution. These authors further concluded that voice quality can be more precisely perceived, if professional terminology is given to the listener.

The reliability of perceptual evaluation can be improved by (1) Operationally defining the voice parameter to be evaluated. (2) Illustrating the voice quality parameters by samples of tape recordings. (3) Serching for acuostic and physiological correlates of perceptual parameters.

i) Importance of Perceptual Evaluation :

Although voice properties can be examined at the physiological, acoustical and perceptual level, the judgement of voice quality is primarily a perceptual matter. Gauffin and Hammerbeg (1995) quoted that eventhough perceptual voice ratings are subjective and impressionistic, there are arguments for such ratings: i.e., (i) perceptual aspects of voice are important, as they play a crucial role in the listeners acceptance of the voice (ii) skilled voice clinicians are able to perceive and distinguish between different voice qualities (iii) perceptual training as a means of patient's self-control of voice function in voice therapy to make the patient improve vocal behavior iv) as pointed out by Kreiman et al (1993), listeners judgements are usually regarded against which other (quantitative) measures are evaluated.

ii) Factors affecting perceptual judgements:

a) Listener groups : Yumato, Sasaki, Okamura (1984) found correlations ranging from 0.51 to 0.79 when 8 laryngologists rated the hoarseness of 87 voices on a 4 point equally appearing interval scale.

b) Rating scales : Most of the perceptual studies use rating scales to simplify the procedures. But it has limitations like scale being too small leading to lack of adequate information.

c) Language: Perceptual analysis is always language specific, semantic contents, idiomatic expressions can vary with different speakers.

d) Age, Sex, social cultural factors: influence interjudge agreement according to sonninen and Sonninen (1976).

e) Voice Sets: There is abundance of evidence that listeners differ systematically in their judgements for different voice sets.

iii) Recent Methods in Perceptual Evaluation:

Toner, Emanuel and Parker (1990) compared two techniques- Direct magnitude estimation and equal appearing interval along with spectral noise level (SNL) measurement and concluded that a high degree of linearity exists between both the techniques and Spectral Noise Leve. Multidimensional scaling analysis has been used by Kreiman, Gerratt and Precoda (1990) and Kreiman, Gerratt; Precoda and Berke (1992). Askenfelt and Hammerberg (1986) found PERC - 7 technique and that also a multidimensional scale. Weavers and Lowe (1990) used a visual analogue scale in which different relevant parameters were represented by a 100 cm. continuous line, the extremes of which corresponding to non-existance and extremely high occurrence of the trait, respectively. Here the judges were supposed to mark a point on the continuum which they thought was is a representative measure of the parameter in the voice under consideration.

iv) Reliability of Perceptual Evaluation:

Kreiman et al (1993) found that reliability of the listeners varied greatly from study to study. Also, their own experiment indicated that the ratings varied widely across individual clinicians. They suggested that the presentation of anchor stimuli or taped reference examples of deviant voice qualities might improve, between listener rating consistency.

v) Scales of Voice quality:

It is obiovious that the variety of voice quality dimensions in a voice sample is to be rated should be reflected in the rating parameters and scales. The most prevalent scale has been the Equal Appearing Interval (EAI) scale (Kreiman et al 1993) which requires the listener to assign a number 1 and n (most of 5 or 7) to a voice sample regarding degree of a certain voice quality. The advantage of EAI scale is easier communication because of the of numbers. Kreiman et al (1993) study indicated, however, that EAI rating driffed in a consistent direction within a listening session.

B) ACOUSTIC ANALYSIS OF HOARSENESS:

The measures used in acoustic analysis of voice are convenient, non invasive, objective, sensitive and quantitative method of studying laryngeal mechanism while producing speech.

Studies have been conducted to identify measurable voice features that are correlated with hoarseness and thus effectively predict the degree of hoarseness perceived by listeners. Some of the advantages of these methods are that quantitative data from the correlated measurement could be easily stored or transmitted to those who need to see them. The measurement are repeatable from audio recorded voice samples and in due course might be obtained by standardised procedures in clinical or research contexts.

The purpose of research in acoustics is aimed at

-> Speaker Identification.

-> Delineating the mechanophysiologic limitations of normal and Pathological laryngeal performance (Scripture 1906).

-> Detecting and descriminating the types of vocal pathology.

-> Monitoring and tracking response to Theorapy

-> Searching for acoustic correlates of voice quality (Moore & Thampson 1965) and checks their variations with voice production conditions caused by various patholigies (Ludlow' 1981).

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

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

METHODOLOGICAL CONSIDERATIONS IN THE ACOUSTIC ANALYSIS OF HOARSENESS:

The estimation of hoarseness and other vocal qualities by acoustic parameters varies with different methodological issues. Some of these are summarised in this section of the review.

a) Modes of Data Acquisition:

The traditional methods used are, recorded speech on a tape with the help of high fidelity microphone and audio cassette recorder/disk. Storing in the computer after digitization is also done. This signal is affected by the transfer effects of the vocal tract Reverberaltion and ambient noise. So other techniques, invasive and non invasive, have been used. Invasive techniques are uncomfertable and can't be used with children. The major non invasive techniques are as follows.

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

* Fourcin (1981) made simultaneous recordings of Electro-glottograph and airflow velocity curves for different modes of phonations and described the method to interpret

laryngeal waveforms. Electroglottograph reflects the vibratory cycle of the VF's with fairly high fidility.

According to Dejonckre & Lebacq (1995) "Electro glottograph reflects the glottal conditions more during the closed phase.... As majority of laryngeal pathologies manifest abnormalities more during the closed phase, Electro glottograph has been considered as a better technique for studying VF movement in dysphanics.

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

* Cepstral analysis was first described by Noll (1964). It relies on the fourier analysis of the speech signal. The speech signal is filtered by a low pass filter and then digitized in order to perform Fast Fourier Transform (FFT). Koike (1986) applied cepstral analysis to study short term perturbation. He compared the findings of the cepstral analysis of residual signal after inverse filtering and an acoustic speech and reported that the residual signal gave a simpler form than that of acoustic speech.

LONG TERM AVERAGE SPECTRUM (LTAS):

Recent research has shown that LTAS is a reasonable index of vocal quality (Carr and Trill 1964).

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

B) Speech sample:

Most of the studies have employed sustained vowels rather than running speech (Horii 1979). According to Horii (1979) there was a paucity of data on large quantities of speech because of the lack of efficient instrumentation and measurement procedures. There is a strong argument for sustained vowels because it gives only the random perturbations associated with physiological limitations of the glottal sound, source and controls supraglottal sources of variations. This allows measurement of only short term perturbations and checks long term systematic perturbation due to phonetic context, stress and intonation. It is to be stressed that most investigators, whatever their choice of speech material, considered that the results achieved by their respective analysis systems confined the feasibility of separating normals and dysphonic subjects at a reasonable level of performance. However looking into these factors, the use of the mid-portion of sustained vowel produced at a natural comfortable pitch and intensity level appears to be the most appropriate phonatory task when changes in perturbation caused by automatic physiologic conditions. of the larynx are in question (Koike 1969; Iwata & Van leden 1970; Iwata 1972; Hosii 1979, 1982).

c) Manual Vs. Automatic analysis:

This includes handmarking of analog oscillograms, semiautomatic methods using interactive digital wave form editors and both handware and software automatic pitch tracers. Some of the earliest studies have involved the use of hand measurements. (Leiberman 1961). This method is extremely tedious and time consuming because of their minute nature. More recent studies (Horii 1972, 82; Wilcox 1980) have used computerised instrumentation which is fast and precise. In between these two are the semi automatic instruments. A lot of subjective judgement is required in both manual and semiautomatic and hence automatic extraction gives much precision.

Many have applied computer techniques which uses a formula or algrithm for the analysis of the acquired. waveform. The majority of acoustic perturbation studies as well as the spectral noise studies have been limited to analysis by means of a single formula (Lieberman 1961,63) & others two formulas (Horii 1980). Regardless of the alogrithm each investigation found that their measure provided some degree of discrimination between normals and pathologic subjects.

Qi. Y. & Shipp-T (1992) devised a new method for tracking irregularities in the acoustic waveform of a sustained phonation using the adaptive Wiener filter. Irregularities were determined by the techniques of correlation cancellation. The alogritm was evaluated using sustained vowals produced by a formant synthesiser and by subjects with and without phonatory disorders. Results indicate that the method is capable of differentiating between normal and abnormal voices.

d) Temporal resolution:

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

Cox, Ito and Morrison (19S9a) reported that increasing the sampling frequency from 10 KHz to 20 KHz had little effect on DFT based Harmonics to noise ratio estimates with all differences being 0.6dB in perturbed data. However, the same in perturbation free data brought HNR from 21.9 dB to 41.2 dB for /i/ vowel and from 29.4dB to 49.0dB for /a/ vowel suggesting that over sampling brings a significant improvement in perturbation free data.

e) Amplitude Resolution:

This is commonly known as bit resolution which gives the resolution of a system along the ordinate where the amplitude of the acoustic wave is represented. This is usually expressed in terms of number of bits which can easily be converted into relatively simpler unit of amplitude resolution: i.e., the number of samples per unit amplitude. Lower bit resolution produces the bit noise contaminating the original analog signal. A minimum of nine bits of resolution are needed to minimise the contaminating bit noise with out intrapolation (Titze etal. 1987).

Interpolation:

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

g) Waveform marking:

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

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

h) Sample duration:

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

* Size of window -> Titze etal (1987) suggested a window of 20-30 cycles with-in a given token of steady vowel phonation.

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

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

i) Vowels:

Perturbation measures have been shown to be different among different vowels by Horii (1979). Mormative data from

Wilcox and Horii (1980) have shown that /u/ was associated with significantly smaller jitter (0.55%) than /a/ & /i/ for which the values were 0.68% & 0.6 9% respectively.

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

j) Fundamental frequency:

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

k) Sex:

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

1) Age:

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

ACOUSTIC PARAMETERS USED TO QUANTIFY HOARSENESS:

Researching acoustical analysis has come out with a number of acoustic parameters to quantify hoarse and breathy voice quality. As perturbation measures and spectral noise parameters are taken up for the study, these will be dealt in detail in this part of the review.

i) SPECTRAL NOISE MEASURES

A frequency domain analysis of the signal called as spectral measures to give frequency specific information to separate the harmonic components from the interhommic noise components. Harmonic components resulting from quasi-periodic interruptions of airflow by the vocal folds bring purity to the signal whereas interharmonic noise components resulting from interrupted turbulent transglottal airflow adds to the noisy perception.

The methods used for the calculation of spectral noise level are developed based on certain assumptions regarding the source of spectral noise and hence each one is sensitive to a particular component of the entire magnitude of noise and hence each one has certain merits and demerits. The common problem with these methods is their limited ability to resolve individual glottal cycles for analysis of fractional error in fundamental period extraction or in synchronisation causes additional spectrum leakage of total harmonics, causing further deterioration of the harmonic structure. As a result of all these technical problems, spectral noise measures are yet to demonstrate clinical usefulness (Muta etal 1988), though they have the potential to become the strongest measure for quantification of hoarseness.

Carchart (1941), Nessel (1960), Ishiki etal (1966) Yanagihara (1967) were among the first to observe that an increase in perceived hoarseness was associated with elevated acoustic spectral noise components.

Yanagihara (1967) reported close correlation between the degree of perceived hoarseness of voice and the amount of noise observed in standard spectrograms. He classified hoarseness into four grades based on the noise relative to that of the harmonic component in the spectrogram. He found that acoustic parameters of hoarseness are mainly determined by the interactions of the following three factors.

> Noise components in the main formant of each vowel.

> High frequency components above 3 KHz.

> Loss of high frequency harmonic components.

Some of the parameters in spectral noise measures are:

1) Spectral noise level
2) Signal to Noise Ratio (SNR)
3) Harmonic to Noise Ratio (HNR)
4) Normalised Noise Ratio (NN energy)
5) Breathiness Index (Br index)
6) First Harmoic Amplitude (H_1 amplitude)

7) Spectral tilt
8) Alpha, Beta, Gama ratios of LTAS.
9) High frequency power ratio (HFPR)
10) RA values.

1. Spectral Noise level:

Emanuel & Sansone (1969) defined it as the lowest peak marking of the vowel spectrum. In a series of articles the authors reported that a strong linear relationship between the Spectral noise level and the perceived degree of hoarseness. They suggested that the Spectral noise level measurement gave a more reliable acoustic index of vowel wave aperiodicity and hoarseness than harmonic level measurment. The disadvantage is that only the lowest peak stylus marking for each section of the spectrum was taken into account and the other points like the level of harmonic components of the spectrum was disregarded. Another demerit is that, it is not feasible for most patients with laryngeal disorders to phonate at an intensity of 75dB SPL for 7 sees. Which is reguired to allow comparison with different measures of Spectral noise level. These measures need visual inspection and hence are subjective. Imaizumi, Hiki, Hirano, Matsushita (1980) have found evidences in support of the above findings.

2. Signal to Noise Ratio (SNR):

Fourier expansion to separate the noise from the periodic components was used by Kojima, Gould, Lambiase and

Isshiki (1980) to compute the signal to noise ratio as an objective estimate of hoarseness.

The resolution of voice into signal and noise components may not be satisfactory, since only three pitch periods used in the fourier transform ie one third of the Fourier components was counted as the signal (Hiraoka etal 1984). This method has theoritical limitations also with regard to the accuracy of estimated noise levels, since the fourier co-efficients derived from a signal with duration T provides estimates of the noise components only at multiple frequencies of 1/T whereas the noise has a continous frequency spectrum. This method is too complex and time consuming to apply to clinical use.

3. Harmonics to Noise Ratio (HNR):

Harmonics to Noise ratio was proposed by Yumoto, Gould & Baer (1982). It was defined by them as the ratio of acoustic energy of the stable harmonics to that of noise. This measure takes into account, the Jitter and Shimmer present in the signal, which is one of its advantages, because jitter affects the spectrum of a sustained vowel by reducing the amplitudes of harmonics and introducing noise between them. Titze etal (1987) reported that Harmonics to Noise Ratio includes waveform perturbation along with peak amplitude and period perturbation. The first step in the calculation of HNR using signal averaging technique of Yumoto is to average the individual pitch pulses. Here, the size of the averaging window is determined by the largest pitch pulse in the signal. For periods shorter than this maximum, the interval between the end of the pitch pulse and the end of the averaging window is filled with zeros. If a sufficient number of periods are averaged, a large proportion of the noise is cancelled. The R.M.S. energy of the average pitch pulse is used as the numerator in the Harmonics to Noise Ratio calculation. The amount of periodic energy is estimated by successive subtractions of the average pitch pul3e from individual periods of the original vowel. The R.M.S. energy in the noise signal is used as the denomination in HNR calculation on a decibel scale HNR is defined as

$$20 \log \frac{\text{R.M.S. (Average)}}{\text{R.M.S. (Noise)}}$$

Yumoto etal (1982) reported HNR values ranging from 7.0 to 17 dB for a group of normals and from -15.2 to 9.6 dB for a group of speakers with a variety of laryngeal disorders. So as the degree of hoarseness increases, Harmonics to Noise Ratio decreases. They also found a highly significant agreement (P=0.849) between HNR calculations and the subjective evaluation of the spectrograms. The HNR proved useful in quantitatively assessing the results of treatment of hoarseness. Subsequent researchers found this index to be superior to the other well established indices of hoarseness. Wolfe, Steinfatt (1987) & Wolfe and Ratusnik (1988) reported that the correlation of severity of hoarseness with spectral noise determined in terms of HNR was higher than that with the jitter and Shimmer values measured.

DEMERITS OF HNR

This algorithm is based on the assumption that a long stationary interval can be obtained from the sustained vowel production, but it can't always be expected in actual recording situations because the speech signal generally has the tendency to change smoothly in amplitude and pitch over a long interval of the sustained phonation. Thus, this method may detect the smooth changes in the waveform incorrectly as noise components. In addition it should be noted that HNR can't quantify noise in the severally hoarse voice that has no recognizable periodic components. Other demerit is that this is highly sansisitive to errors in pitch period demarcation, and a dependency on jitter perturbation, Fo and vowel type was also demonstrated. Hiraoka etal (1984) suggested that a voice spectrum should be resolved into three points - Fo component, the harmonic component and the noise & that the relative increase of Fo component in hoarse voice spectra is important. So relative harmonic intensity (Hr) was proposed to evaluate hoarse voice. Relative Harmonic intensity is defined as the intensity of the second and higher harmonics expressed percent of the total voice intensity.

$$Hr. = \frac{Pi}{P} \times 100$$

Where pi = intensity of the i th component

P = total voice spectral intensity.

They reported Relative Harmonic intensity values of 67-72% larger than critical values, for normals and lesser Relative Harmonic intensity for Hoarseness thus providing a good discrimination between and hause voice.

The HNR given by Yumoto et al (1982) is considered best suited for the quantification of spectral noise due to the reason that HNR is sensitive to both jitter and additive noise. However it used the vocal output near the lip as a signal for calculation, which can't be considered as the glottal signal because the transfer function of the supra glottal structures modify the glottal source before it is picked up by the mic near the lips.

For this reason, cepstral analysis was employed by Anantha Padmanabha (1992) to nullify the effects of Transfer function of a supraglottal cavity. He found that the HNR calculation based on the cepstral analysis of the lip radiated vowels are even sensitive to detect the severe and profound categories of hoarseness, whereas the techniques of Yumoto fails to do so. He developed a software program called Harmonics to Noise Ratio based on cepstral analysis where he U3ed a haming window of only 4 pitch period at a time at an interval of every 10 ms for cepstral analysis and a final cepstrum was obtained by finding a cepstral average (LTCA). Two different measures, peak Harmonics to Noise Ratio and

Average Harmonics to Noise Ratio, can be obtained. This technique eliminates another disadvantage of traditional Harmonics to Noise Ratio calculation which was found to be sensitive to the smooth changes in amplitude and pitch in addition to actual shimmer and jitter.

A new method of computation of HNR called the Dynamic Time Warping (DTW) was proposed by Qi. y (1992) to avoid the demerits of the earlier methods. In this method, noise components of voice were calculated from the discrepancies between wavelets after they had been optimally aligned in time. The optimal time normalisations of wavelets was accomplished by DTW. This method was evaluated using both synthetic and natural voices and significant reductions in noise were obtained. HNR measure obtained by this technique was free of frequency perturbations.

Another capstrum based technique was used by Kekrom (1993) to calculate spectral HNR descreases in speech signals. He found that HNR almost linearly with both increasing noise levels and increasing jitter continuaum. He concluded that the method could be considered as a valid technique for determining the amount of spectral noise and as a useful measure in the analysis of voice quality.

Pathak (1995) also agreed with the above findings. He reported that Peak HNR measures together obtained from the samples of vowels /a/ and /u/ had the potential to be included in the diagnostic battery for the classification of various degrees of hoarseness and also as a screening measure.

4. First Harmonic Amplitude:

Fairbanks (1940) found enhanced First Harmonic Amplitude in breathy vocal quality, and it was attributed to the more nearly sinusoidal shape of breathy glottal waveforms. Still other investigators (Biekley 1982; Fischer & Jorgenson 1967) also reported similar findings. Huffman (1987) used inverse filtering to derive glottal waveforms from samples of four phonation types used in breathy samples showed stronger first harmonics than non-breathy samples. Hillenbrand, Cleveland and Erickson (1994) studied this measure as a possible correlate of breathy voice. They found that the relative amplitude of the first harmonic correlated moderately with breathiness ratings. Though there are a number of studies to report First Harmonic Amplitude as a quantification of breathy vocal quality, there are very few studies to report this measure as a correlate of hoarse voice and the validity of this measure, hence needs further investigations.

5) Spectral tilt:

Several investigators noted that breathy signals tend to have more high frequency energy than normally phonated signals (Hillenbrand etal. 1994). Klich (1982) reported strong correlations between perceived breathiness and several measures of spectral tilt calculated as energy ratios of low.

medium and high frequency bands. Fukazawa etal (1988) reported that the spectral tilt measure accounted for approximately half of the variance in breathiness ratings obtained from the sustained vowels of speakers with various vocal pathologies. But there have been contradictory factors on the reliability of spectral tilt as an indication of breaty vocal quality.

6. Normalised Noise Energy:

Kasuya etal (1986) proposed normalised noise energy (NNE) which was considered to be superior to other measures of spectral noise. The Normalised noise energy, was automatically computed from the voice signals using an adaptive comb filtering method performed in the frequency domain. Experiments with the voice samples have show that Normalised Noise Energy is especially effective for detecting the glotic cancers. Since the Normalised Noise Energy measures primarily the turbulence noise caused by the closing insufficiency of the glottis during the phonation. It is very useful in the detection of these diseases. But Normalised Noise Energy is not sensitive to the noise caused by irregular vibratory motion of the vocal folds. Hence Normalised Noise Energy is not an effective measure for those laryngeal conditions which produce hoarseness because of a periodicity of vocal fold movements.

7) Breathiness Index (Br Index):

Proposed by Fukazawa. El Assuooty, and Honjo (1987) the Breathiness Index was an indicator of the turbulent noise in breathy voice. The parameter was determined by the ratio between energy of the second derivative of the high pass filtered wave and that of the non-derived high pass filtered wave. The principle was to utilize the difference in the frequency range between the turbulent noise and other components present. The parameter was found to correlate with the perception of breathiness. Hence this could be applied for the screening purposes to detect the pathologies which generated turbulent noise and could not be used with pathologies that generate excessive perturbations.

8) High frequency power ratio (HFPR):

It was proposed by Shoji, Rogenbogen, Yu and Blaugrund (1992) for quantification of breathy voice. It is defined by the investigators as the ratio of high frequency power versus total power, calculated as the lower limit of the high frequency range (F_c) and varied from 1 to 10 KHz. They reported that the HFPR values measured at an Fc of 6KHz. significantly separated normal from breathy voice.

9) R.A. values:

Gobi and Ni chasaide (1992) study revealed that Breathy Voice quality had high R.A. values showing considerable dynamic leakage. However unlike whispery voice where RA values are high throughout the few glottal pulses were rather closer to modal values. EE values (EE-corresponds to over intensity of a signal so that an increase in EE amplifies all frequency components equally) are also lower than for model but not as low as whispery voice for most of the interval in question. The open quotient and RK parameters serve to differentiate between breathy and whispery voice qualities. Breathy Voice has a higher RK and open quotient which would indicate a more symmetrical glottal flow, pulse with a relatively longer closing branch.

[Where RA = (TA/To) is a measure of the return phase which is the residual flow from the point of excitation to complete (or maximum) closure. It affects the steepness of the source spectrum. A large RA corresponds to greater attenuation of the high frequencies.

RK - is a measure of symmetry/asymmetry of the glottal pulse : a larger value means a more symmetrical pulse.

OQ - open quotient is the ratio between the open phase and fundamental period].

10) ALPHA, BETA AND GAMMA RATIOS OF LTAS:

There are a number of methods by which speech can be analysed spectrally. One such analysis procedure takes a time average of the sound pressure level per cycle across frequency. This measurement is commonly referred to as the

Long Term Average Spectrum of speech (LTAS) (Formby and Monsen, 1982).

The measurement of the long term average speech spectrum is obtained by passing the speech energy through a series of contingenous band pass filters and integrating the energy at the output of each filter. These average values are then plotted to arrive at the visual representation, a smoothened plot by the envelope of the power spectrum of the speech sample (Formby and Monsen, 1982).

LTAS has been used for studies of the human voice source. The speech signal represents the product of the sound source and the vocal tract transfer functions. The vocal tract transfer function differs for different sound segments, but in the averaging process, the short term variations due to phonetic structure will be averaged out and the resulting spectrum can be used to obtain information on the sound source (Lofqvist and Mander3on, 1987). Frokjaer - Jensen and Prtytz (1976) used the ratio of energy below and above 1 Khz and named it as Alpha parameter. According to them, since the amplitude above 1000 Hz is normalized relative to the amplitude below 1000 Hz, is independent of the microphone distance, amplitude level, etc. Nataraja (1986) studied three spectral parameters in the voice of dysphonics, they are:

1) The ratio of intensities between 0-1 KHz and above 1-5 KHz and named AA.

2) Ratio of intensities of harmonics and noise in 2-3 KHz and named it as AC.

3) Frequency of first formant (AD) which is defined as the frequency with maximum intensity in the range of 300-1000 Hz.

Nataraja (1986) found the following results:

i) The males and females of the dysphonic group show no significant difference in terms of AA ratio. The males and females of the normal groups also showed no significant difference. A statistically significant difference was found between the dysphonic and the normal groups. The dysphonic groups showed lower AA values than the normal groups i.e., the dyshonics had higher intensities in the frequencies above 1 KHz than normals.

ii) The AC ratio values shown by the males and females of the dysphonic groups were found to be not significant. The normals also showed similar results. However, a significant difference between the males of the two groups was found.

iii) The first formant frequency of the males and females within the dysphonic group showed a significant difference. This was similar to the results seen within the normal groups. The dysphonic group and the normal groups of males and females did not differ from each other. Thus he concluded that out of three parameters, only AA parameter significantly differentiated normal and dysphonics in both genders.

Lofqvist and Manderson (1987) made two measurements on the calculated long term spectrum. They were:

1. The ratio of energy between 0-1 KHz to 1-5 KHz.

According to them, "This ratio provides a measure of the overall tilt of the sound spectrum. A high value of this ratio indicates that the fundamental and the lower harmonics dominate the spectrum which thus falls off rapidly. A low value of this ratio shows, on the other hand, that the sound spectrum has a lower spectral tilt.

2. Measurement of the energy between 5-8 KHz.

A high level of energy at these frequencies can be associated with noise components of the source in a hypofunctional voice (Yanagihara, 1967). Rashmi (1985) has made an attempt to study the ratio of intensities below and above 1 KHz, in the spectra of vowel /i/. She has concluded that:

i) The energy level above 1 KHz is less than the energy level below 1 KHz.

ii) The alpha parameter shows no significant different till the age of 9 years in both males and females. The female group in the age range 9 to 14 and the male group age ranging from 9 to 15 years had shown some changes, and

iii) No significant difference between males and females has been found. The age group above 9 years of age showed a change in the voice quality both in the case of males and females as reflected by the changes in ratio. The mean value ranged from 0.78 to 0.92.

Gopal (1986) reports no significant difference between males and females upto the age of 55 years. A significant difference was observed between males and females in the age range 56 to 65 years group i.e. males showing a higher score (0.73) than females (0.70). The value ranged from 0.71 to 0.76 in the age range 16 to 55 years both in the case of males and females. Similar to the results of Rashmi's (1985) study, the average intensity above lKHz has been less than below 1 KHz.

Wendler, Doherty and Hollien (1980) met with encouraging results when they used long term speech spectra to objectively differentiate between four classes of voices according to auditive judgements (normal, mild, moderate or severe degree) of hoarseness. In addition, they attempted to differentiate between certain degrees of roughness and breathiness as well as to carry out differential diagnosis based on acoustic analysis.

Hartman and Cramon (1984) studied the amount of spectral energy in the 1 to 5 KHz range and above 5 KHz in two sub groups of patients. The first subgroup showed a voice quality compound of breathiness and tense, which gradually normalized in the follow-up period. The other subgroup initially exhibited a normal or lax, breathy voice, which subsequently became more tense. They found that the variations of 3pectral energy and the duration of aspiration preceding voice on-set indicates signs of tense and breathy voice production. Thus they concluded that this measure is sufficient to differentiate these two subgroups.

II) PERTURBATION MEASURES

a) Jitter measures:

Michael & Wendahl (1971) defined jitter as the cycle to cycle variation in pitch period that occured when an individual is attempting to sustain phonation at a constant frequency. This gives direct information on the status of the phonatory system.

Jitter is a measurement of variation of a given period differs from the period that immediately follows it. Jitter is highly sensitive to pathological changes in the phonatory process. Normals present some amount of jitter but pathological voices have higher magnitude of jitter.

Various measures of jitter differ among themselves in either one or more of these factors like basic assumption regarding the source of perturbation, the rational behind the techniques, statistical treament of data, the degree of automation in computing, magnitude verses connected speech. Each of the method has it3 own advantages and drawbacks.

The measures used were:

(1) Mean jitter (2) Percent jitter (3) Jitter Ratio (JR) (4) Directional perturbation factor (DPF) (5) Jitter factor (6) Relative average perturbation (RAP) or frequency perterbation quotient (FPQ) or pitch perturbation quotient (PPQ) (7) Period variability Index (PVI) (8) Deviation from linear trend (DLT).

i) Mean Jitter:

"Mean jitter" is the average absolute difference in fundamental period between adjacent pitch pulses. This is measured in milliseconds.

$$M.J. = \frac{1}{n-1}\left[\sum_{i=1}^{n-1}\left|P_i - P_{i+1}\right|\right]$$

Moore and Thompson (1965) found mean jitter values of 0.30 msec for severe hoarse voice and 0.06 msec for a moderate hoarse voice. Horii (1985) reported mean jitter value for adult males ranging from 14-40 years to be 0.0176, 0.102 and 0.078 msec for /a/, /i/ and /u/ respectively. Kane and Wellen (1985) reported the mean jitter varying from 0.0023 to 0.0472 msec with a mean of 0.0123 msec for children of age ranging from 6 to 11 years with vocal nodule. Sridhara (1986) studied the mean jitter in 30 young normals using /a/, /i/ and /u/ vowels and reported the following data:

TABLE 1 : Data for "Mean Jitter" (milliseconds) obtained by Sridhara (1986).

Sex	/a/	/i/	/u/
Male	0.065	0.110	0.067
Female	0.058	0.030	0.048

But, this is an absolute measure and tends to be proportional to the mean fundamental period (Lieberman, 1963; Hollien et al, 1973; Horii, 1979).

ii) Percent Jitter:

Percent jitter is defined as mean jitter in milliseconds divided by the mean period in milliseconds, multiplied by 100

$$\text{Percent Jitter} = \frac{\frac{1}{n-1}\left[\sum_{i=1}^{n-1} \left| P_i - P_{i+1} \right|\right]}{\frac{1}{n}\sum_{i=1}^{n} P_i} \times 100$$

Moore and Thompson (1965) found that percent jitter was 4.9% and 1.4% for severally and moderately hoarse voices respectively. Jacob (1968) found a median jitter of about 0.6% for phonation produced at a comfortable pitch and intensity level. Hollien et al (1973) found 0.5% and 1.1% jitter for 102 Hz and 276 Hz sustained vowel phonations. Results of jitter analysis of normal sustained phonation by Young adults indicated that jitter of the order of 0.5 to 1.0% was typical (Hollien, Girard and Coleman, 1977; Horii, 1979). Smith, Weinberg, Feth and Horii (1978) established a range from 5.4 to 14.5% of jitter for esophageal voice. Lastly, Nataraja and Savithri (1990) reported that a jitter greater than 3% is considered abnormal.

iii) Jitter ratio:

The value of percent jitter is very small in case of normal in sustained phonation. So Jacob (1968) used another index termed "jitter ratio", which can be obtained by multiplying the percent jitter by 10.

Jitter Ratio (JR) - % Jitter x 10

In other words, JR is defined as the mean absolute jitter (in milli seconds) divided by the mean pitch period (in milliseconds) multiplied by 1000.

$$\text{Jitter Ratio} = \frac{\frac{1}{n-1}\left[\sum_{i=1}^{n-1} \left| Pi - P_{i+1} \right|\right]}{\frac{1}{n}\sum_{i=1}^{n} Pi} \times 1000$$

Horii (1979) reported a range of 5.3 to 7.6 of jitter ratio for six normal males age ranging from 28 to 43 years.

iv) Directional Perturbation Factor:

"Directional Perturbation Factor" as the percentage of time period difference between adjacent period which differed in algebraic sign. Hecker and Kreul (1971) found that the directional perturbation factor could separate 5 patients with cancer of the vocal folds from 5 normal speakers whereas Lieberman's perturbation factor did not separate the two groups of speakers. Higgins and Saxman (1989) compared the intra subject variation across sessions of three measures of jitter and reported that directional perturbation factor (DPF) was more temporally stable measure as compared to jitter factor and pitch perturbation quotient.

Sorensen and Horii (1984) reported the more completely available data obtained for 20 men and 20 women and this may be considered as tentative norm.

TABLE 2 : Normative data for DPF (%) obtained by Sorensen and Horii (1984).

Sex	/a/	/i/	/u/
Male	46.24	46.37	49.26
Female	48.79	52.04	52.77

v) Jitter Factor:

The jitter factor is defined by Hollien et al (1973) as the mean difference betweent he frequencies of adjacent cycles divided by the mean frequency multiplied by 100.

$$\text{Jitter Factor} = \frac{\dfrac{1}{n-1}\left[\displaystyle\sum_{i=1}^{n-1} \left| F_i - F_{i+1} \right|\right]}{\dfrac{1}{n}\displaystyle\sum_{i=1}^{n} F_i} \times 100$$

Hollien et al (1973) established the values of jitter factor for normal adult males of age ranging from 21 to 37 years and reported the mean jitter factor to be 0.47, 0.53, 0.43 and 0.97 for 102 Hz, 142 Hz, 198 Hz, and 276 Hz respectively of sustained phonation. Later on, Murry and Doherty (1980) reported a higher mean jitter factor of 0.99 for the elderly male group of age ranging from 55 to 71 years whose mean fundamental frequency was 115.3 Hz.

vi) Relative Average pitch perturbation:

This is also called the frequency perturbation quotient (FPQ) or pitch perturbation quotient (PPQ).

The RAP is defined as the ratio of a moving average of fundamental period difference to the average fundamental period where the length of the moving average is equal to either 3 or 5 periods.

$$\text{Rap (3)} = \frac{\dfrac{1}{n=1} \sum_{i=2}^{n=1} \left| \dfrac{P_i - 1 + P_i + 1}{3} - P_i \right|}{\dfrac{1}{n} \sum_{i=1}^{n} P_i}$$

Koike (1973) found that the mean frequency RAP for 30 adult speakers of both sexes and various ages was about 0.0046 for the midsection of the sustained /a/ vowel. Later on, Takahashi and Koike (1976) reported a mean RAP value of 0.0057 for 7 males and 0.0061 for 2 females. Koike (1973), Takahashi and Koike (1976), Koike et al (1977) and Davis (1979, 1981) applied this technique and demonstrated the feasibility of screening laryngeal pathology like tumors, unilateral paralysis etc.

vii) Period Variability index (PVI):

Deal and Emanuel (1978) defined PVI as the mean of the squares of the deviations of each period in the sample from the mean period divided by the square of the mean period, multiplied by 1000.

$$PVI = \frac{\frac{1}{n}\left[\sum_{i=1}^{n} (Pi - \bar{P})^2\right]}{(\bar{P})^2} \times 1000$$

They reported mean PVI of 0.4412, 0.4898 and 0.4451 for /a/, /i/, /u/ vowels respectively sustained for seven seconds at 75 dB SPL for 20 normal adult males.

viii) Deviation from Linear Trend (DLT):

Ludlow, Coulter and Gentges (1983) proposed a different index of jitter called "deviation from linear trend" (DLT). Here the pitch periods of two cycles away from the cycle in question in both the directions are averaged and the difference of this period from the average is calculated for all cycles except the four cycles at both extremities (2 each) of the vowel sample. Finally, the average of all the diff erences in data is found which is termed the mean deviation from linear trend (DLT).

DLT might not detect perturbation caused by a short cycle regularly alternating with a long one, as it occurs in the pulse register. (Cavallo, Baken and Shaiman, 1984). Ludlow (1981) proposed a diplophonia ratio which is sensitive to such alternating changes. Thi3 can be obtained after division of mean DLT by mean jitter in milliseconds. The same ratio can be converted into percentage by multiplying it by 100. Ludlow et al (1983) established an overall Deviation from Linear Trend vaule of 28.43 for 17 normals of both sexes

B) Shimmer Measures:

Michel & Wendahl (1971) defined Shimmer as the "cycle to cycle variation in amplitude that occurs when the individual attempts to sustain phonation at a constant frequency and intensity. This refers to glottal function, and also serves to quantify short term instability of the vocal signal.

Usefulness of Shimmer information in the description of voice characteristics has been clearly indicated. Wendahl (1966) claims that Shimmer is as important as jitter in its contribution to the perception of hoarseness. Researchers have found shimmer to be more important than jitter in terms of sensitiveness to laryngeal pathology. It can also be concluded that it can be used a diagnostic tool.

The measures are:

1) Mean Shimmer

2) Percent shimmer

3) dB shimeer

4) Directional perturbation factor for amplitude (Amplitude DPF)

5) APQ

6) AVI

i) Mean Shimmer:

It is the averaging absolute differences in the peak amplitudes between adjacent pitch cycles. This can be expressed in terms of millivolts or millimeters.

$$\text{Mean shimmer} = \frac{1}{n-1}\left[\sum_{i=1}^{n-1} |Ai-Ai+1|\right]$$

ii) Percent Shimmer:

This measure tends to be proportional to the absolute amplitude. Hence a correction is required to make this measure free from absolute amplitude, which makes it necessary to divide this measure by mean peak amplitude of these cycles. This ratio is analogous to jitter ratio or jitter factor. This ratio can be converted into percentage by multiplying it by 100 which is called percent shimmer.

$$\text{Percent Shimmer} = \frac{\frac{1}{n-1}\left[\sum_{i=1}^{n-1} \left|Ai - Ai+1\right|\right]}{\frac{1}{n} \sum_{i=1}^{n} Ai} \times 100$$

Nataraja and Savithri (1990) reported that the percent shimmer of 3% can be considered normal and above 3% is abnormal.

$$\text{dB shimmer} = \frac{\sum_{i=1}^{n-1} \left| \log \frac{Ai}{Ai+1} \right|}{n-1} \times 20$$

iii) dB shimmer:

Shimmer in dB is defined as the mean of logarithms of the ratio of the peak amplitudes of the successive peaks, multiplied by 20.

Kitajimma and Gould (1976) studies normal males and females and reported that average shimmer in normal phonation is on the order of 0.1 dB for the vowel /a/, with a critical value of 0.19 dB. The data ranged from 0.04 dB for normals. They also studied 25 subjects of vocal polyps who produced a result ranging from 0.08 dB to 3.23 dB for /a/ vowel.

Horii (1980) found an overall average shimmer of 0.39 dB with a critical value of 0.98 dB for the sustained vowel phonations of /a/, /i/ and /u/ for 31 normal males of age range 18-38 years. The individual shimmer values for these vowels were 0.47 dB, 0.37 dB and 0.33 dB respectively. Later, in another study with 12 adult males age ranging from 24-30 years, Horii (1982) found shimmer value as 0.62 dB, 0.48 dB and 0.34 dB for /a/, /i/ and /u/ vowels respectively with an average fundamental frequency of 104.3 Hz.

Kane and Wellen (1985) using 10 children (6-11 years) with vocal nodules found shimmer values of 0.0151 dB to 0.0911 dB with a mean of 0.0577 dB. Sridhara (1986) studied 30 young normal males and females using /a/, /i/ and /u/ vowels. He reported the following values of mean shimmer in dB.

TABLE 3 : Normative data for DPF (%) obtained by Sorensen and Horii (1984).

Sex	/a/	/i/	/u/
Male	0.033	0.066	0.156
Female	0.7	0.37	0.44

iv) Amplitude (DPF):

Hecker and Kreul (1971) defined Directional perturbation factor for amplitude (Amplitude DPF) as the percentage of peak amplitude difference between adjacent cycles which differed in algebraic sign. So the measure tallies the number of times that the amplitude changes between two successive waves shifts direction. Sorensen and Horii (1984) have studied the amplitude DPF in 40 normal males and females

(2- each) with age ranging from 25 to 49 years using the three vowels /a/, /i/ and /u/. They reported the values as 59.47%, 61.13% and 58.91% for males and 63.13 & and 59.76% for females respectively.

v) Amplitude perturbation Quotient (APQ):

Takahashi and Koike (1976) defined APQ as the variation in signal amplitude measured at the fundamental period divided by the mean amplitude.

$$APQ = \frac{\frac{1}{n-10} \sum_{P=6}^{n-5} \left| \frac{(Ai-5 + Ai-4 + \ldots Ai+5)}{n1} - Ai \right|}{\frac{1}{n} \sum_{i=1}^{n} Ai}$$

In other words, it is the ratio of moving average of peak amplitude differs to the average amplitude (peak) where the moving average is equal to 11 periods.

Takahashi and Koike (1976) reported mean APQ values of 0.00403 and 0.00329 for males and females respectively for sustained /a/ vowel.

vi) Amplitude variability Index (AVI):

According to Deal and Emanuel (1978) PVI is defined as the mean of the squares of the deviations of peak amplitude of each cycle in the sample from the mean peak amplitude divided by the square of the mean peak amplitude, multiplied by 1000.

$$AVI = \log 10 \left[\frac{\frac{1}{n} \sum_{i=1}^{n} \left\{ (Ai - \bar{A})^2 \right\}}{(\bar{A})_2} \times 1000 \right]$$

They evaluated their index with samples of sustained vowels produced at 75 dB SPL for 7 seconds by normal adult males and by clinically hoarse males (20 each) and reported mean AVI of - 0.0619, - 0.1330 and - 0.1287 for /a/, /i/ and /u/ vowels respectively for the normal males and 0.2163, 0.5706 and 0.4142 respectively for the same vowels for clinically hoarse males.

Gauffin, Hammarberg and Hertegard (1996) reported that, if the degree of perturbation is high enough to be perceived as a separate feature, we will usually describe the voice as hoarse, harsh or rough. Jitter and shimmer measurements which are made by comparing the durations and amplitudes of neighbouring periods have been used by many investigators to study this phenomenon. In some cases perceived hoarseness is highly correlated with jitter and shimmer measurements. But it is easy to find voices where this correlation is very low or even negetive.

These investigation have concluded that methods using period to period variability as a way to rate perceptual voice qualities might fail. Differences in the narrow band spectra of the stimuli may explain the differences in perceptual rating of the stimuli. For voices with repetitive patterns in period to period variation, this suggests that a method analyzing spectral characteristics, considering marking and so on, might yield better results than jitter and shimmer methods.

VENKATESH et al (1992) reported Jitter Ratio (JR) Relative Average Perturbation, 3 point (RAP 3), 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 added that in the clinical application, Shimmer is most effective parameter and act like a quick screening device and in pitch perturbation measures like jitter ratio (JR), relative average perturbation (3 point) and DLT are most useful in differentiating laryngeal disorders.

III FREQUENCY AND INTENSITY MEASURES

HANSON, GARRATT 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 i.e. noise, fundamental frequency and intensity variations. This random energy and aperiodicity of fundamental frequency is perceived by the human ears as hoarseness. The aerodynamic parameters measures the respiratory airflow. They do not provide adequate information regarding the voice and its production. Whereas spectral parameters are more appropriate in quantifying the phonatory functions. However, spectral measurements are complex to obtain and the instrumentation is highly sophisticated and expensive. Hence, for clinical purposes these measurements are not desirable. Although intensity related measurements are useful in describing the phonatory function and are relatively easy to measure, the values are highly variable. So, they have reduced reliability. Among the various measurements, the measurements of intensity variations are very useful in early identification and assessment of severity of voice disorder.

They are:

i) Amplitude perturbation (Shimmer)

ii) Extent of fluctuation in intensity

iii) Speed of fluctuation in intensity

A few studies of these acoustic parameters have been carried out for the normals in the Indian population (KUSHAL RAJ, 1984; RASHMI, 1985; RAJANIKANTH, 1986).

a) Fundamental frequency:

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

Of the three major attributes of voice the underlying basis of speech are namely pitch, loudness and quality. ".... 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).

The study of fundamental frequency has important clinical implications. COOPER (1974) had used spectrographic analysis, as a clinical tool to describe and compare the Fo and hoarseness in dysphonic patients before and after vocal rehabilitation. JAYARAM (1975) found a significant difference in habitual frequency measures between normals and dysphonics.

b) Frequency range in phonation and speech:

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

Variations in fundamental frequency and the extent of range used also relate to the intent of the speaker (FAIRBANKS and PRONOVAST, 1939). More specifically, the spread of frequency range used corresponds to the mood of the speaker, that is, as SKINNER (1936) reports, cheerful animated speech exhibits greater range than serious, thoughtful speech.

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

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

NATARAJA (1986) found that the frequency range did not change much with age i.e. in the age range 16-45 years. He also found that females showed a greater frequency range than males in both phonation and speech.

GOPAL (1986) from a study of normal males from 16-65 years, reported slightly lower frequency range in speech.

HANSON, GARRATT and WARD (1983), suggested that majority of phonatory dysfunctions are associated with abnormal and irregular vibrations lead to the generation of random acoustic energy, i.e. noise, fundamental frequency and intensity variations. This random energy and a periodicity of Fo is perceived by human ear3 as hoarseness. Hence, the spectral, intensity and Fo parameters are more appropriate in quantifying phonatory dysfunctions. The frequency related parameters are the most rugged and sensitive in detecting anatomical and sensitive in detecting anatomical physiological changes in the larynx (HANSON, GARRATT and WARD, 1983).

c) Extent and speed of fluctuation in Frequency and Intensity:

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

i) Extent of fluctuation in fundamental frequency

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

ii) Speed of fluctuation in fundamental frequency

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

iii) Extent of fluctuation in intensity

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

iv) Speed of fluctuation of intensity

This was defined as the number of positive peaks on an amplitude display within 1 sec. Peaks of 3 dB or greater from adjacent trough have been counted.

The results of KIM, KAKITA and HIRANO (1982)'s study have indicated that among the above mentioned acoustic parameters significant differences were found between the control and the diseased groups in terms of fluctuation of fundamental frequency. VANAJA (1986), THARMAR (1991) and SURESH (1991) have reported that as the age increases there was increase in fluctuations in frequency and intensity of phonation and this difference was more marked in females.

NATARAJA (1986) found that speed of fluctuation in fundamental frequency and extent of fluctuation in intensity parameters were sufficient to differentiate the dysphonics from the normals.

Correlation between perceptual and acoustic measures:

Many studies have been done to find out the correlation between the perceptual and acoustic measures. Such studies were done by Hartman and Cramon (1984) and Imaizumi (1986). These studies reveal that there is a good correlation between acoustic parameters studied and amount of hoarseness perceived.

vi) Generally, the greater the amount of acoustic perturbation, the more dysphonic the voice (Karnel, 1991).

It is clear from the review that many studies have been done on various acoustic parameters to differentiate normals from hoarse voice. Most of these studies have concluded that the acoustic measures taken could differentiate between normals and hoarseness. Jitter measures were considered for the quantification by Lieberman (1961, 1963), Moore and Thompson (1965) and Hollien et al (1973). Shimmer measures were used by Koike et al (1977) and Deal and Emanuel (1967), Livery and Emanuel (1970); Sansone and Emanuel (1970) and Emanuel et al (1973) studied spectral noise measures for hoarseness quantifications.

Pathak (1995) studied jitter, shimmer and spectral noise measures to quantify hoarseness. He concluded that all these acoustic measures were found to be sensitive to detect hoarseness and were able to classify degree of hoarseness. Jitter ratio, Relative average perturbation, dB shimmer, Amplitude perturbation quotient and Peak harmonics to noise ratio measures together were recommended as a diagnostic test battery for the classification of various degrees of hoarseness. Peak harmonics to noise ratio found to be most sensitive measure for screening. Spectral noise measures were less sensitive to quantify the hoarseness as compared to perturbation measures.

vi) Generally, the greater the amount of acoustic perturbation, the more dysphonic the voice (Karnel, 1991).

It is clear from the review that many studies have been done on various acoustic parameters to differentiate normals from hoarse voice. Most of these studies have concluded that the acoustic measures taken could differentiate between normals and hoarseness. Jitter measures were considered for the quantification by Lieberman (1961, 1963), Moore and Thompson (1965) and Hollien et al (1973). Shimmer measures were used by Koike et al (1977) and Deal and Emanuel (1967), Livery and Emanuel (1970); Sansone and Emanuel (1970) and Emanuel et al (1973) studied spectral noise measures for hoarseness quantifications.

Pathak (1995) studied jitter, shimmer and spectral noise measures to quantify hoarseness. He concluded that all these acoustic measures were found to be sensitive to detect hoarseness and were able to classify degree of hoarseness. Jitter ratio, Relative average perturbation, dB shimmer, Amplitude perturbation quotient and Peak harmonics to noise ratio measures together were recommended as a diagnostic test battery for the classification of various degrees of hoarseness. Peak harmonics to noise ratio found to be most sensitive measure for screening. Spectral noise measures were less sensitive to quantify the hoarseness as compared to perturbation measures.

The review of literature reveals that inspite of several studies have been done on hoarse voice, none of a single study attempted to cover all four acoustic measures (spectral, perturbation, frequency and intensity measures).

Further, none of these studies have clearly indicated the parameters clinically useful in differentiating and quantifying hoarse voice from normal voice. Thus this study aims to identify, differentiate and quantify the hoarseness using spectral noise measures, perturbation measures and frequency and intensity measures.

The following parameters were taken for this study.

A) Spectral measures

1) Harmonics to Noise Ratio (HNR)
2) H_1 amplitude
3) Number of Harmonics
4) Alpha, Beta, Gamma ratios of LTAS

B) Perturbation Measures

i) Jitter Measures:

1) Mean Jitter
2) Percent Jitter
3) Period Variability Index (PVI)
4) Relative Average Perturbation (RAP)
5) Directional Perturbation Quotient (DPQ)
6) Deviation from Linear Trend (DLT)

ii) Shimmer Measures:

1) dB Shimmer
2) Amplitude Variability Index (AVI)
3) Amplitude Perturbation Quotient (APQ)
4) Amplitude DPQ

C) Frequency Measures

1) Mean Fo
2) Range of frequency
3) Extend of Fluctuation in frequency
4) Speed of Fluctuation in frequency

D) Intensity measures

1) Mean intensity
2) Range of intensity
3) Extend of fluctuation in intensity
4) Speed of flutuation in intensity

Therefore, the purpose of present study was to identify the parameters clinicaly useful in differentiating and quantifying horse voice from normal voice.

METHODOLOGY

The purpose of present study was to

1. differentiate hoarse voice from normal voice
2. classify the hoarse voice.
3. compare perceptual estimation with acoustic estimation of hoarse voice.

using the combination of spectral, perturbation, frequency and intensity measures. A total of 24 parameters as listed below were considered for the study.

A) Spectral measures

1) Harmonics to Noise Ratio
2) First Harmonic Amplitude
3) Number of Harmonics
4) Alpha, Beta, Gamma ratios of LTAS

B) Perturbation Measures

i) Jitter Measures:

   1) Mean Jitter
   2) Percent Jitter
   3) Period Variability Index
   4) Relative Average Perturbation
   5) Directional Perturbation Quotient
   6) Deviation from Linear Trend

ii) Shimmer Measures:

1) dB Shimmer
2) Amplitude Variability Index
3) Amplitude Perturbation Quotient
4) Amplitude Directional Perturbation Quotient

C) Frequency Measures

1) Mean Fundamental Frequency
2) Range of Frequency
3) Extend of Fluctuation in Frequency
4) Speed of Fluctuation in Frequency

D) Intensity measures

1) Mean Intensity
2) Range of Intensity
3) Extend of Fluctuation in Intensity
4) Speed of Flutuation in Intensity

SUBJECTS:

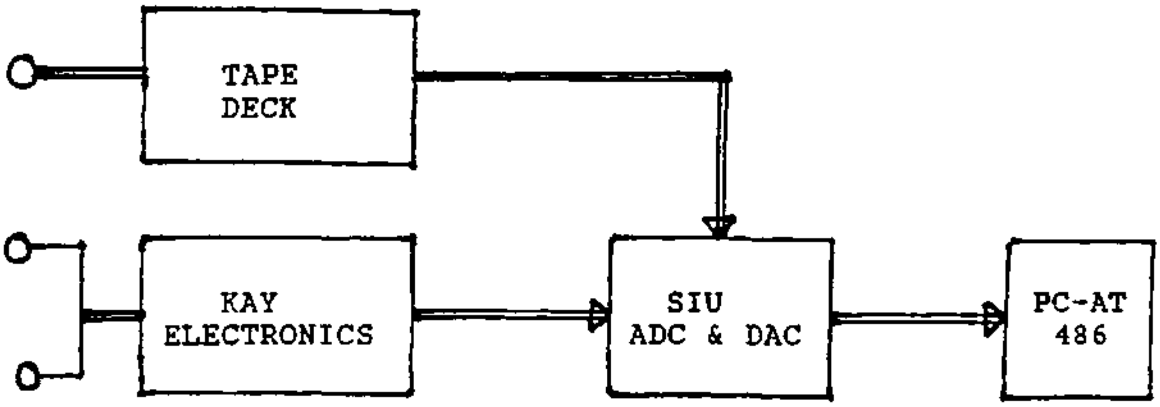
50 normals (25 each sex) with an age range of 18 to 30 years and 30 dysphonics (15 each sex) with an age range of 20 to 50 years were studied. The group of dysphonics were chosen from among the patients who visited the All India Institute of Speech & Hearing with the complaint of voice problem. These cases were diagnosed as having voice disorder after the routine speech science and speech pathology, otolaryngalogical and pyschological examinations. The normal subjects selected for the study were also evaluated by qualified Speech Language Pathologist and considered normal.

PROCEDURE:

Instrumentation:

The instruments used in the study were:

1. Unidirectional microphone (H-Legend D 800)
2. Sony Tape deck (TC FX 170)
3. Meltrack CR-X90 audio-cassette.
4. Portable Electrolaryngograph (Kay Elemetrics corporation)
5. 12 bit ADC with speech interface unit (VSS)
6. Computer (Pentium 200 MHz Processor)
7. Headphones (Sukawa)
8. Philipamp 60.



Block diagram of instrumental setup:

Recording the Data:

Phonation of \a\, \i\ & \u\ for about 5 to 6 seconds and the reading of a paragraph of rainbow passage were recorded on an audio cassette and also on the computer using record programme.

The following instructions were given to each subject prior to recording:

"Say \a\ three times each with comfortable loudness until I signal you to stop. Then read this paragraph, pointing to the paragraph to be read. Do not move unnecessarily".

An unidirectional microphone was connected to the tape deck (Sony TC FX170) with mouth to mic distance being about 6 cms (in order to reduce the ambinent noise level from the surronding). In this set up, phonation of vowels and speech samples were recorded on Meltrak CR-X90 audio cassette.

STEP - I

Electrodes connected to Electroglottograph were placed on both the thyroid alae of the subject. The electro glottograph was connected to speech interface unit (VSS) which in turn was connected to the PC -AT 486 computer. In this set up, glottal signal of phonation of vowel /a/ were recorded for duration of 2 seconds. Thus simultaneous recording of glottal signals on computer and the acoustic signal on the tape recorder were carried out. Precaution was taken to see that the signal was not distorted. This was done by asking the subjects to monitor the loudness of voice based on LED lights on the fron pannel of the Speech Interface Unit. The subject was also instructed not move to the head during recording.

STEP - II

Then, with the same setup of microphone connected to computer through Speech Interface Unit, instead of Electroglottograph. Following the same procedure reading of rainbow passage as speech sample was recorded for 10 seconds. While recording on audio cassette and on the computer, the microphone was placed at 70° angle of incidence because the microphone had a flat frequency response (± 1dB) from 30Hz to 15KHz at this angle.

Similarly /i/ and /u/ vowels were also recorded. Thus the phonation and glottel signals in Step - I and Speech signal in Step - II were recorded simultanesly on tape recorder and computer. Thus the glottal signal, phonation and speech signals for each subjects of both the groups were recorded. At the same time the phonation was recorded on a Sony (TC FX170) cassette deck using a microphone which was kept at a distance of 6 cms from the mouth of the subjects.

ACOUSTIC ANALYSIS:

The signal was modified by a low pass fitter with a cut-off frequency of 7.5KHz and a roll-off rate of 48dB/0ctave. The filtered signal was digitized with 12 bit precision at a sampling rate of 16KHz. The amplitude of the input signal was adjusted not to exceed the full scale of ADC and it was stored on the hard disk of computer for further acoustic analysis.

The voice signal thus stored on the hard disk of the computer was submitted for analysis using the "Inton Analysis" of VAGHMI (VSS, Bangalore). The phonation signal was read in blocks or frames of 40 mesc. duration each. Autocorrelation technique was used to estimate the average Fo over this block of 40 msec. Intensity was measured as the RMS value in dB. Successive blocks were spaced by 10 msec. The minimum and maximum limits for Fo measurement were 50 and 800 Hz. The analysis of the voice signal yielded the following parameters.

a) Frequency Measures

1) Mean Fundamental Frequency
2) Range of frequency
3) Extend of Fluctuation in frequency
4) Speed of Fluctuation in frequency

b) Intensity measures

1) Mean intensity
2) Range of intensity
3) Extend of fluctuation in intensity
4) Speed of Fluctuations in intensity

PERTURBATIONS:

The purpose of this part of the study was to analyse the frequency and amplitude perturbations during phonation. The voice samples of the subjects (sustained phonation of /a/)

reocrded using Electroglotography were used for this. The analysis was carried out with the help of 'JITSHIM' program, a module of VAGHMI software (Voice and Speech Systems, Bangalroe). The following parameters were obtained for each:

a) Jitter Measures:

1) Mean Jitter
2) Percent Jitter
3) Period Variability Index
4) Relative Average Perturbation
5) Directional Perturbation Quotient
6) Deviation from Linear Trend

b) Shimmer Measures:

1) dB Shimmer
2) Amplitude Variability Index
3) Amplitude Perturbation Quotient
4) Amplitude Directional Perturbation Quotient

Procedure:

The recording that was made for Laryngography was used for this part of experiment. A sample duration of approximately 5 seconds was fed to the JITSHIM program of VAGHMI. The computer displayed the results in terms of above parameters.

Long Term Average Spectrum (LTAS):

To compute LTAS and measure relevant parameters digitized speech data (three sentences from standard passage) which was already stored on the hard disk of the computer was used. For analysis of this data the LTAS module of VAGHMI was used. The speech signal was read in blocks or frames of aobut 16 msec. FFT Technique was used to computer magnitude squared spectrum on this block of 16 msec. Successive blocks or overlapped by 8 msec, spectra were accumulated. At the end of the specific duration, the average was determined. The spectral analysis thus obtained contained a graphic display of spectral patterns in the frequency range of 0-8 kHz. The data of energy of all the different points which were analysed by the computer were:

1. Alpha Ratio

   It is the ratio of energy betwen 0-lKHz and l-5KHz.

2. Beta Ratiio

   It is the ratio of energy betwen 0-2KHz and 2-5KHz.

3. Gamma Ratio

   It is the ratio of energy betwen 0-lKHz and 5-8KHz.

Procedure:

The speech signal of approximately ten seconds was used for analysis. The LTAS program calculated and displayed the results in the form of Alpha, Beta and Gamma Ratios.

Harmonics to Noise Ratio:

The purpose was to compare the relative intensities of harmonics and noise in the voice source. Sample of sustained /a/ used for acoustic analysis earlier were used here. These were fed to the program "HNR" of VAGHMI (VSS, Bangalore). Phonation signal was divided into blocks or frames of 40 msec, duration each. Successive frames were spaced by 10 msec, giving 100 measures per second. For each frame, Cepstrum was computed. The amplitude of the peak was assumed to represent the strength of the harmonics. The strength of harmonics was represented in dB scale. The (square) root mean (average) of squared (rms) value of the noise sample was computed and expressed in dB. The Harmonics to noise ratio (HNR) was the difference in dB of the two (signal and noice).

Procedure:

The phonation signal of approximately 5 seconds (sustained phonation of /a/) was used for analysis. The HNR program calculated and displayed the results of harmonics to noise ratio on the screen.

First Harmonic Amplitude and Number of Harmonics:

The purpose was to measure the amplitude of First Harmonics and the Number of Harminics present in the glottal signal. The voice samples of sustained phonation of /a/ recorded using Electroglottography were used for this purpose. For analysis of this data, the LTAS module of VAGHMI was used.

Procedure:

The glottal signal of approximately 5 seconds was used for analysis. The LTAS program calculated and displayed the glottal wave form on the screen. The cursor was moved and placed at the peak of first harmonic to measure First Harmonic Amplitude. The number of harmonics visible and clearly present in the glottal wave form was manually counted to measure the Number of Harmonics.

PERCEPTUAL ANALYSIS:

The phonation and speech samples of both normals and dysphonics were randomly dubbed into another audio cassette. No identity was revealed about the subject on dubbing, except code number, age and sex. For intrajudge reliability check, 10 samples were randomly selected and recorded second time.

Seven judges (4 males and 3 females) from M.Sc Speech and Hearing students were selected for perceptual evaluation. Scoring was done on a 4 point scale (1-Normal, 2-mild, 3-moderate, 4-severe hoarse voice). For this purpose, perception lab in Speech Science department of All India Institute of Speech and Hearing was used. The samples were presented through ear phones (Sukawa) at a comfortable loudness.

STATISTICAL ANALYSIS:

Intrajudge Reliability was computed by taking the ratings of repeated judgements of 10 samples out of 80 samples and by finding the Pearson's coefficient of correlation (r) of the two judgements for each judge. Interjudge reliability was computed by finding the Person's coefficient of correlation (r) of the judgements of all the ratings for all judges.

SPSS (Statistical Package of Social Science) program was used for descriptive and descriminant analysis. Descriptive analysis was done to calculate the mean and the standard deviation and range of parameters. Then the data was treated with Paired Sample-T Test to find out the significance of difference of means and standard deviations of all parameters between and within the groups.

Further, the data was treated with Canonical Discriminant analysis for classification of parameters within and across groups.

RESULTS AND DISCUSSIONS

The purpose of the present study was:

1. Differentiating hoarse voice from normal voice
2. Classifying the hoarse voice
3. Comparing perceptual estimation with acoustic estimation of hoarse voice.

using the combination of spectral, perturbation, frequency and intensity measures. A total of 24 parameters were studied. Further subjective rating of severity of hoarseness has been used for determining parameters related to perception of severity of hoarsseness.

The results of this study has been presented as under:

A. Spectral parameters
B. Perturbation parameters.
C. Frequency related parameters.
D. Intensity related parameters.

PERCEPTUAL ANALYSIS:

The rating of severity based on perceptual judgement was done on a four point scale employing 7 judges for all voice and speech samples. The ratings by four or above judges for each sample was considered as the rating of hoarseness for that sample.

Judges	r
J1	1.0
J2	0.98
J3	0.87
J4	0.94
J5	0.93
J6	0.83
J7	0.89

TABLE 4: Shows correlation (r) used for varifying intra Judge variability.

To check the intrajudge reliability Karl Pearson's co-efficient correlations was done. From the above table, it was clear that there was a high correlation between the repeated ratings made by each judge (+1.0 to +0.83). Further the inter Judge realiability showed high corealation (0.95 to 0.63). Hence the evaluation of these judges were considered reliable.

(A). SPECTRAL PARAMETERS

(i). Alpha Ratio (AR)

The comparisions within and accross the groups were made with 'Paired T test'. The test revealed the following results for Alpha Ratio.

	NORMALS		DYSPHONICS	
	MALE	FEMALE	MALE	FEMALE
MEAN	48.83	47.14	45.45	72.88
S.D.	19.03	36.54	25.45	51.27
RANGE	72.71	131.06	101.51	189.54

TABLE 5 : Shows Mean, S.D and Range for Alpha ratio in normals and dysphonic3.

The values for dysphonic females were higher compared to of dysphonic males. There was a significant difference between dysphonic males and females only (t=2.314). The dysphonics females showed a higher ratio than the males. However, there was no statistically significant difference between the values of normal males and females (t=0.201). This parameter did not show a significant difference between normal and dysphonic males (t= -0.119; -1.465). However there was a significant difference between females and dysphonic groups. Therefore the hypothesis (1) stating that there is no significant difference between normal and dysphonic males in terms of Alpha ratio has been accepted, whereas with respect to females it has been rejected. Further, the hypothesis (2) stating that there is no significant difference between males and female has been accepted with reference normals, whereas it was rejected with regard to dysphonic male and females, as females had higher values than males. Hence could not be used to differentiate hoarse voice vs normal voice in case of males.

Rajashekhar (1991) gave the value of alpha ratio as 4.0 for normal adult males. Wedin and Argen (1982) used this measure to quantify the progress in voice training program. Other authors have suggested that this parameter is useful in identifying hoareseness (Fritzell and Hammer berg 1973; Wendler and Doherty and Hollein 1980; Nataraja 1986; Rajkumar 1998). There have been only few studies on the importance of quantifying hoarseness with this parameter and in this study this parameter has not found to be useful in quantify hoarse voice.

(ii). BETA RATIO (BR):

The values of normal males and females and dysphonic males and females are shown in the Table - 6.

	NORMALS		DYSPHONICS	
	MALE	FEMALE	MALE	FEMALE
MEAN	126.03	118.73	147.75	190.98
S.D.	57.74	88.66	79.14	70.65
RANGE	233.31	449.14	566.02	550.88

TABLE 6: Shows mean, S.D , Range for Beta ratios in normals and dysphonics.

The values didn't show stastically significant difference in any of the comparisions made i.e..

> normal males vs normals females — (t = 0.271)

> dysphonic males vs dysphonic females — (t = 0.1004)

> normal vs dysphonic males — (t = 0.702)

> normal vs dysphonic females — (t = -1.463)

The males showed a higher ratio values in the normals group. When compared to females of the respective group. Whereas the females of dysphonic group had higher values than males of the group.

Study by Rajkumar (1998) revealed a significant difference between normals and pathological cases both in males and females on this measure. Kiar, Kakita and Hirano (1982) also showed similar reults. But the present study was not in agreement with the other studies as it has not been found useful in differentiating normal voice from dysphonic voice.

(iii) GAMMA RATIO (GR):-

The gamma ratio for the normals and dysphonics are shown below:-

	NORMALS		DYSPHONICS	
	MALE	FEMALE	MALE	FEMALE
MEAN	4868.80	2119.35	3129.18	4058.56
S.D.	993.86	652.15	498.39	499.48
RANGE	4583.99	4858.12	3466.85	660.42

TABLE 7: Shows mean, S.D and Range for gamma ratio in normals and dysphonics.

This value showed a significant difference between normal males and females (t = 5.202) and normal males and dysphonic males (t = 2.202) but not between dysphonic males and females (t = 0.737) and between dysphonic and normal females (t = -1.696), there was a difference which was statistically significant. Thus this parameter was not found to be useful in differentiating between normals and dysphonics.

Gamma ratio has been reported to indicate degre of hoarseness as it was similar to the other spectral measures (Kior, Kahita and Hirano 1984 ; Rajkumar Pandit 1998) Though there was a significant difference between normal and pathological males, this measure could not be relied upon to differentiate in both the sexes and hence could not be used as a clinical tool to identify the hoarseness.

(iv). Harmonics to Noise Ratio (HNR):

	NORMALS		DYSPHONICS	
	MALE	FEMALE	MALE	FEMALE
MEAN	25.97	27.89	21.40	23.82
S.D.	3.06	3.32	5.21	3.49
RANGE	15.54	16.37	21.16	12.68

TABLE 8: Shows the means, S.D and Range for normals and dysphonics for HNR ratio.

Like in the previous studies reported in the literature, this study also revealed a statistically significant difference between normal males and dysphonic males (t = 5.182) and normal females and dysphonic females (t = 5.800) and also between normal males and females (t = -4.618) and dysphonic males and females (t = 2.135).

Therefore the hypothesis (1) stating that there was no significant difference between normals and dysphonics in terms of Harmonic to Noice Ratio was rejected. Further, the hypothesis (2) stating that there is no significant difference between males and female in terms of Harmonic to Noice Ratio was rejected.

Kine, Kakita, Hirano (1982) showed that the characteristic feature of hoarseness was the replacement of harmonics by noise energy. Pathak (1997) found Harmonic to Noice Ratio values in normal males as 26.51 and 27.82 for females. Rajkumar (1998) found the values to be 24.92 for males and 27.33 for females. In the present study the females of both the groups had showed higher values.

Hence this parameter could be used as a reliable tool to differentiate normals and dysphonics.

(v) FIRST HARMONIC AMPLITUDE (H1A):

The values obtained by the 2 groups are given below.

	NORMALS		DYSPHONICS	
	MALE	FEMALE	MALE	FEMALE
MEAN	84.37	83.93	83.85	83.62
S.D.	0.28	0.78	0.77	0.90
RANGE	1.45	3.03	3.93	3.62

TABLE 9: Shows mean, S.D and Range obtained for First Harnomic Amplitude in normals and dysphonics.

From the study of above table it was clear that there was no difference in values of First Harmonic Amplitude between sexes in both the groups. Statistical analysis of this parameter revealed significance of difference across three comparisons. The comparison showed significant difference between normal males and females (t = 4 .927) ; normal males and dysphonic males (t = 4.815) and normal females and dysphonic females (t = 2.676). This parameter did not show significant difference between the dysphonic males and females.

Therefore the hypothesis (1) stating that there was no significant difference between normals and dysphonics in terms of First Harmonic Amplitude was rejected. Further, the hypothesis (2) stating that there is no significant difference between males and females was rejected with reference to normal, whereas it was accepted in dysphonicmales and females.

Fairlbanks (1940) found enhanced First Harmonic Amplitude in breathy voice. Other investigators (Bicklay 1982, Fischer and Jorgenson 1967) had also reported similar findings. Hillen brand et.al (1994) found that relative amplitude of the first harmonic correlated moderately with breathiness ratings. The present study showed significant difference between normal and dysphonic groups in both sexes.

(vi) Number of Harmonics (NOH);

	NORMALS		DYSPHONICS	
	M	F	M	F
Mean	20.29	22.93	12.31	13.78
S.D	4.86	4.63	6.11	6.2
Range	31.00	31.00	22.00	22.00

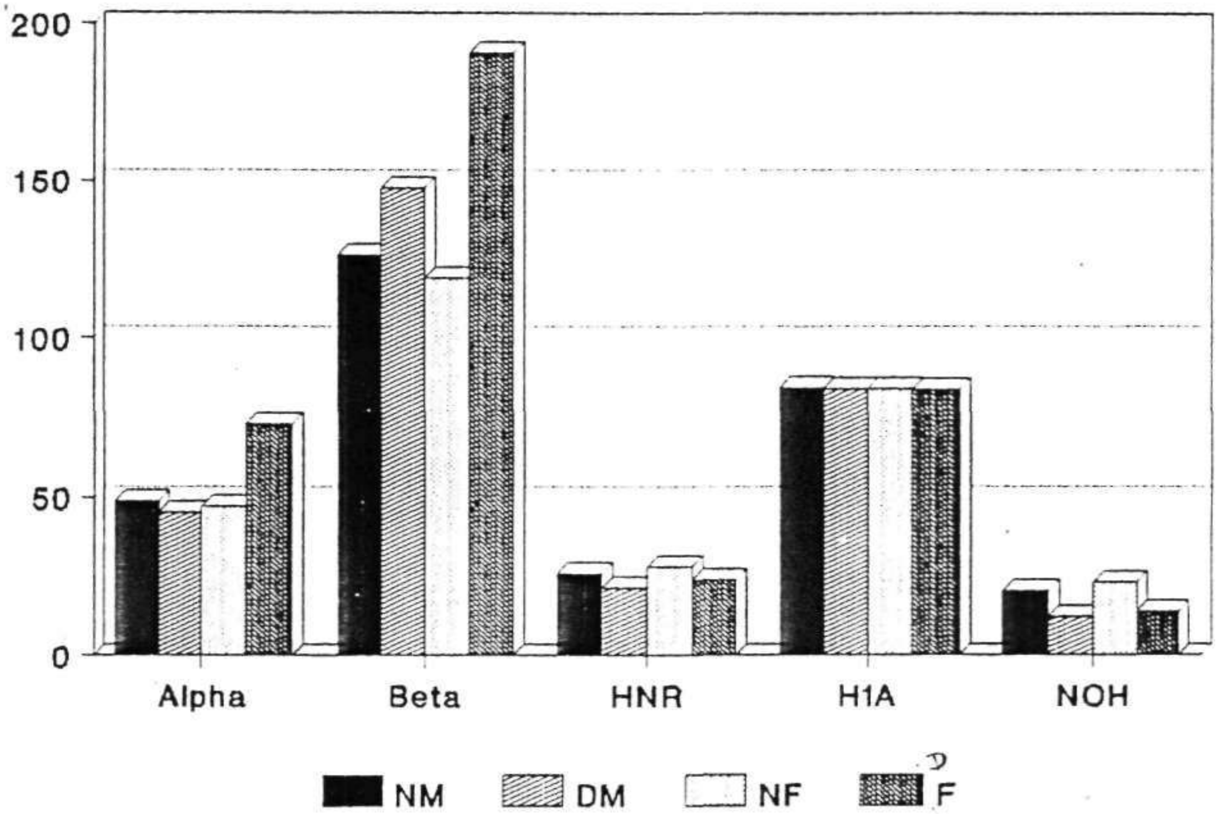
TABLE 10: Indicates Mean, S.D and Range obtained for Number of harmonics in normals and dysphonics.

The dysphonics in the present study were found to had lesser number of harmonics compared to normals which was due to the replacement of noise in the spectrum. The males had slightly lesser number of harmonics than females in both the groups.

Analysis of this measure revealed that it differentiated the normal and dysphonic males (t = 8.342) but not the normal females and dysphonic females (t = 0.251). No significant difference was found between normal males and females (t = -1.073) and between dysphonic males and females (t = 1.121). Since this did not differentiate normals and dysphonic3 of both the sexes, it was not considered useful in differentiating normals from dysphonics.

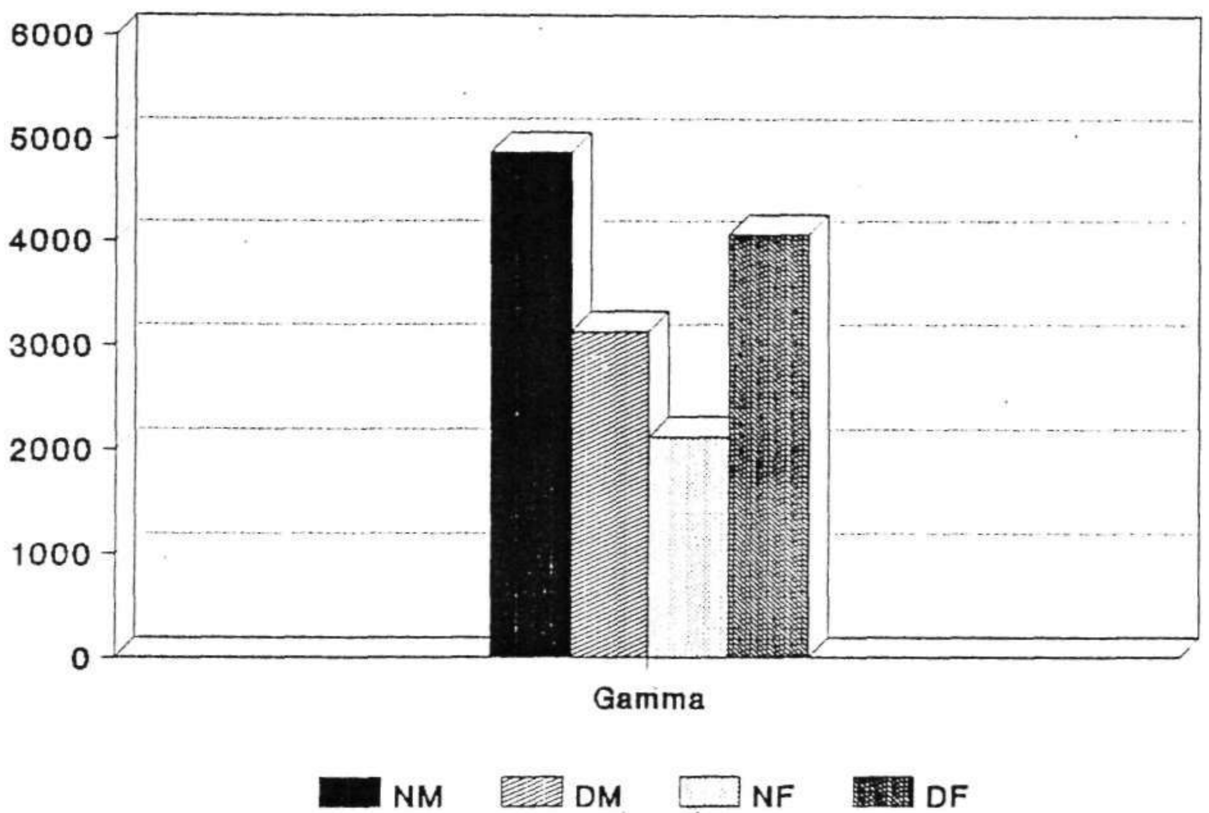
Therefore the hypothesis (1) stating that there was no significant difference between normal and dysphonic males in terms of Number of Harmonics was rejected, but accepted for females. Further the hypothesis (2) stating that there is no significant difference between males and female in normals and dysphonics was rejected.

SPECTRAL MEASURES



Graph 1: Showing Mean values of Alpha & Beta Ratios, Harmonics to Noise Ratio (HNR), First Harmonic Amplitude (H1A) and Number of Harmonics (NOH) for Normal Males (NM), Dysphonic Males (DM), Normal Females (NF) and Dysphonic Females (DF).

SPECTRAL MEASURES



Graph 2: Showing Mean values of Gamma Ratio for Normal Males (NM), Dysphonic Males (DM), Normal Females (NF) and Dysphonic Females (DF).

B) PERTURBATION PARAMETERS:

The results of perturbations in frequency and intensity are discussed below

JITTER MEASURES:

(i) Jitter Mean Fundamental Frequency (JFO):

The scores obtained for both sexes of normals and dysphonics are tabulated below.

	Normals		Dysphonics	
	M	F	M	F
Mean	127.23	237.24	155.85	212.82
S.D.	11.34	25.15	54.41	43.80
Range	51.37	114.20	231.03	205.50

TABLE 11: The Mean , S.D and Range of Mean Fo for normals and dysphonics.

The scores obtained by females of both normal and dysphonic groups had higher values than males. Significant differences were found between normal males and females (t = - 34.943), dysphonic males and females (t = 4.340), normal males and dysphonic males (t = - 3.229), and normal females and dysphonic females (t = 2.630). Therefore the hypothesis (1) stating that there was no significant difference between normals and dysphonics in terms of Jitter Mean Fundamental Frequency was rejected. Further, the hypothesis (2) stating that there is no significant difference between males and female with reference to normals and dysphonics was rejected. Hence, Jitter mean Fo could be used to differentiate normals and dysphonics.

(ii) Percent Jitter (PJ):-

	Normals M	Normals F	Dysphonics M	Dysphonics F
Mean	2.63	6.86	3.90	9.11
S.D	0.88	4.26	4.71	7.95
Range	3.88	16.93	32.23	35.49

TABLE 12: The mean, S.D and Range for Oercent Jitter.

The values obtained by females of normal and dysphonic groups were greater than the values obtained by males of the respective groups. The comparison between males and females of normals (t = -8.342), and between males and females of dysphonics (t = 3.516) showed significant differences. A comparison between normal males and dysphonic males (t = -1.650) and between normal females and dysphonic females (t = - 1.235) showed no statistically significant difference. Hence this parmeter was not considered useful in differentiating normals from dysphonics.

Therefore the hypothesis (1) stating that there was no significant difference between normals and dysphonics in terms of Percent Jitter was rejected. Further, the hypothesis

(2) stating that there is no significant difference between males and females in terms of normals and dysphonics was rejected.

(iii) Jitter - Period Variability Intex (JPVI)

	Normals		Dyphonics	
	M	F	M	F
Mean	0.45	6.63	1.06	13.44
S.D.	0.22	12.34	1.01	21.62
Range	1.08	76.83	5.94	79.10

TABLE 13: The mean, S.D and Range of Period Variability Index.

The values observed in females of both normal and dysphonic groups were higher than males of respective groups. This parameter was found statistically significant on all four comparisons made i,e., males vs females of normal groups (t = -4.336), males vs females of dysphonics (t = 3.889), males of normal vs dyphonic males (t = - 3.858) and females of normals and females of dyphonic groups (t = -2.006).

Therefore the hypothesis (1) stating that there was no significant difference between normals and dysphonics in terms of Jitter Directional Perturbation Quotient was rejected. Further, the hypothesis (2) stating that there is no significant difference between males and female in terms of normals and dysphonics was rejected.

Deal and Emanuel (1978) reported mean normative PVI of 0.4412, 0.4898 and 0.4451 for \a\,\i\, and \u\ vowels respectively for 20 adult males. In dysphonics, they found the mean values as 0.8295. Rajkumar (1998) found mean values of 6.69 and 13.41 for normal males and females respectively and 25.12 and 23.45 for dysphonic males and females respectively. The findings of this study correlated with the findings of Deal and Emanuel (1978). Hence, PVI was useful to differentiate normals from dysphonics.

(iv) Jitter - Directional Perturbation Quotient (JDPQ):

	Normals		Dysphonics	
	M	F	M	F
Mean	69.25	67.53	69.67	68.79
S.D	4.43	5.39	5.46	5.15
Range	25.59	31.44	28.17	30.69

TABLE 14: The values of mean, S.D and Range of Directional Perturbation Quotient.

The mean values obtained for females in both normal and dysphonic group were lower than the values of males of respective groups. There was no significant difference in the scores of males and females of normal and dysphonic groups. The significant difference was found only between males and females of normal group (t = 1.995), while no statistically significant difference was found in other three comparisons i.e. between males and females of dysphonic group (t = -0.754), between males of normal and dysphonic groups (t = - 0.236), and between females of normal and dysphonic groups (t = - 1.245).

Therefore the hypothesis (1) stating that there was no significant difference between normals and dysphonics in terms of Jitter Directional Perturbation Quotient was accepted. Further, the hypothesis (2) stating that there is no significant difference between males and females in terms

of normals was rejected, whereas it was accepted for dysphonics.

Hecker and Kreul (1971) found that DPQ ranged from 27.7% to 39.2% (average 33.3%) in normals and from 42.5% to 54.0% in dysphonics. Murry and Doherty (1980) found that the mean DPQ for normals was 58.5% with a range of 45.8% to 65.3% and the subjects with laryngeal cancer had a range of 55.1% to 76.7% and a mean value of 64.5%. Izdebski and Murry (1980) found DPQ of 58.4% for normal adult. Sorenson and Horii (1984) reported the value of DPQ as 46.24 and 48.79 in normal males and females. Rajkumar (1998) reported the DPQ mean values of 66.15 and 72.46 for normal males and females respectively and 68.65 and 66.76 for dysphonic males and females. The findings of the present study were correlating with the findings of Rajkumar (1998) and were found to be slightly higher than the findings of Murry and Doherty (1980). Since, this parameter significantly differed between both sexes of normal group alone, it was not considered as a useful parameter to differentiate between normals from dysphonics.

v) Jitter-Relative Average Perturbation - 3 point (JRAP)

	Normals		Dysphonics	
	M	f F	M	F
Mean	0.0159	0.0395	0.0204	0.0531
S.D.	0.0063	0.0231	0.0088	0.0422
Range	0.02	0.08	0.99	0.17

TABLE 15: The values of mean, S.D and Range of JRAP

The comparison of mean values of between males and females within the normal and dysphonic groups showed higher values for females than in males in both the groups. This was true with the comparison of males and females between normal and dysphonic groups. There was no significant difference seen on any of the comparison of both sexes within and between the normal and dysphonic groups. i.e., in normals, between males and females (t=0.211); in dysphonics, between males and females (t=0.466); in males between normals and dysphonics (t=0.417) and in females, between normals and dysphonics (t= -1.597).

Therefore the hypothesis (1) stating that there was no significant difference between normal and dysphonic in terms of Jitter Realative Average Perturbation was acepted. Further, the hypothesis (2) stating that there is no significant difference between males and female was accepted.

Investigator	Normals M	Normals F	Dysphonics M	Dysphonics F
1) Koike (1973)	0.0046	0.0046	-	-
2) Takahashi and Koike (1975)	0.0057	0.0061	-	-
3) Krishnan (1992)	0.0062	0.0058	-	-
4) Pathak (1997)	0.0052	0.0052	0.0811	0.0531
5) Rajkumar (1998)	0.140	0.250	0.750	0.460
6) Present study (1998)	0.0159	0.0395	0.0204	0.0531

TABLE 16: indicates the mean values of other studies of JRAP

It was clear from the 3tudy of above table that the findings of present study was slightly higher than the findings of other investigators except for findings of Rajkumar (1998). Thus the values of the present study valid but they were not useful in differentiating between dysphonics from normals.

vi) Jitter Deviation from Linear Trend (JDLT):

	Normals M	Normals F	Dysphonics M	Dysphonics F
Mean	0.158	0.233	0.256	1.566
S.D	0.060	0.152	0.448	7.998
Range	0.27	0.58	2.96	53.96

TABLE 17: Shows the mean, S.D and Range of JDLT.

The mean values of females of both normals and dysphonics were graeter than that of males of both the groups. Between normal and dysphonic groups, dysphonic males and females showed higher values than normal males and females.

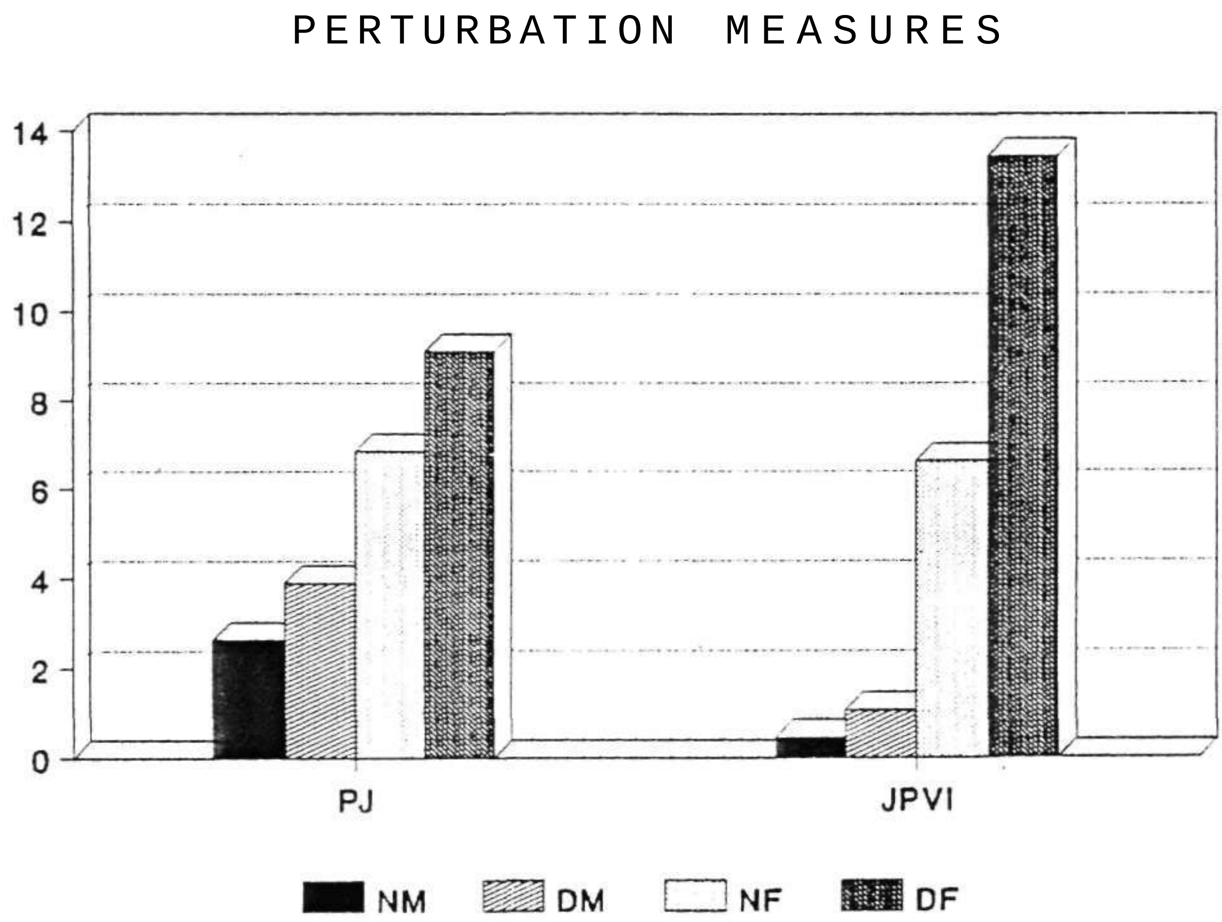
On comparison of both sexes in normal group it was found that there was significant difference between the two (t= -4.108). While the other comparisons between males and females in dysphonic group (t= 1.094), between males of normal and dysphonic groups (t= - 1.330) and between females of normal and dysphonic groups (t= - 1.100) indicated no significant difference between the groups.

Therefore the hypothesis (1) stating that there was no significant difference between normal and dysphonic in terms of Jitter Deviation from Linear Trend was accepted. Further, the hypothesis (2) stating that there is no significant difference between males and female with reference to normals was rejected, whereas it was accepted for dysphonics.

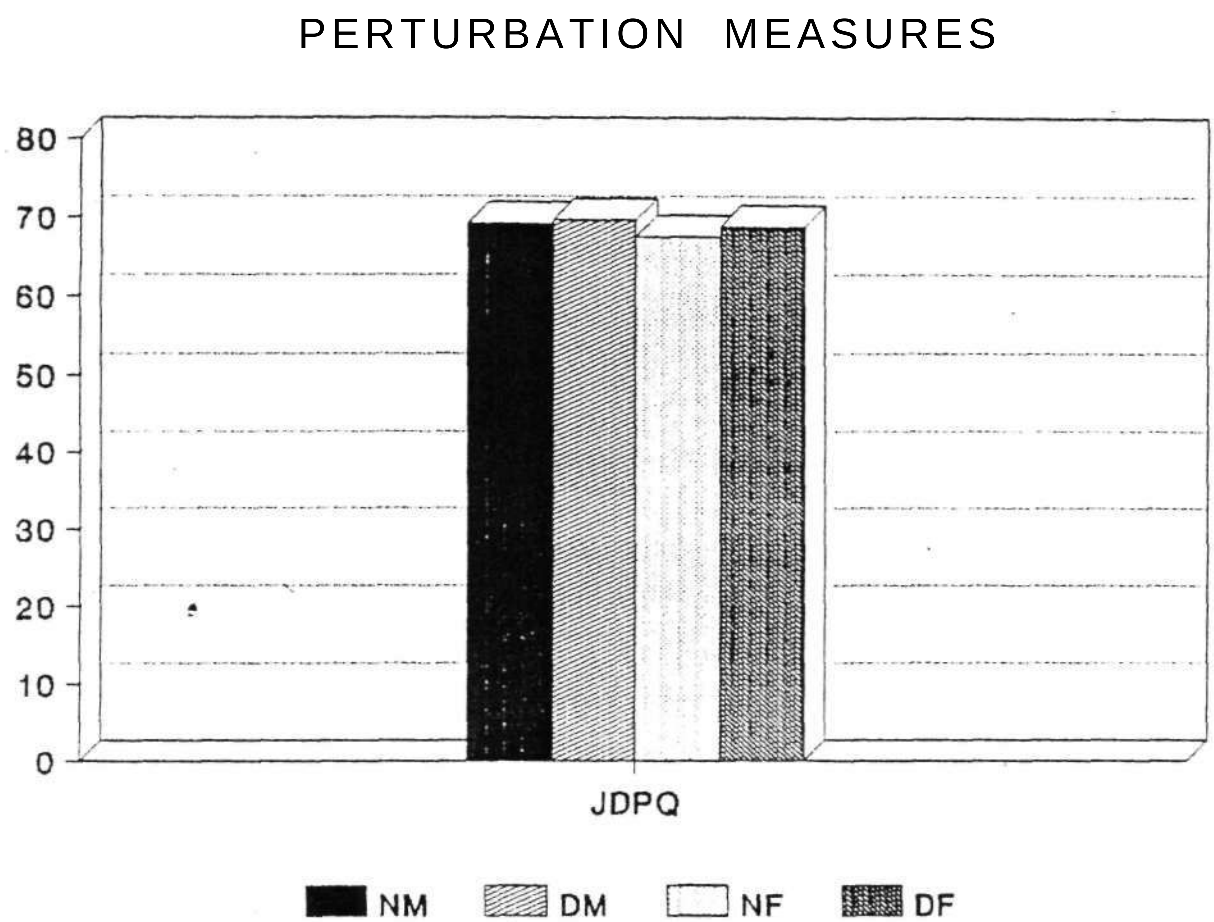
Investigator	Normals		Dysphonics	
	M	F	M	F
1) Ludlow et.al (1988)	28.43	28.43	35.76	35.76
2) Rajkumar (1998)	0.140	0.250	0.750	0.460
3) Present study (1998)	0.158	0.233	0.256	1.566

TABLE 18: Shows the mean values as found in other studies of JDLT.

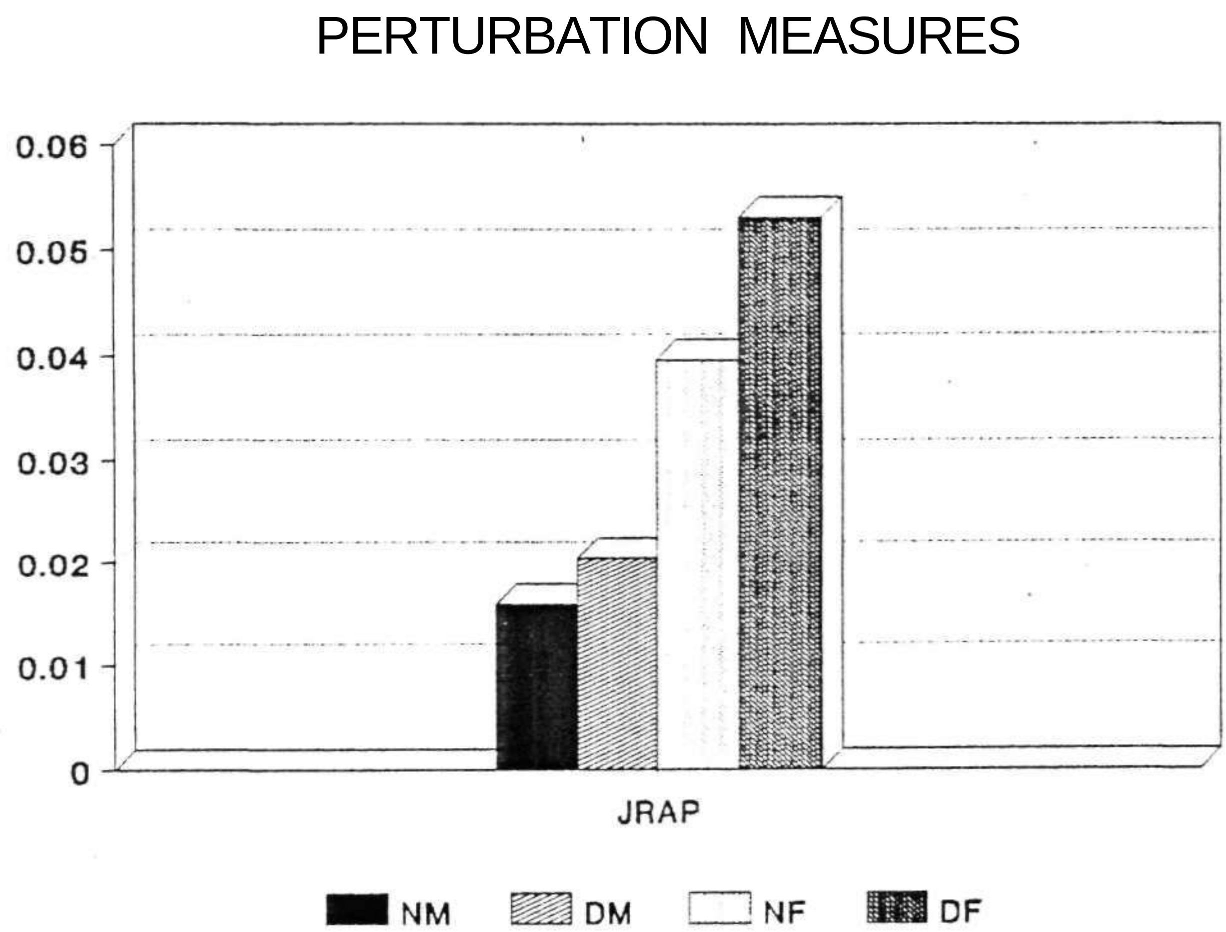
The values of this study correlated with the values of both sexes in normal group and showed slight variations in dysphonic groups from the findings of Rajkumar (1998).



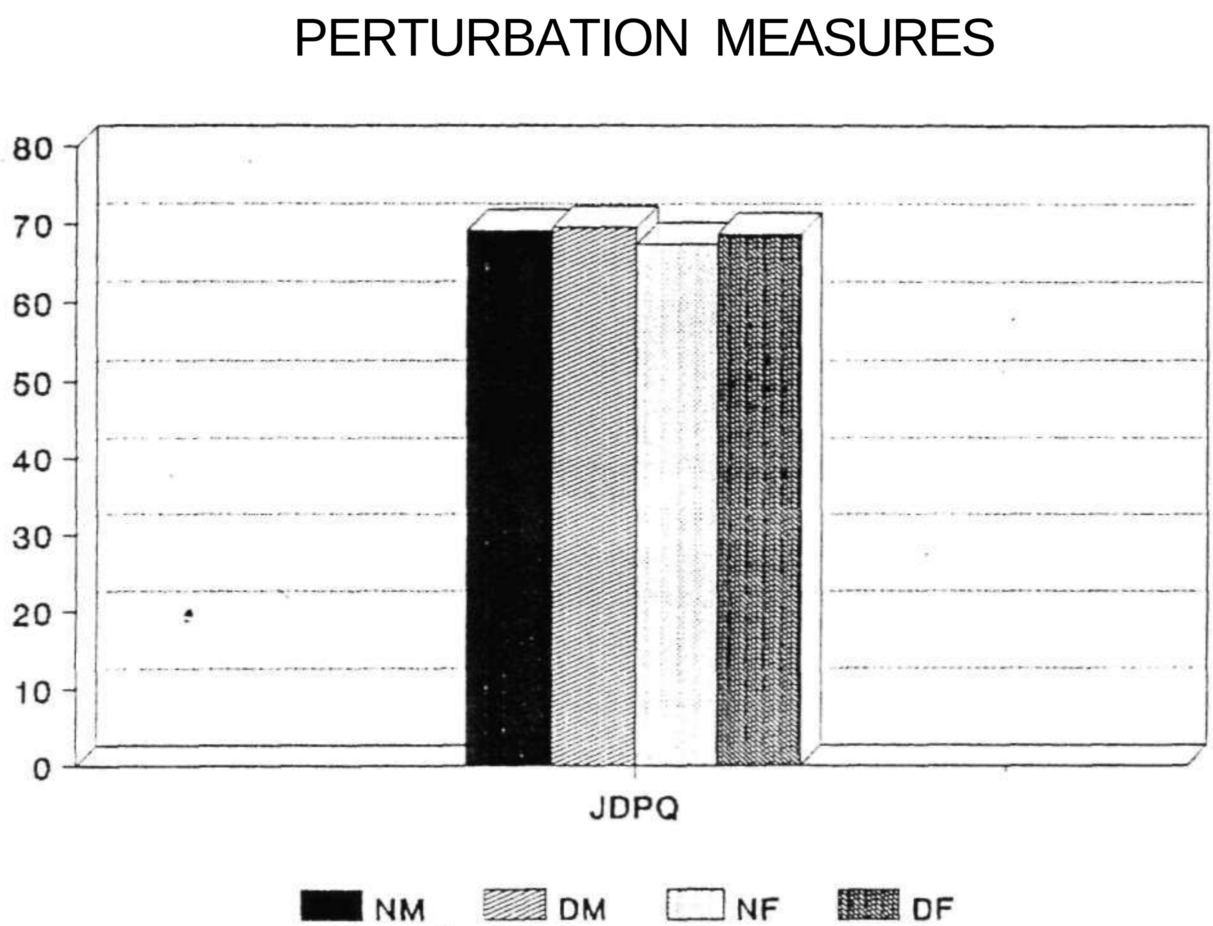
Graph 4: Showing Mean values of Percent Jitter (PJ) and Jitter Period Variability Index (JPVI) for Normal Males (NM), Dysphonic Males (DM), Normal Females (NF) and Dysphonic Females (DF) .



Graph 5: Showing Mean values of Jitter Directional Perturbation Quotient (JDPQ) for Normal Males (NM), Dysphonic Males (DM), Normal Females (NF) and Dysphonic Females (DF).



Graph 6: Showing Mean values of Jitter Relative Average Perturbation (JRAP) for Normal Hales (NM), Dysphonic Hales (DM), Normal Females (NF) and Dysphonic Females (DF).



Graph 5: Showing Mean values of Jitter Directional Perturbation Quotient (JDPQ) for Normal Males (NM), Dysphonic Males (DM), Normal Females (NF) and Dysphonic Females (DF).

SHIMMER MEASURES:

i) Shimmer in dB (SdB):

	Normals		Dysphonics	
	M	F	M	F
Mean	0.211	0.756	0.519	0.799
S.D	0.222	0.721	0.753	0.808
Range	1 .78	2.64	4.91	3.46

TABLE 19: Indicates the mean, S.D and Range for Shimmer in dB.

The mean values obtained for females in both normal and dysphonic groups were larger than in case of males of both the groups. Between normal and dysphonic groups, dysphonic males and females showed greater values than their normal counter parts . The comparison between males and females in normal group (t= -6.944), between males of normal and dysphonic groups (t= -2.423) and between females in normal and dysphonic groups (t= -0.061) significant differences were found. Between males and females of dysphonic group (t= 1.585) a statistically significant (at 0.1 level) difference was found.

Investigator	Normals M	Normals F	Dysphonics M	Dysphonics F
1) Sridhara (1986)	0.03	0.70	-	-
2) Chandra Shekar (1987)	0.03	0.70	1.78	0.87
3) Manjula (1987)	0.08	0.65	-	-
4) Balaji (1988)	0.03	0.70	2.32	1.23
5) Sanjay Kumar (1991)	3.68	2.08	7.30	4.75
6) Battacharya (1991)	1.93	2.16	-	-
7) Krishnan (1992)	0.28	0.25	-	-
8) Pathak (1997)	0.31	0.42	1.35	1.96
9) Rajkumar (1998)	0.23	0.79	3.38	3.52
10) Present study (1998)	0.21	0.76	0.52	0.80

TABLE 20: List3 of the mean scores found in other studies on shimmer in dB.

Therefore the hypothesis (1) stating that there was no significant difference between normal and dysphonic in terms of Shimmer in dB was rejected. Further, the hypothesis (2) stating that there is no significant difference between males and female was rejected.

The values obtained in this study was in agreement with the results of other studies. Shimmer in dB found to be a significant parameter in differentiating normals from dysphonics.

0.4142 respectively for the same vowels for males with hoarse voice. Rajkumar (1998) found mean values of 1.40 and 1.86 for males and females of normal group and 2.14 & 2.50 for males and females of dysphonic group. The findings of present study was in agreement with the results of Rajkumar's (1998) study. Since this parameter showed statistical significance in above mentioned three comparison. It was considered that it could be used to differentiate dysphonics from normals.

iii) Shimmer-Directional Perturbation Quotient (SDPQ)

	Normals		Dysphonic	
	M	F	M	F
Mean	47.36	40.68	50.75	45.83
S.D	8.92	12.41	11.11	9.90
Range	36.37	60.58	43.08	48.74

TABLE 22: indicates the mean, S.D and Range for SDPQ.

The mean values obtained for males in both normal and dysphonic groups were greater than the values obtained for females of respective groups. Comparison of males and females between both groups revealed that scores obtained for dysphonic group for both sexes were higher than their normal counterpart. Comparison between both sexes in normal group (t= 3.794), and of dysphonic group (t= - 2.069) and between normal and dysphonic group for males (t= - 1.903, significant at 0.064 level) and for females (t= 1.981) were found to have significant differences.

Therefore the hypothesis (1) stating that there was no significant difference between normal and dysphonic in terms of Shimmer Directional Perturbation Quotient was rejected. Further, the hypothesis (2) stating that there is no significant difference between males and female in terms normals and dysphonics was rejected.

Sorenson & Horii (1984) found 59.47%, 61.13%, & 58.19 for \a\,\i\ & \u\ of males and 63.13%, 61.71% & 59.76% for females respectively. Rajkumar (1998) reported 35.4 for males and 74.13 for females in normals and 48.72 for males and 53.02 for females is dysphonics. The present results correlated with findings of previous study. Since, SDPQ found significant on all comparison, this parameter was considered useful in differentiating dysphonics from normals.

iv) Shimmer - Amplitude Perturbation Quotient (SAPQ) :

	Normals		Dysphonics	
	M	F	M	F
Mean	1.16	3.50	2.46	3.66
S.D.	0.71	3.22	2.04	4.37
Range	4.21	16.15	8.23	24.50

TABLE 23: Shows the mean, S.D and Range for SAPQ.

The scores observed for females in both normal and dysphonic groups were slightly higher than the scores of males. The comparison of scores in both sexes between both groups indicated that dysphonics had scored slightly higher

ii) Shimmer-Amplitude Variability Index (SAVI)

	Normals		Dysphorics	
	M	F	M	F
Mean	0.76	1.76	1.13	1.73
S.D	0.48	1.01	0.61	0.74
Range	2.58	5.97	2.83	2.63

TABLE 21: Shows the mean, S.D and Range for SAVI.

The mean values obtained for females in both normals and dysphonics were higher than those values obtained for males. Significance of differences were found on comparison between males and females in both normal (t= -8.045) and dysphonic groups (t= 3.811) and between males of normal and dysphonic groups (t= - 3.457). Comparison between females of normal and dysphonic groups showed no statistical significance of difference (t= 0.512).

Therefore the hypothesis (1) stating that there was no significant difference between normal and dysphonic females in terms of Shimmer Amplitude Variability Index was accepted. Whereas it was rejected for males. Further, the hypothesis (2) stating that there is no significant difference between males and female with reference to normals and dysphonics was rejected.

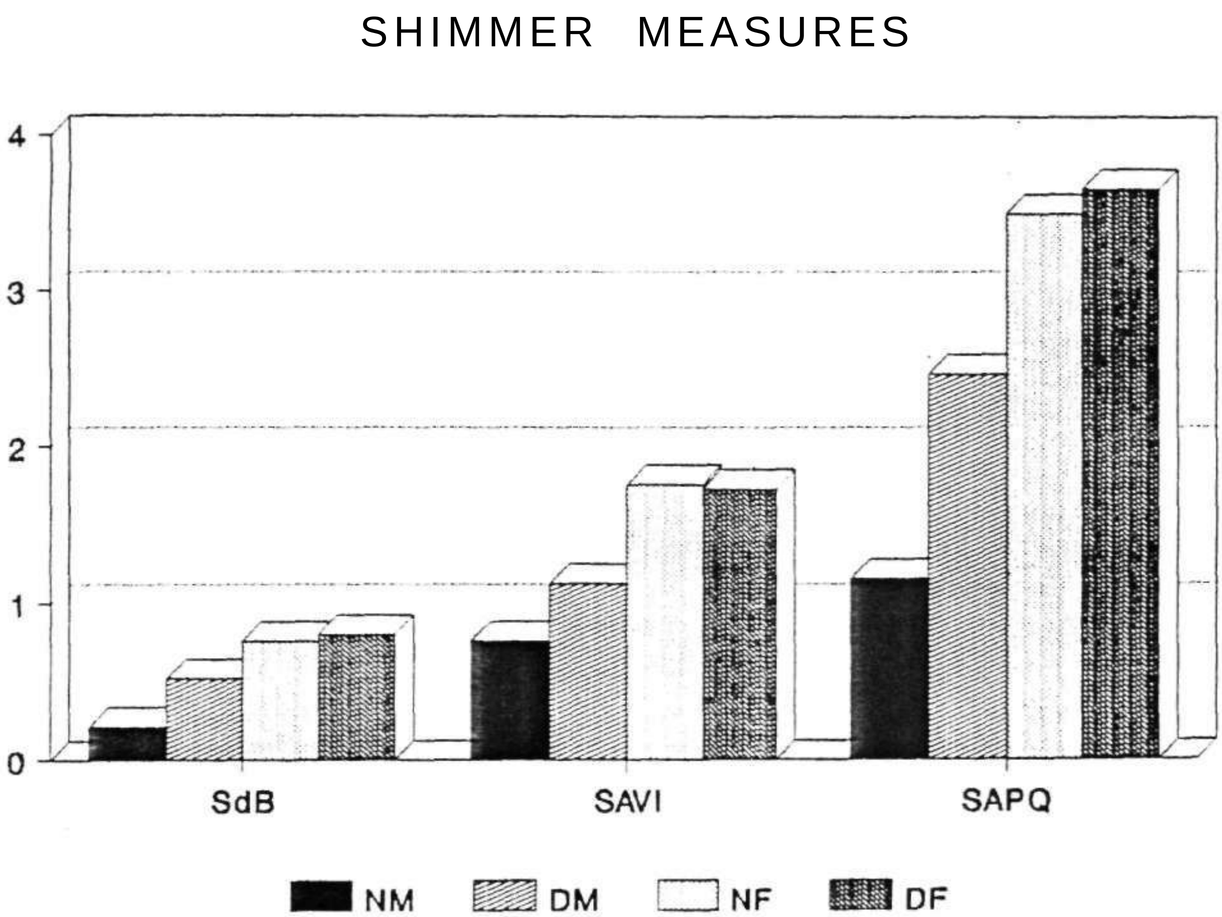
Deal and Emanuel (1978) reported mean amplitude variability index of 0.0619, -0.1330 and - 0.1287 for \a\,\i\, & \u\ respectively for males and 0.2163, 0.5706 and than normals. When compared males and females of normal group, (t= - 6.302) and of dysphonic group (t= 1.686) and males of normal and dysphonic groups (t= - 4.091), differencces were found to be statistically significant. No significant difference was found on comparison of females of normals and dysphonics (t= 0.499).

Therefore the hypothesis (1) stating that there was no significant difference between normal and dysphonic females in terms of Shimmer Amplitude Perturbation Quotient was accepted, whereas it was rejected for males. Further, the hypothesis (2) stating that there was no significant difference batween males and female in terms normals and dysphonics was rejected.

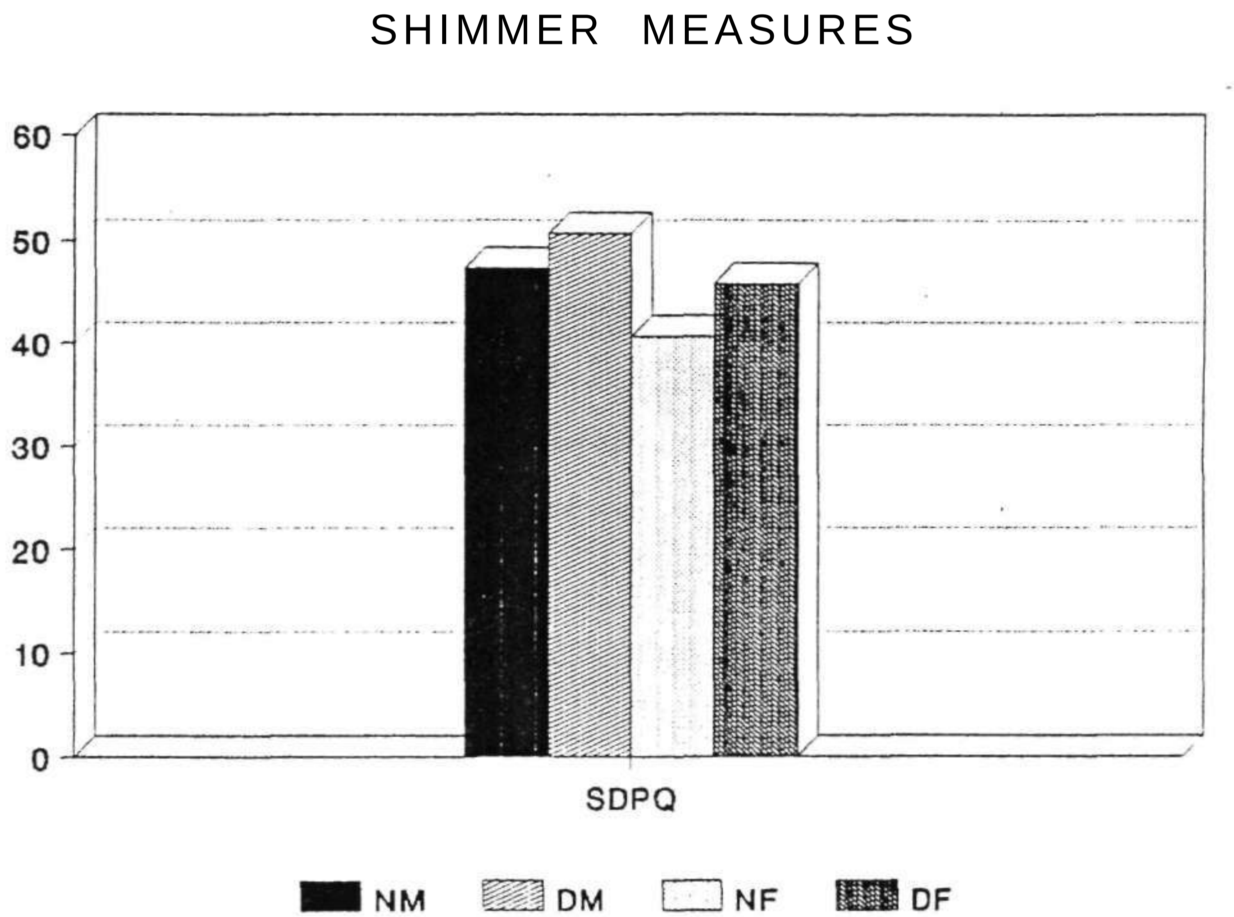
Investigator	Normals		Dysphonics	
	M	F	M	F
1) Krishnan (1992)	1.87	1 .80	-	-
2) Pathak (1997)	2.25	2.94	9.33	6.15
3) Rajkumar (1998)	1.82	7.03	7.77	7.09
4) Present study (1998)	1.16	3.50	2.46	3.66

TABLE 24: Shows the mean values of other studies on SAPQ.

From the study of above table it was clear that values found in the present 3tudy were similar to the reports of earlier studies. In the present study, SAPQ was found to be not an important parameter in differentiating between normals and dysphonics.



Graph 8: Showing Mean values of Shimmer in dB (SdB), Shimmer Amplitude Variability Index (SAVI), Shimmer Amplitude Perturbation Quotient (SAVI) for Males (NM), Dysphonic Males (DM), Normal Females (NF) and Dysphonic Females (DF).



Graph 9: Showing Mean values of Shimmer Directional Perturbation Quotient (SDPQ) for Normal Males (NM), Dysphonic Males (DM), Normal Females (NF) and Dysphonic Females (DF).

(C) FREQUENCY RELATED PARAMETERS

(i) Mean Fundamental Frequency (FFO):

The Mean Fo was compared for normal males, females and dysphonic males and females.

	Normals		Dysphonics	
	Males	Females	Males	Females
Mean	126.65	232.12	155.52	209.22
Standard Deviation	9.45	24.19	55.63	46.19
Range	44.71	96.26	249.60	217.76

TABLE 25: Reveals the mean, S.D and Range for Mean Fundamental Frequency.

This parameter differentiated the normal males and females from dysphonic males and females (t= 3.708; 2.524). There was also a significant difference between normal males and females (t= 34.063) and dyphonic males and females (t= 4.937).

Therefore the hypothesis (1) stating that there is no significant difference between normals and dysphonics in terms of Mean Fundamental Frequency was rejected. Further, the hypothesis (2) stating that there is no significant difference between males and female in terms normals and dysphonics was rejected.

The results of this study correlated with the findings of earlier studies.

INVESTIGATORS	MALE	FEMALE
1. Sheela (1974)	126	217
2. Jayaram (1975)	123	225
3. Nataraja and Jagadi3h (1984)	141	237
4. Vanaja (1986)	127	234
5. Nataraja (1986)	119	223
6. Sreedevi (1987)	119	218
7. Tharmar (1991)	124	233
8. Suresh (1991)	123	219
9. Sanjay (1991)	131	220
10. Rajashekhar (1991)	148	-
11. Krishnan (1992)	122	231
12. Pathak (1997)	126	231
13. Prabha (1997)	125	214
14. Pradeep (1997)	136	240
15. Rajkumar (1998)	140	240
16. Present study (1998)	127	232

TABLE 26 : The values of Mean Fundamental Frequency (in Hz.) for phonation on normal Indian population as reported by various investigators.

INVESTIGATORS	MALE	FEMALE
1. Jayaram (1975)	174	202
2. Nataraja (1986)	152	200
3. Sanjay (1991)	157	233
4. Pathak (1997)	141	234
5. Prabha (1997)	159	217
6. Rajkumar (1998)	172	229
7. Present study (1998)	156	209

TABLE 27: The values of Mean Fundamental Frequency (in Hz.) for phonation on 'dysphonic' Indian population as reported by various investigators.

The above results clearly indicated that this parameter was useful in differentiating normals and dysphonics.

(ii) Frequency Range (FRAN).

	NORMALS		DSYPHONICS	
	M	F	M	F
Mean	4.84	9.63	21.89	23.70
S.D.	2.01	2.94	39.68	36.38
Range	9.30	12.06	169.00	163.50

TABLE 28: Shows mean, S.D , Range for Frequency Range obtained in normals of dysphonics.

The above table clearly indicates a larger frequency range for females, in both normal and dysphonic groups compared to males. This could be owing to higher maximum frequency obtained in case of females.

The studies done so far in the Indian context have revealed a higher frequency range for females compared to males.

INVESTIGATORS	MALE	FEMALE
1. Nataraja (1986)	7.80	9.43
2. Sreedevi (1987)	20.42	24.08
3. Rajashekhar (1991)	11.20	—
4. Tharmar (1991)	6.20	8.59
5. Suresh (1991)	9.06	19.80
6. Prabha (1977)	4.78	18.21
7. Pradeep (1977)	10.18	18.88
8. Rajkumar (1998)	21.81	29.81
9. Present Study (1998)	4.84	9.63

TABLE 29 : The values of Fundamental Frequency Range (in Hz.) in phonation of /a/ by males and females as reported by previous investigators and the present study.

INVESTIGATORS	MALE	FEMALE
1. Nataraja (1986)	218.33	202.85
2. Prabha (1997)	49.33	53.20
3. Rajkumar (1998)	25.80	20.53
4. Present Study (1998)	21.89	23.70

TABLE 30: The values of Fundamental Frequency Range (in Hz.) in phonation shown by 'dysphonic' males and females as reported by previous investigators and the present study.

Shobha (1996) also found a good frequency range for female professional voice users (27.16 Hz). Sreedevi (1987) also attempted measures of fundamental frequency range in good voice and the values were 16.29 & 24.33 Hz for males and females.

In the present study, this measure has proved useful in differentiating the normals and dysphonics i.e., there were statistically significant differences between normal and dysphonic females (t= -2.295) and normal and dysphonic males (t= -2.897). There was also significant difference between normal males and females (t= -11.465) but not between dysphonic males and females (t= - 0.005). The results of this study were in agreement with Nataraja (1986) and Rajkumar (1998) who found frequency ranges to be an important parameter in differentiating between normals and dysphonics and also for differential diagnosis of different types of dysphonia.

Therefore the hypothesis (1) stating that there is no significant difference between normals and dysphonics in terms of Speed of Fluctuation in Frequency was rejected, whereas it was accepted for dysphonics. Further, the hypothesis (2) stating that there is no significant difference between males and female in terms normals and dysphonics was rejected.

(iii). SPEED OF FLUCTUATION IN FREQUENCY (SFF)

The values of normals of dysphonics are shown in the table below.

	NORMALS M	F	DYSPHONICS M	F
Mean	1.30	6.13	6.61	10.09
S.D.	2.431	3.53	10.48	9.12
Range	14.43	14.43	64.18	43.78

TABLE 31: Shows mean, S.D, Range for Speed of Fluctuation in frequency in normals and dysphonics.

The males of the normal group and dysphonic groups showed a low scores in trems of fluctuations compared to the females of the same group. Studies reported in the literature also showed a lower values for males than for females.

INVESTIGATORS	MALE	FEMALE
1. Nataraja (1986)	5.60	6.18
2. Tharmar (1991)	0.83	1.95
3. Rajashekhar (1991)	5.70	–
4. Suresh (1991)	2.54	4.49
5. Krishnan (1992)	8.73	15.42
6. Prabha (1997)	1.12	4.14
7. Pradeep (1997)	6.20	8.37
8. Rajkumar (1998)	14.86	22.01
9. Present Study (1998)	1.30	6.13

TABLE 32 : The values of Speed of Fluctuation in Frequency (in fluc./sec) for sustained phonation of /a/ shown by normal males and females as reported by various investigators.

INVESTIGATORS	MALE	FEMALE
1. Nataraja (1986)	47.59	48.31
2. Prabha (1997)	5.86	9.31
3. Present Study (1998)	17.33	19.24

TABLE 33 : The values of Speed of Fluctuation in Frequency (in fluc./sec) for sustained phonation of /a/ shown by normal males and females of the 'dysphonic' population as reported by various investigators.

On analysis it was seen that there was a statistically significant difference between the values obtained by males and females of two groups i.e., normal males and females (t= -10.393); and dysphonic males & females (t= 2.229).

This parameter had also proved useful in differentiating dysphonics and normals of both the sexes. A significant difference was found between normal and dysphonic males (t= -3.361) and normal and dysphonic females (-2.295).

Therefore the hypothesis (1) stating that there is no significant difference between normals and dysphonics in terms of Speed of Fluctuation in Frequency was rejected. Further, the hypothesis (2) stating that there is no significant difference between males and female in terms normals and dysphonics was rejected.

These results correlated with the findings of Rajkumar (1996) who also found a statistically significant difference between males and females of normal and dysphonic groups Nataraja (1986) had concluded that SFF was an important parameter in differentiating normals from dysphonics. Shobha (1996) had reported Speed of Fluctuation in Frequency in good voices (in Indian professional voice users) and got similar results.

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

(iv) Extent of fluctuations in frequency (EFF)

For this parameter, the following values were obtained by normals and dysphonics.

	Normals		Dysphonics	
	M	F	M	F
Mean	1.58	4.87	8.94	8.87
S.D.	1.68	5.48	17.35	16.33
Range	4.12	42.98	82.02	79.69

TABLE 34: Shows mean, S.D and Range of Extent of fluctuations in frequency for normals of dysphonics.

The normal males obtained lesser values compared to their female counterparts. But the extent of fluctuations in the dysphonics were relatively equal between the sexes.

There was a significant difference between the values of normal males and females (t= 4.770) and between dysphonic males and normal males (t= -2.827) but not between dysphonic males and females (t= 0.021) and between dysphonic females and normal females (t= -1.738).

Therefore the hypothesis (1) stating that there is no significant difference between normal and dysphonic males in terms of Extent of Fluctuation in Frequency was rejected, whereas it was accepted for females. Further, the hypothesis (2) stating that there is no significant difference between males and female with reference to dysphonics was accepted, whereas it was rejected for normals.

In most of the studies in the literature the males have obtained a lesser value compared to females.

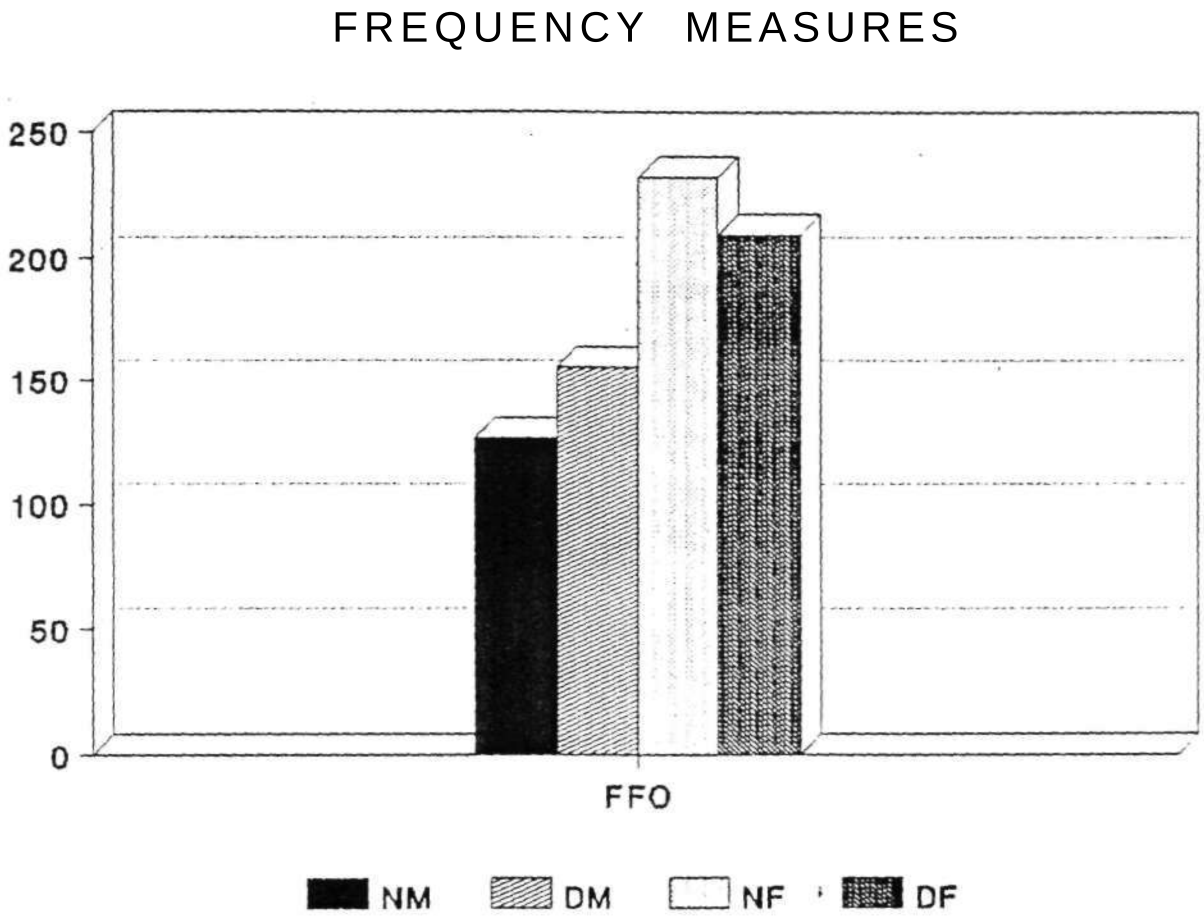
INVESTIGATORS	MALE	FEMALE
1. Vanaja (1986)	1.38	1 .29
2. Nataraja (1986)	3.87	3.56
3. Rajashekhar (1991)	3.0	–
4. Tharmar (1991)	2.75	3.59
5. Suresh (1991)	3.44	4.12
6. Krishnan (1992)	19.13	8.55
7. Prabha (1997)	1.94	2.36
8. Pradeep (1997)	2.95	3.41
9. Rajkumar (1998)	3.89	4.64
9. Present Study (1998)	1.58	4.87

TABLE 35: The values of Extent of Fluctuation in Frequency (in Hz.) for phonation of vowel /a/, in normal male and female as reported by various investigators.

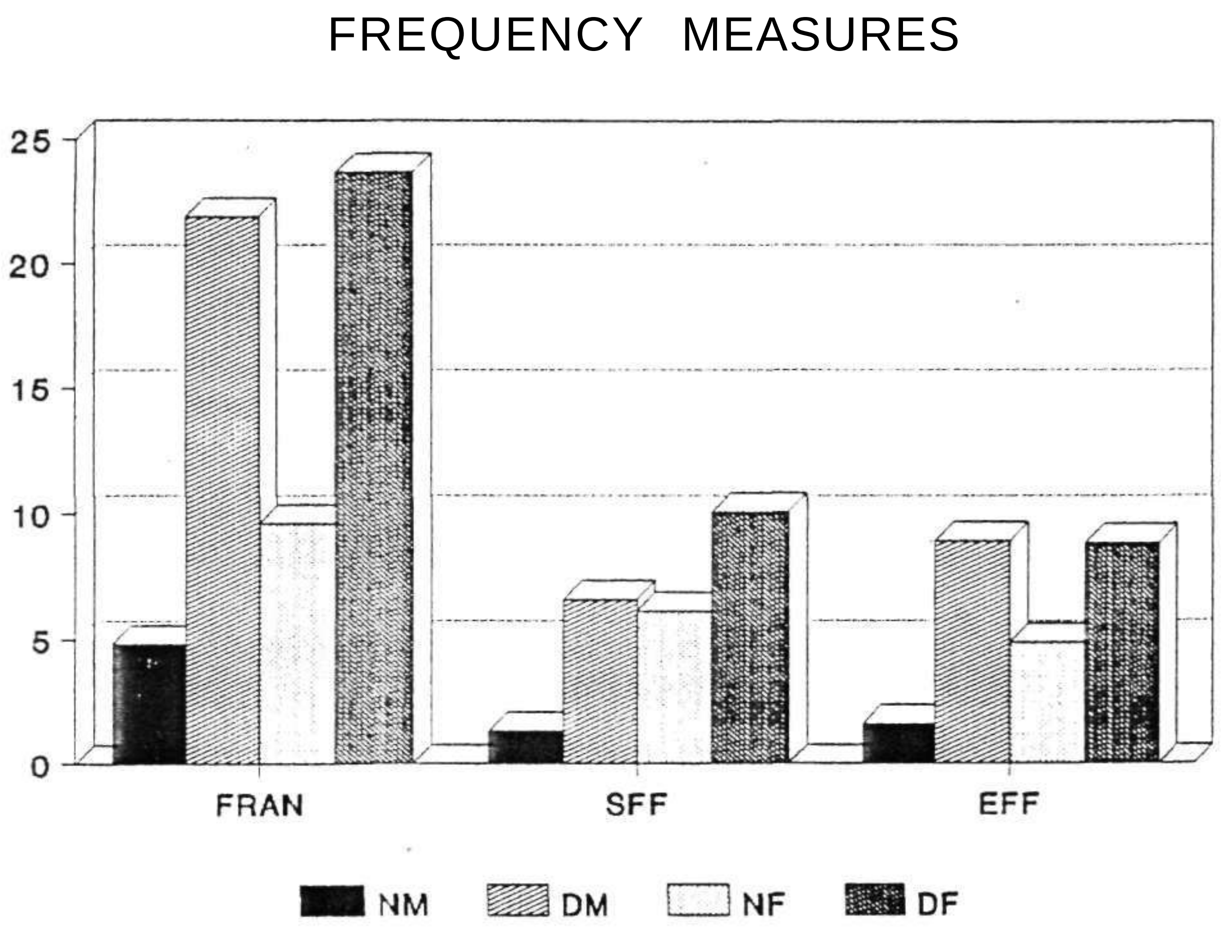
INVESTIGATORS	MALE	FEMALE
1. Nataraja (1986)	28.90	24.79
2. Prabha (1997)	3.64	2.37
3. Present Study (1998)	5.04	4.92

TABLE 36: The values of Extent of Fluctuation in Frequency (in Hz.) for phonation of vowel /a/, in the 'dysphonic' male and female as shown by various investigators.

Rajkumar (1996) found a statistically significant difference between normal and dysphonic groups which correlated with the results of this study. Since this measure indicated the ability of the subject to control voice and he found that in the good group, the extent of fluctuation was lesser than in the other two groups, Shobha (1996) found EFF values in professional voices as 3.97 for females. However this parameter was considered not useful in differentiating the normals from dysphopnics of both the sexes.



Graph 10: Showing Mean values of Mean Fundamental Frequency (FFo) for Normal Males (NM), Dysphonic Males (DM), Normal Females (NF) and Dysphonic Females (DF).



Graph 11: Showing Mean values of Frequency Range (FRAN), Speed of Fluctuations in Frequency (SFF) and Extend of Fluctuations in Frequency (EFF) for Normal Males (NM), Dysphonic Males (DM), Normal Females (NF) and Dysphonic Females (DF).

(D) INTENSITY RELATED PARAMETERS:-

(i) Intensity - Mean Amplitude (IMAO): -

	NORMALS M	F	DYSPHONICS M	F
Mean	51.84	52.80	51.83	55.97
S.D.	3.68	4.14	4.84	4.63
Range	17.69	18.39	20.25	21.40

TABLE 37: Shows mean, S.D and Range of Mean Amplitude obtained for normals and dysphonics.

From the analysis it was clear that this parameter was not useful to differentiate neither normal and dysphonics nor between the sexes, because there was no statistically significant difference between groups on comparisons made; normal males and females (t= -1.697); dysphonic males and females (t= 0.319); normal males and dysphonic males (t= -0.248); normal females and dysphonic females (t= - 4.356).

Therefore the hypothesis (1) stating that there is no significant difference between normals and dysphonics in terms of Mean Amplitude was rejected. Further, the hypothesis (2) stating that there is no significant difference between males and female in terms normals and dysphonics was rejected.

The mean intensity reported by differed authors are presented in table - 38.

Investigators	Male	female
Prabha (1997)	49.73	48.44
Pradeep (1997)	45.72	50.65
Rajkumar (1998)	54.62	47.28
Present study (1998)	51.84	52.80

TABLE 38: Shows the Mean Amplitude obtained by various Indian investigation.

The values obtained in this study were in agreement with the values obtained by the reports of studies quoted above.

It can be concluded that mean intensity was not proved useful in differentiating normals from dysphonics in the present study.

(ii) Extent of fluctuation in Intensity (EFI);

	Normals		Dysphonics	
	M	F	M	F
Mean	0.13	0.28	0.95	0.22
S.D	0.63	0.89	1.62	0.83
Range	3.38	3.06	4.10	3.45

TABLE 39: Shows mean, S.D and Range of EFI obtained for normals and dysphonics.

The males of the dysphonic group had a greater value of Extent fo Fluctuation in Intensity compared to the females, but in contrast the females of the normal group obtained a lower Extent of Fluctuation in Intensity compared to their male counterparts. The males in the dysphonic group obtained the highest Extent of Fluctuation in Intensity values compared to all others.

Kim, Kakita and Hirano (1982) found the extent of fluctuation in intensity for \a\ in adult males and females. The mean value was 4.2 dB, S.D was 1.1 and the range was 3-6 for males and 5.8,0.8 (S.D) and Range (5 to 7dB) for females. A similar study by Yoon, Kakita and Hirano (1984) gave the range as 1.6 to 4dB. Some of the studies which had obtained higher Extent of Fluctuation in Intensity values were.

Investigators	Male	female
Nataraja (1986)	2.45	1.59
Rajashekhar (1991)	1.80	-
Suresh (1991)	2.39	1.32
Krshnan (1992)	4.81	5.58
Pradeep (1997)	2.13	1.71
Rajkumar (1998)	1.16	1.21

TABLE 40: Shows the values of Extent of Fluctuation in Intensity obtained by various Indian investigators.

The results of this study approximated that of Tharmar (1991) who found Extent of Fluctuation in Intensity values for males and females as 0.26 and 0.46 and of Prabha (1997) as 0.42 and 0.54. Shobha (1986) also found the Extent of Fluctuation in Intensity in professional voice users as 0.66.

Extent of Fluctuation in Intensity in the present study was found to differentiate only normal males from dysphonic males (t= -3.168) but not the normal females from dysphonic females (t= -1.738). There was a significant difference between dysphonic males and females (t= -4.784), but not between normal males and females (t=-4.770).

Therefore the hypothesis (1) stating that there is no significant difference between normal and dysphonic females in terms of Extent of Fluctuation in Intensity was accepted, whereas it was rejected for males. Further, the hypothesis (2) stating that there is no significant difference with reference to normals was accepted whereas it was rejected for dysphonics.

Nataraja (1986) and Rajkumar (1998) had opined that Extent of Fluctuation in Intensity was a valued parameter for deferential diagnosis of dysphonics and for differentiating between normal and dysphonic cases. Though this parameter was considered as reflecting the ability of the individual to control his voice production system, it was clear from this study that dysphonic males had a poor control over their vocal mechanism owing to a pathology, but the females did not show such a difficulty and hence Extent of Fluctuation in Intensity was considered as a not useful parameter in differentiating normals from dysphonics.

(iii) Speed of fluctuation in Intensity (SFI)

-	Normals		Dysphonics	
	M	F	M	F
Mean	0.027	0.053	0.586	0.044
S.D.	0.140	0.155	1.73	0.179
Range	1.00	0.50	7.96	1 .00

TABLE 41: Shows mean, S.D and Range of SFI obtained in normals and dysphonics.

The males of the normal group had shown a lower values compared to the males of the dysphonic group. This could be attributed to the limited ability on the control of the vocal mechanism in case of dysphonics.

On analysis it was found that Speed of Fluctuation in Intensity was an important parameter which could differentiate normals and dysphonics of both the sexes i.e., there was a significant difference across all the comparisons made i.e., between normal males and females (t= -0.830) ; dysphonic males and females (t= -2.136) normal males and dysphonic males ((t= -2.174) and normal females and dysphonic females (t= 0.000).

Therefore the hypothesis (1) stating that there is no significant difference between normals and dysphonics in terms of Speed of Fluctuation in Intensity was rejected. Further, the hypothesis (2) stating that there is no significant difference between males and female in terms normals and dysphonics was rejected.

The values of Speed of Fluctuation in Intensity obtained in the present study were in agreement with the following reports;

INVESTIGATORS	MALE	FEMALE
1. Nataraja (1986)	1.40	1.00
2. Rajashekhar (1991)	0.40	–
3. Tharmar (1991)	1.06	2.44
4. Suresh (1991)	1.43	0.45
5. Krishnan (1992)	2.23	0.30
6. Prabha (1997)	0.22	0.88
7. Pradeep (1997)	2.43	1.20
8. Rajkumar (1998)	0.80	1.62
9. Present study (1968)	0.03	0.05

TABLE 42: The values of Speed of Fluctuation in Intensity (in Fluc./sec) for the phonation of /a/ in normal males and females as shown by various investigators.

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

iv) Intensity Range (IRAN);

	Normals M	F	Dysphonics M	F
Mean	2.13	3.30	3.30	2.03
S.D	0.90	9.84	1.84	1.17
Range	6.79	86.09	9.65	6.13

TABLE 43: Shows mean, S.D and Range of mean Intensity range for normals and dysphonics.

From the above table it is clear that the females of the normal group had obtained the same value as the males of the dysphonic group.

There was statistically significant difference in the values of intensity range obtained by normal males compared with dysphonic males (t= -4.214) but not between normal females and dysphonic females (t= 0.116). There was significant difference between males and females of both normal and dysphonic groups.

Therefore the hypothesis (1) stating that there is no significant difference between normal and dysphonic females in terms of Intensity Range was accepted, whereas it was rejected for males. Further, the hypothesis (2) stating that there is no significant difference between males and female in terms normals and dysphonics was rejected.

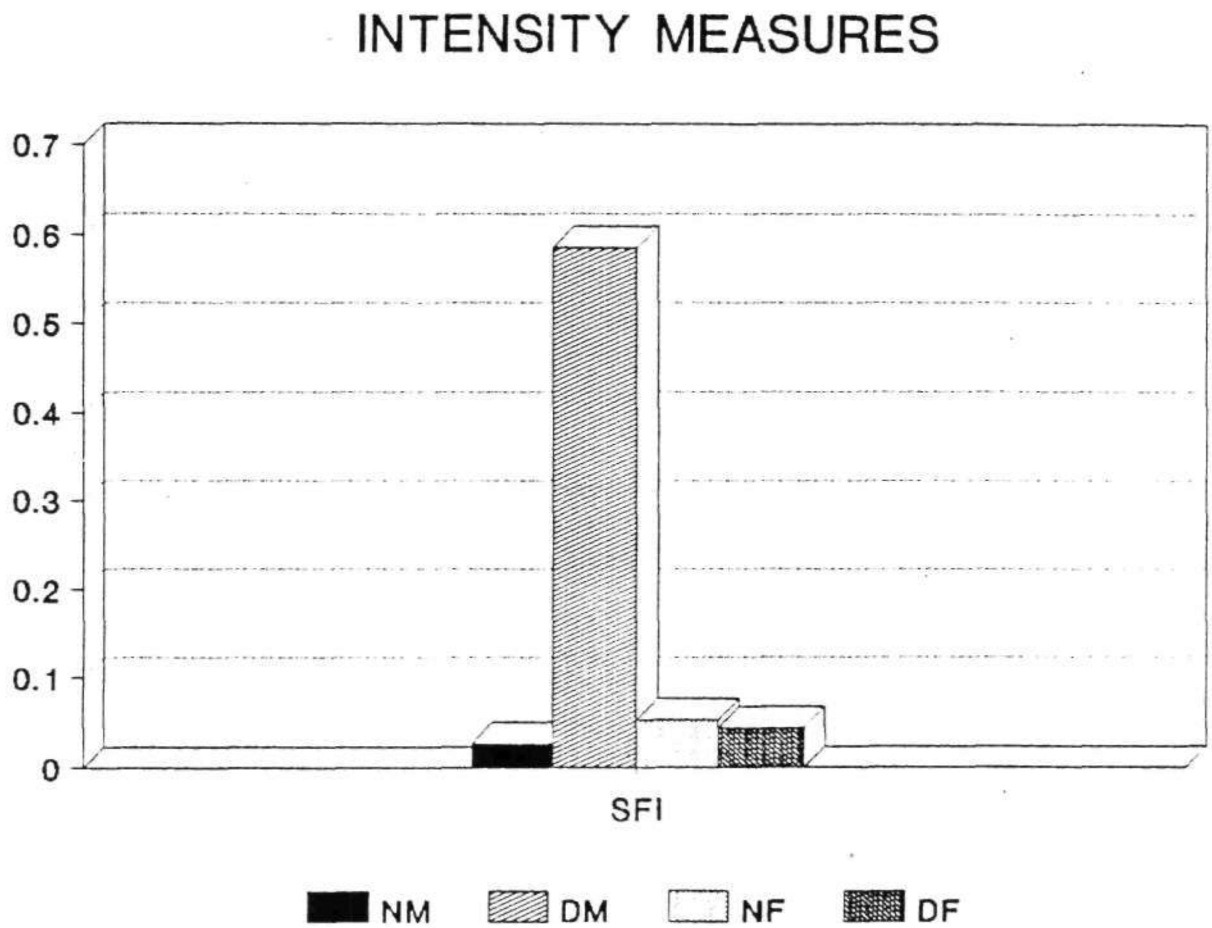
Earlier studies had also shown values similar to that in the present study.

INVESTIGATORS	MALE	FEMALE
1. Nataraja (1986)	3.80	4.18
2. Sreedevi (1987)	5.06	4.04
3. Rajashekhar (1991)	5.20	–
4. Tharmar (1991)	4.56	4.73
5. Suresh (1991)	5.36	3.80
6. Prabha (1997)	2.29	3.00
7. Pradeep (1997)	4.56	4.50
8. Rajkumar (1998)	2.57	2.45

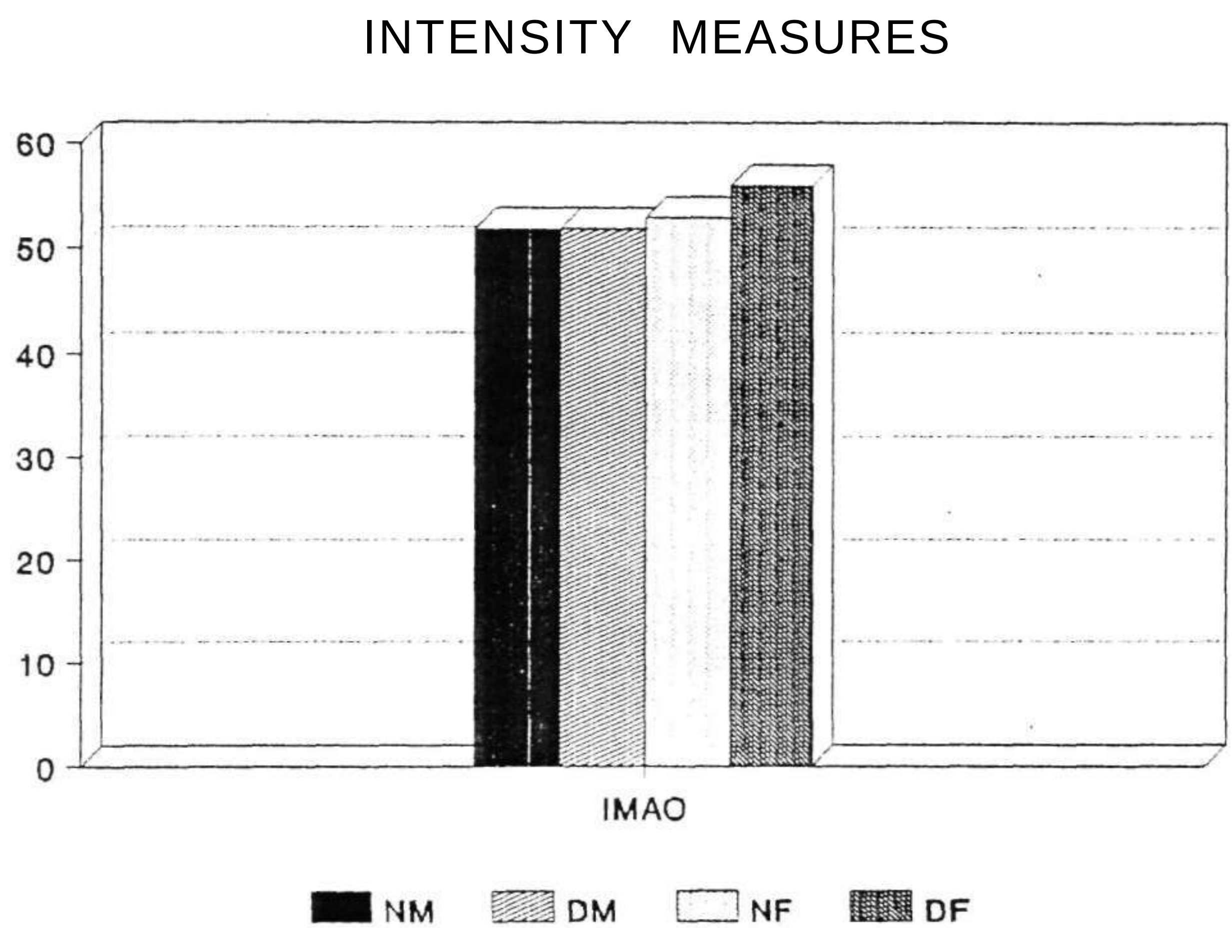
TABLE 44: The values of Intensity Range (in dB.) for phonation of /a/ in normal males and females as reported by various investigators.

In the dysphonic males and females Nataraja (1986) had measured the intensity range in phonation as 8.74 and 9.46 dB respectively.

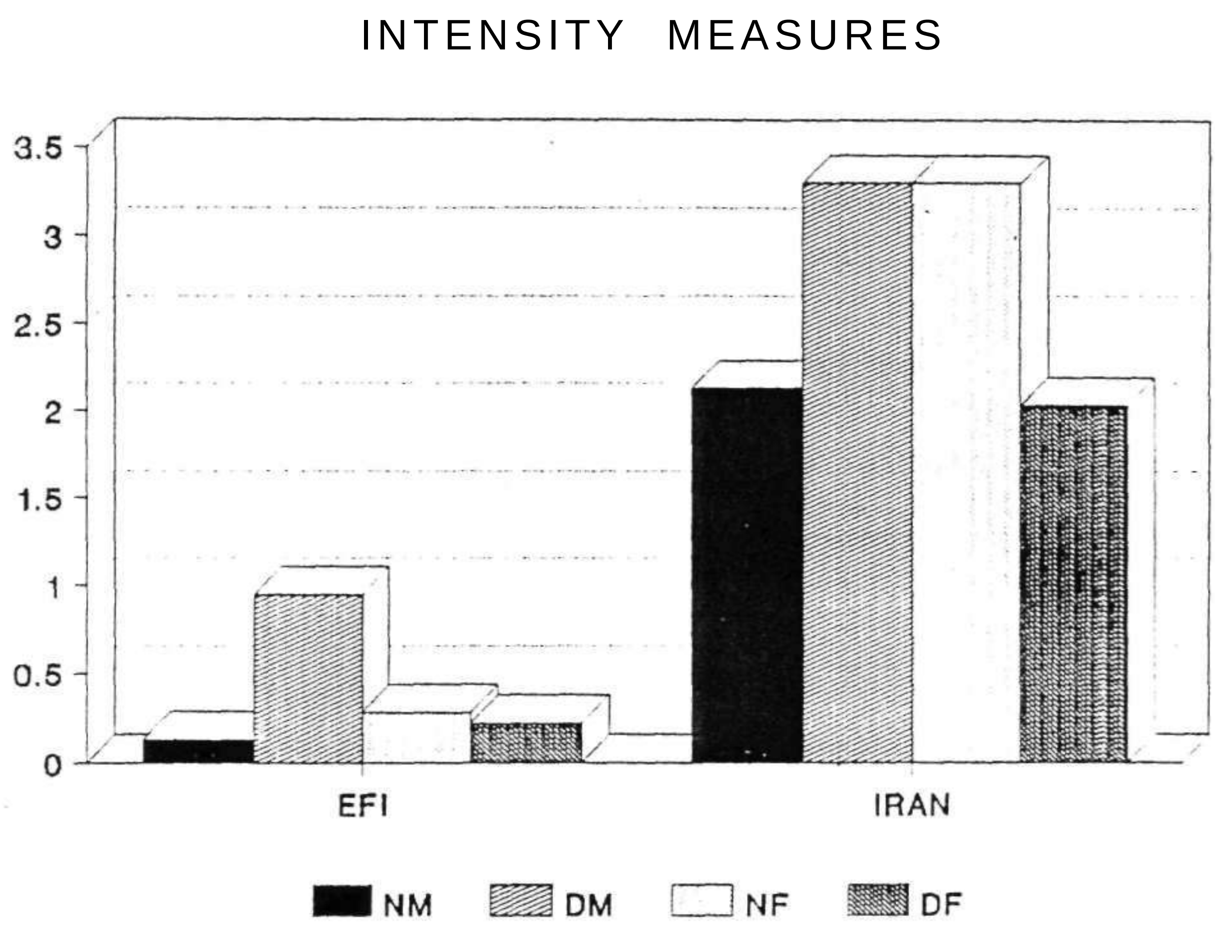
So it was concluded that intensity range was not a useful measure to differentiate normals and dysphonics of both the sexes.



Graph 12: Showing Mean values of Speed of Fluctuations in Intensity (SFI) for Normal Males (NM), Dysphonic Males (DM), Normal Females (NF) and Dysphonic Females (DF).



Graph 13: Showing Mean values of Intensity—Mean Amplitude (IMAO) for Normal Males (NM), Dy3phonic Males (DM), Normal Females (NF) and Dysphonic Females (DF).



Graph 14: Showing Mean values of Extend of Fluctuations in Intensity (EFI) and Intensity Range (IRAN) for Normal Males (NM), Dysphonic Males (DM), Normal Females (NF) and Dysphonic Females (DF).

I) COMPARISON OF SCORES BETWEEN MALES AND FEMALES WITHIN NORMALS AND DYSPHONICS.

The comparison of scores of both sexes within normals and dysphonic groups was made. The parameters common to both groups (eleven only) where there was a significant difference between the scores of males and females were (Table 45):

A) Spectral Measures:-

1) Harmonics to noise ratio

(B) Perturbation Measures:-

2) Jitter Mean Fundamental Frequency
3) Percent Jitter
4) Jitter Period Variability Index
5) Shimmer in dB
6) Shimmer Amplitude Variability Index
7) Shimmer Directional Perturbation Quotient
8) Shimmer Amplitude Perturbation Quotient

(C) Frequency Measures;-

9) Fundamental Frequency
10) Speed of Fluctuations in Frequency

(D) Intensity Measures:-

11) Mean Intensity

On comparing the scores of males and females in both normals and dysphonic groups there was a difference only in the above mentioned parameters. The parameters which had not shown significant differences were as follows:

Parametes	Between Nm & Nf	Between Dm & Df	Between Nm & Dm	Between Nf & Df
l.AR	X	S5	X	X
2. BR	X	X	X	X
3.GR	S	X	S5	S1
4.HNR	S	S5	S	S
5.H1A	S	X	S	S
6.N0H	X	X	S	X
7.JF0	S	S	S	S
8.PJ	S	S	S1	X
9.JPVI	S	S	S	S
10.JRAP	X	X	X	X
11.JDLT	S	X	X	X
12.JDPQ	S5	Y	X	X
13.SdB	S	S1	S	X
14.SAVI	S	S	S	X
15.SDPQ	S	S5	S1	S5
16.SAPQ	S	SI	S	X
17.FFO	S	S	S	S
18.FRAN	S	X	S	S
19.EFF	S	X	S	SI
20.SFF	c	S5	S	S5
21.IMAO	SI	S	X	X
22.EFI	X	S1	S	X
23.SFI	X	S5	S5	X
24.IRAN	X	X	S	X

TABLE 45: Shows the comparison within and between the groups of both sexes for all parameters.

(S - significant at all levels, S5 - significant at 0.05 level, SI - significant at 0.1 level, X - not significant, Nm - Normal males, Nf - normal females , Dm - Dysphonic males, Df - Dysphonic females).

RESULTS OF DISCRIMINATE ANALYSIS

I) Comparison of Normals and dysphonics group:

To check the sensitivity of each parameter in differentiating the normals from dysphonics, the canonical discriminant analysis was done.

The following parameters were listed below in the order of canonical discriminant function

(a) Harmonics to Noise Ratio
(b) Speed of Fluctuations in Frequency
(c) First Harmonic Amplitude
(d) Extent of Fluctuations in Frequency
(e) Frequency Fange
(f) Speed of Fluctuation in Intensity
(g) Beta Ratio
(h) Shimmer in dB
(i) Shimmer Amplitude Variability Index
(j) Intensity Range
(k) Mean Intensity
(l) Alpha Ratio
(m) Shimmer Amplitude Perturbation Quotient
(n) Jitter Period Variability Index

The criterion for discrimination was taken as 0.05 level. It was clearly evident that the above listed parameters were most sensitive to differentiate normals and dysphonics.

Classification Results:

Actual group	No. of Samples	Predicted group 1	Membership 2
Group 1	150	148	2
(Normals)	100%	98.7%	1.3%
Group 2	90	21	69
(Dysphonics)	100%	23.3%	76.7%

90.4% of originaly grouped cases were correctly classified.

TABLE 46: Shows results of discriminant analysis for classifying the groups (group 1 - normals; group 2 -dysphonics).

From the above table it was found that 98.7% was correctly classified into normal groups and 76.7% was correctly classified as dysphonic groups. In dysphonic group, 23.3% was incorrectly classified into normal groups. i.e. two samples of normals voice (1.3%) were considered as belonging to dysphonic group and similarly 21 samples out of 90 (23.3%) dysphonic voice were considered as belonging to normal group by the Descriminent Analysis. The parameters used had not differentiated mild cases of hoarseness i.e the cases had provided scores similar to normals. This may happen as mild cases of dysphonics may use voices inconsistently and might have used normal or near normal voice dring testing. Further it may be that the parameters used in the present study are not sensitive enough to identify the mild hoarse voice cases. Eventhough several other studies (Yanagihara 1967) have claimed that these parameters were useful in differentiating cases of hoarseness, the present study does not support such a view.

II) COMPARISON OF PERCEPTUAL AND OBJECTIVE EVALUATION.

Perceptual evaluation of all the voice (normal and dysphonic) samples rated by 7 judges on a 4 point scale (1 - normal; 2 -mild; 3 - Moderate; 4 - severe hoarse voice) was considered. The ratings of 4 judges and above was considered as the score for that particular voice sample.

In order to find out the parameters which could differentiate the different degrees of hoarseness, the perceptual ratings were compared with objective values i.e., scores of different parameters. For this purpose the canonical discriminant analysis was employed using the SPSS software program.

The following parameters were found to be useful in differentiating various degrees of hoarse voice. They were listed below in the order of Canonical Discriminant function:

1. Speed of Fluctuations in Frequency
2. Mean Fundamental Frequency in both Jitter and Frequency measures
3. Extend of Fluctuations in Frequency
4. Jitter Directional Perturbation Quotient
5. Shimmer Directional Perturbation Quotient
6. Frequency Range
7. Percent Jitter

(a) Spectual Measures:-

1) Alpha ratio
2) Beta ratio
3) Gamma ratio.
4) First harmonic amplitude
5) Number of harmonics

b) Perturbation Measures:-

6) Jitter relative average perturbation
7) Jitter deviation from linear trend
8) Jitter = directional perturbation Quotient

c) Frequency Measures:

9) Frequency range
10) Extent of fluctuations in frequency

d) Intensity measures:-

11) Extent of fluctuations in Intensity
12) speed of fluctuations in intensity
13) Intensity range

II. Comparison of scores in males and females between normals and dysphonics.

The parameters that differentiated both the sexes between two groups were ten only. (Table 45)

a) Spectral measures:-

1) Gamma Ratio
2) First Harmonic Amplitude
3) Hormonics to Noise Ratio

III) COMPARISON OF THE SCORES OF BOTH SEXES WITHIN AND BETWEEN BOTH GROUPS i.e., NORMALS AND DYSPHONICS:

Only six parameters listed below showed a significant statistical difference across all the 4 types of comparisons made in this study. (Table 45)

a) Spectral Measures

1) Harmonics to noise ratio

b) Perturbation Measures:-

2) Jitter fundamental frequency
3) Jitter period variability Index.
4) Shimmer directional perturbation Quotient

c) Frequency Measures:-

5) Fundamental frequency
6) Speed of fluctuations in frequency

b) Perturbation Measures:-

4) Jitter Mean Fundamental Frequency

5) Jitter Period Variability Index

6) Shimmer Directional Perturbation Quotient

c) Frequency Parameters:-

7) Fundamental Frequency

8) Frequency Range

9) Extent of Fluctuations in Frequency

10) Speed of Fluctuations in Frequency

The hypothesis that there was no significant difference between normal and dysphonic groups in terms of parameters studied, was rejected based on the above findings.

Based on the above mentioned findings, the second hypothesis earlier formulated stating that there was no significant difference among the four groups based on degree hoarse voice in terms of parameters studied was rejected.

Classification results:-

Actual group	No. of samples	Predicted group Membership 1	2	3
Group 1 (Mild)	63	56 88.9	2 3.2	5 7.9
Group 2 (Moderate)	12	5 41.7	6 50.0	1 8.3
Group 3 (Severe)	15	1 6.7	4 26.7	10 66.7

* 80.00% of cases were correctly classified.

TABLE 47: Results of discriminant analysis for degrees of hoarseness, (group 1 - mild; group 2- moderate; group 3 -severe).

From classification results it was clear that:

1) In the mild group, out of 63 mild dysphonic samples, only 56 samples (88.9%) were correctly identified. 2 samples (3.2%) and 5 samples (7.9%) were grouped into moderate and severe degrees of hoarse voice respectively.

2) In the moderate group, out of 12 samples, 6 samples (50%) were correctly identified. 5 samples (41.7%) and 1

sample (8.3%) were grouped into mild and moderate degrees of hoarse voice respectively.

3) In the severe group, out of 15 samples, 10 samples (66.7%) were correctly identified. 4 samples (26.7%) and 1 sample (6.7%) were grouped into moderate and mild degrees of hoarse voice respectively.

From these results, it was inferred that there was a thin margin between normals and mild degrees if hoarse voice, because many of the mild degree of hoarse voices were grouped as belonging to normal group.

Futher it was also noticed that some of the normals were classified/identified as mild hoarse voice cases. This suggests that the quality of voice, normal to sever hoarseness occures on the same continuum and the boundaries between the normal and mild, mild and moderate and moderate to severe or not very clear, particularly perceptually. Therefore over lap cross this boundaries has to be excepted/accepted. However the classification of normals and different degrees of hoarseness has been possible.

Thus the study has achieved the objective of classifying the normal and different degrees of hoarseness based on 7 parameters, namely;

1. Speed of Fluctuations in Frequency
2. Mean Fundamental Frequency in both Jitter and Frequency measures

3. Extend of Fluctuations in Frequency
4. Jitter Directional Perturbation Quotient
5. Shimmer Directional Perturbation Quotient
6. Frequency Range
7. Percent Jitter

The measurement of these parameters would help clinically to evaluate the hoarseness of voice and in synthesis of hoarseness to obtain better quality of voice.

J

SUMMARY AMD CONCLUSION

Voice is considered as multidimensional series of measurable events. Many have suggested various means of analysing voice to note the factors that are responsible for creating an impression of a particular voice to determine the underlying mechanism (Micheal and Wendahl 1971; Jayaram, 1975; Hanson and Laver, 1981; Hirano, 1981).

Objective of the present study was to find out the parameters which could:

1. Differentiate hoarse voice from normal voice
2. Differentiate the hoarse voice based on severity
3. Determine parameters which lead to the perception of different degrees of hoarse voice.

For this purpose, the following parameters:

A) Spectral measures

1) Harmonics to Noise Ratio
2) First Harmonic Amplitude
3) Number of Harmonics
4) Alpha, Beta, Gamma ratios of LTAS

B) Perturbation Measures

i) Jitter Measures:
   1) Mean Jitter
   2) Percent Jitter
   3) Period Variability Index

4) Relative Average Perturbation

5) Directional Perturbation Quotient

6) Deviation from Linear Trend

ii) Shimmer Measures:

1) dB Shimmer

2) Amplitude Variability Index

3) Amplitude Perturbation Quotient

4) Amplitude Directional Perturbation Quotient

C) Frequency Measures

1) Mean Fundamental Frequency

2) Range of Frequency

3) Extend of Fluctuation in Frequency

4) Speed of Fluctuation in Frequency

D) Intensity measures

1) Mean Intensity

2) Range of Intensity

3) Extend of Fluctuation in Intensity

4) Speed of Flutuation in Intensity

were studied. All the 24 parameters were measured in a group of 80 subjects of 50 normals (25 each sex) and 30 dysphonics (15 each sex). Three trials of /a/, /i/ and /u/ vowels were recorded using a tape deck for the extraction of both the spectral frequency and intensity measures as well as for perceptual evaluation. The electroglottograph outputs were simultaneously recorded using laryngograph for the extraction of perturbation measures. Both the signal from electroglottograph and audio signals were digitized at 16 KHz sampling frequency using 12 bit analog to digital converter for further analysis using appropriate programme from vaghmi (VSS Bangalore)

For perceptual evaluation, the audio recording of speech was used. Seven judges were asked to rate the overall severity of hoarse voice on 4 point scale for each sample presented to them in a random order. Inter and Intrajudgements have been found to be reliable.

The results were subjected to Canonical Discriminant analysis using the SPSS computer program.

The following results were obtained:

a. It was found that following fifteen parameters were sensitive to differentiate between normals and dysphonics.

   1. Harmonics to noise ratio
   2. Speed of fluctuations in frequency
   3. First harmonic amplitude
   4. Extent of fluctuation in frequency
   5. Frequency range
   6. Speed of fluctuation in intensity
   7. Beta ratio
   8. dB Shimmer
   9. Shimmer Amplitude
   10. Shimmer Amplitude variability Index
   11. Intensity Range

12. Mean Intensity
13. Alpha Ratio
14. Shimmer Amplitude Perturbation quotient
15. Jitter period variability index.

b. It was found that 98.7% was correctly classified into normal group and 76.7% was correctly classified into dysphonic group (i.e.) two samples of normal voice (1.3%) were considered as belonging to dysphonic group and similarly 21 samples out of 90 (23.3%) dysphonic voice were considered as belonging to normal group by the discriminant analysis. Therefore, the parameters used were not sensitive to differentiate mild hoarse voice.

c. It was found that the following parameters were found to be useful in differentiating various degrees of hoarse voice.

1. Speed of Fluctuations in Frequency
2. Mean Fundamental Frequency in both Jitter and Frequency
3. Extent of Fluctuations in Frequency
4. Jitter Directional Perturbation Quotient
5. Shimmer Directional Perturbation Quotient
6. Frequency Range
7. Percent Jitter

d. It was found that 88.9% 50% and 66.7% were corectly identified into mild, moderate and severe degrees of hoarse voice respectively

CONCLUSIONS:

1. Fifteen out of 24 parameters studied were useful in differentiating between normal and dysophonic groups.

2. 7 parameters out of 24 parameters studied were found to be useful in differentiating various degrees of hoarse voice.

3. The norms generated from this study could be useful to the Speech Language Pathologist in diagnosis and treatment.

4. The findings also validate perception based vocal hoarseness ratings of listeners.

RECOMMENDATIONS:

1. A similar study may be conducted in larger samples

2. Studies using these parameters on peadiatric and geriatric subjects may also be tried.

BIBLIOGRAPHY

Anderson, V (1961): Training the Speaking Voice Oxford University, New York.

Ananthapadmanabha, T.V. (1992): HNR program based on cepstral analysis. Personal Communication.

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

Aronson, A.E. (1985): Clinical Voice Disorders. Second Edition. New York, Thieme Inc.

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

Balaji, O.G. (1988): Long term average spectrum and electrolottcgraphy in dysphonics-Unpublished master's dissertation. University of Mysore.

Baynes, R.A (1996):An incidence study of chronic hoarsenesses among children. JSHD, 31(2), 172-176.

Boone, D.R. and Plante, E. (1993): Human Communication and its Disorders. New Jercy, Printice Hall Inc.

Boone, D.R. (1971): The Voice and Voice Therapy. Second Edition, New Jercy, Printice Hall Inc.

Carhart, R. (1941): The spectra of model larynx tones. Speech Monographs, 8, 76-84. Cited in Emanuel, F.W. (1991): Spectral noise. Seminars in Speech and Language, 12(2), 115-130.

Carr, P.B. and Trill, D. (1964): Long term excitation spectra. JASA, 36, 2033-2040. Cited in Baken, R.J. (1987): Clinical measurement of speech and voice. London: Taylor and Francis Ltd.

Casper, M. , Abramson, A.L. and Forman - Franco, B. (1981): Hoarseness in children: Summer camp study. Int. J. Pediatric Otorhinol. , 3, 85-89. Cited in Sederholm, E. , McAllister, A., Sunberg, J. and Dalkvist, J. (1992): Perceptual analysis of child hoarseness using continuous scales. Speech Transmission Laboratory Quarterly Progress and Status Report, 1, 99-133.

Coleman, R.F. (1960): Some acoustic correlates of hoarseness- Unpublished master's thesis. Vanderbilt University, Nashville, T.N. Cited in Childers, D.G. and Lee, C.K. (1991): Vocal quality factors: Analysis, synthesis and perception. JASA, 90(5), 2394-2410.

Coleman, R.F. (1971): Effect of waveform changes upon roughness perception. Folia Phoniatrica, 23, 314-322. Cited in Askenfelt, A.G. and Hammarberg, B. (1986): Speech waveform perturbation analysis: A perceptual acoustical comparison of seven-measures. JSHR, 29(1), 50-64.

Cooper, M. (1974): Spectrographic Analysis of Fundamental frequency and hoarseness before and after vocal rehabilitation. Journal of speech and Hearing Disorders, 39, 286-296.

Cox, N.B., Ito. M.R., and Morrison, M.D. (1989): Data labeling and sampling effects in harmonic-to-noise ratios. JASA, 85(5), 2165-2178.

Curtis. J.F. (1978): Processes and Disorders of Comminication. New York, Harper and Row Publications Page 6.

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

Dejonckere, P.H. and Lebacq, J. (1985): Electroglotography and vocal nodules: An attempt to quantify the shape of the signal. Folia Phoniatrica, 37, 195-200.

De Krom, G. (1993): A Cepstrum-based technique for determining a harmonics-to-noise ratio in speech signals. JSHR, 36(2), 254-266.

Emanuel, F.W., Lively, M.A. and McCoy J.F. (1973): Spectral noise levels and roughness ratings for vowels produced by males and females. Folia Phoniatrica, 25, 110-120. Cited in Emanuel, F.W. (1991): Spectral noise. Seminars in Speech and language, 12(2), 115-130.

Emanuel, F.W. and Sansone, F.F. (1969): Some spectral features of normal and simulated rough vowels. Folia Phoniatrica, 21, 401-415. Cited in Emanuel, F.W.(1991): Spectral noise. Seminars in Speech and Language, 12(2), 115-130.

Fairbanks, G. (1960): Voice and articulation drillbook. New York: Harper and Bros.

Fourcin, A. (1981): Laryngographic assessment of phonatory function. ASHA Reports, 11, 116-127. Cited in Haji, T., Horiguchi, S., Baer, T. and Gould, W.J. (1986): Frequency and amplitude perturbation analysis of electroglottograph during sustained phonation. JASA, 80(1), 58-62.

Fukazawa, T. , El-Assuooty, A. and Honjo, I. (1988): A new index for evaluation of the turbulent noise in pathological voice. JASA, 83(3), 1189-1193.

Gill, J.S. (1961): Automatic extraction of the excitation function of speech with particular reference to the use of correlation methods - Proceedings of the third international congress on acoustics. Amsterdam: Elsevier, 1, 217-220. Cited in Horii, Y. (1979): Fundamental perturbation observed in sustained phonation. JSHR, 22(1), 5-19.

Gopal, N.K. (1986): Acoustic analysis of the speech in normal adults. An unpublished master's dissertation submitted to the Mysore University, Mysore.

Gobi, C and Ni Chasaide, A (1992): Acoustic characteristics of voice quality. Speech communication, 11, 481-490.

Gauffin, J., Hammarberg, B. and Hertegard, S. (1996): Irregularities in the voice: A perceptual experiment using synthetic voices with subharmonics. Cited in Davis, P.J. and Fletcher, N.H. (Eds): Vocal Fold Physiology, London, Singular publishing group.

Greene, M.C.L. (1964): The Voice and its Disorders. Whurr-AITBS, Delhi, Page 51.

Hartmann, E. and Cramon, D.V. (1984): Acoustic measurement of voice quality in central dysphonia. JCD, 17(6),425-440.

Hecker, M.H. and Kreul, E.J. (1971): Descriptions of the speech of patients with cancer of the vocal folds. Part I: Measures of fundamental frequency. JASA, 49, 1274-1282.

Heiberger, V.L. and Horii, Y. (1982): Jitter and shimmer in sustained phonation. In Lass, N.J. (Ed) Speech and language: Advances in basic research and practice. Vol. 7 (pp. 299-332). New York: Academic Press.

Hillenbrand, J., Cleveland, R.A., and Erickson, R.L. (1994): Acoustic correlates of Breathy vocal Quality. JSHR, 37, 769-778.

Hirano, M. (1975): Phonosurgery: Basic and clinical investigations. Otologia (Fukuoka), 21, 239-240. Cited in Baken, R.J. (1987): Clinical measurement of speech and voice. London: Taylor and Francies Ltd.

Huffman, M.K. (1987): Measures of phonation type in Hmong. JASA, 81, 495-504.

Hirano, M., Kakita, Y., Matsushita, H., Hiki, S., and Imaizumi, S. (1977): Correlation between parameters related to vocal cord vibration and acoustical parameters in voice disorders, Pract. Otol. Kyoto. 70(5), 393-403.

Hiraoka, N., Kitazoe, Y., Ueta, H., Tanaka, S. and tanabe, M. (1984): Harmonic intensity analysis of normal and hoarse voices. JASA, 76(6), 1648-1651.

Horii, Y. (1979): Fundamental frequency perturbation observed in sustained phonation. JSHR, 22(1), 5-19.

Horii, Y. (1980): Jitter and shimmer sustained phonation. JSHR 23(l) 202-209.

Horii, Y. (1980): Jitter and shimmer difference among sustained vowel phonations. JSHR, 25, 12-14.

Horii, Y. (1985): Jitter and shimmer in sustained vocal fry phonation. Folia Phoniatrica, 37(2), 81-86.

Imaizumi, S. (1986): Acoustic measures of roughness in pathological voice. Journal of Phonetics, 14, 457-462.

Imaizumi, S., Hiki, S., Hirano, M. and Matsushita, H. (1980): Analysis of pathological voices with a sound spectrograph (in Japanese). J. Acoust. Soc. Jpn., 36, 9-16. Cited in Kasuya, H. , Ogawa, S., Mashima, K. and Ebihara, S. (1986): Normalized noise energy as an acoustic measure to evaluate pathologic voice. JASA, 80(5), 1329-1334.

Isshiki, N. (1966): Classification of hoarseness. Jpn. J. Logop. Phoniat., 7, 15-21. Cited in Hirano, M. (1981): Clinical examination of voice. New York: Springer-Verlag Wien.

Is3hiki, N., Yanagihara, N. and Morimoto, M. (1966): Approach to the objective diagnosis of hoarseness. Folia Phoniatrica, 18(6), 393-400. Cited in Emanuel, F.W. (1991): Spectral noise. Seminars is Speech and Language, 12(2), 115-130.

Kreiman, J., Gerratt, B.R Precoda, K. and Berke, G.S. (1992): Individual differences in voice quality perception. JSHR, 35(3), 512-520.

Lieberman, P. (1961): Perturbation in vocal pitch. JASA, 33, 597-603. Cited in Askenfelt, A.G. and Hammarberg, B. (1986): Speech waveform perturbation analysis: A perceptual acoustical comparison of seven measures. JSHR, 29(1), 50-64.

Lively, M.A. and Emanuel, F.W. (1970): Spectral noise levels and roughness severity ratings for normal and simulated rough vowel3 produced by adult females. JSHR, 13(3), 503-517.

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

Michel, J.F. and Wendahl, R. (1971): Correlates of voice production (chapter - 18). In Travis, L.E.(Ed) Handbook of speech pathology and audiology. New Jersey: Prentice Hall pp. 465-480.

Moore, G.P. (1971): Voice disorders organically based (Chapter - 21). In Travis, L.E. (Ed) Handbook of speech pathology and audiology. New Jersey: Prentice Hall, pp. 535-569.

Moore, G.P. (1975): Observation on the physiology of hoarseness - Proceedings of the fourth international congress of phonetic science. Helsinki: Fin and (pp.92-95). Cited in Childers, D.G. and Lee, C.K. (1991): Vocal quality factors: Analysis, synthesis and perception. JASA, 90(5), 2394-2410.

Moore, P.,and Thompson, C.L. (1965): Comments on physiology of hoarseness. Archives of Otolaryngology, 81, 97-102.

Muta, H., Baer, T., Wagatsuma, K., Muraoka, T. and Fukuda, H. (1988): A pitch synchronous analysis of hoarseness in running speech. JASA, 84(4), 1292-1301.

Nessel, E. (1960): Uber das tonfrequenzspektrum der pathologisch veranderten stimme. Acta Otolaryngologica, Supplement 1457, 3-45. Cited in Emanuel, F.W. (1991): Spectral noise. Seminars in Speech and Language, 12(2), 115-130.

Noll, A.M. (1964): Short term spectrum and cepstrum techniques for vocal pitch detection. JASA, 36, 296-302. Cited in Bakery, R.J. (1987): Clinical measurement of speech and voice. London: Taylor and Francis Ltd.

Pathak, C.B., (1995): The quantitative evaluation of hoarseness using harmonics to noise ratio and voice perturbations. An unpublished master's Dissertation submitted to Mysore University, Mysore.

Paul Moore (1971): Voice disorders Organically based. Cited in Travis. L.E. (1971), Handbook - Speech Pathology and Audiology, New Jersey, Printice Hall Inc. Page 543-545.

Qi, Y. and Shipp, T. (1992): An adaptive method for tracking voicing irregularities. JASA, 91(6); 3471-77.

Rajanikanth, B.R., (1986): Acoustic analysis of the speech of the Hearing impaired. An unpublished master's Dissertation submitted to Mysore University, Mysore.

Rajkumar Pandita (1998) : Objective Analysis of Efficient Voice. Doctoral Thesis, University of Mysore.

Schoentgen, J. (1989): Jitter in sustained vowels and isolated sentences produced by dysphonic speakers. Speech Communication, 8(1), 61-80.

Scripture, E.W. (1906): Researches in experimental phonetics: The study of speech curves. Washington, D.C.:Carngegie Institute, (p.41). Cited in Horii, Y. (1979).-Fundamental frequency perturbation observed in sustained phonation. JSHR, 22(1), 5-19.

Sederholm, E., McAllister, A., Sundberg, J. and Dalkvist, J. (1992): Perceptual analysis of child hoarseness using continuous scales. Speech Transmission Laboratory Quarterly Progress and Status Report, 1, 99-113.

Shoji, K., Regenbogen, E., Yu, J.D., and Blaugrund, S.M. (1992): High frequency power ratio of breathy voice. Laryngoscope, 102(3), 267-71.

Silverman, E.M. and Zimmer, C.H. (1974): Incidence of chronic hoarseness among school age children. JSHD, 40(2), 211-215.

Simon, C. 91927): The variability of consecutive wavelengths in vocal and instrumental sounds. Psychological Monographs, 36, 41-83. Cited in Titze, I.R., Horii, Y. and Scherer, R.C. (1987): Some technical considerations in voice perturbation measurements. JSHR, 30(2), 252-260.

Sonninen, A. and Sonninen, E. (1976): Logopedic-phoniatric observations on eight year old school children in the town of Jyvaskyla. Nord, Tidskr. Logopediog Foniatri., 2, 74-82. Cited in Sederholm, E., McAllister, A., Sundberg, J. and Dalkvist, J. (1992): Perceptual analysis of child hoarseness using continuous scales. Speech Transmission Laboratory Quarterly Progress and Status Report, 1, 99-113.

Sridhara, R. (1986): Glottal waveforms in normals-Unpbulished master's dissertation. University of Mysore.

Stemple, J.C., Glaze, L.E., and Gerdeman, B.k. (1995): Clinical voice pathology, California, Singular publishing group, Inc.

Titze, I.R., Horii, Y. and Scherer, R.C. (1987): Some technical considerations in voice perturbation measurements. JSHR, 30(2), 252-260.

Wendhal, R.W. (1963): Laryngeal analog synthesis of harsh voice quality. Folia Phoniatrica, 15, 241-250. Cited in Baken, R.J. (1987): Clinical measurement of speech and voice. London: Taylor and Francis Ltd.

Wendahl, R.W. (1996): Laryngeal analog synthesis of jitter and shimmer auditory parameters of harshness. Floia Phoniatrica, 18, 98-108. Cited in Baken, R.J. (1987): Clinical measurement of speech and voice. London: Taylor and Francis Ltd.

Wewers, M.E. and Lowe, N.K. (1990): A critical review of visual analogue scales in the measurement of clinical phenomena. Research in Nurshing and Health, 13, 227-236. Cited in Sederholm, E., McAllister, A., Sundberg, J. and Dalkvist, J. (1992): Perceptual analysis of child hoarseness using continues scales. Speech Transmission Laboratory Quarterly Progress and Status Report, 1, 99-113.

Wilcox, K. (1978): Speaker age and vowel differences in vocal fundamental frequency perturbation-Unpublished master's thesis. Purdue University. Cited in Sorensen, D. and Horii, Y. (1984): Directional perturbation factors for jitter and for shimmer. JCD, 17(3), 143-151.

Wilcox, K. and Horii, Y. (1980): Age and changes in vocal jitter. Journal of Gerontology, 35, 194-198. Cited in Higgins, M.B. and Saxman, J.H. (1989): A comparison in intrasubject variation across sessions of three vocal frequency perturbation indices. JASA, 86(3), 911-916.

Wolfe, V.I. and Ratusnik, D.L. (1988): Acoustic and perceptual measurements of roughness influencing judgements of pitch. JSHD, 53(1), 15-22.

Wolfe, V.I. and Steinfatt, T.M. (1987): Prediction of vocal severity within and across voice types. JSHR, 30, 230-240.

Yanagihara, N. (1967): Significance of harmonic changes and noise components in hoarseness. JSHR, 10(3), 531-541.

Yumoto, E., Gould, W.J. and Baer, T. (1982): Harmonics-to-noise ratio as an index of the degree of hoarseness. JASA, 71(6), 1544-1550.

Yumoto, E., Sasaki, Y. and Okamura, H. (1984): Harmonics-to-noise ratio and psychophysical measurement of the degree of hoarseness. JSHR, 27, 2-6.