"SYNTHESIZED VOICE : AN APPROACH FOR VOICE THERAPY"

REG. NO : M-9420

Sheeja (G)

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

ALL INDIA INSTITUTE OF SPEECH AND HEARING

MYSORE - 570 006

MAY 1996.

DEDICATED TO

My

Achan, Amma

"You are the part of my life that gives me reason for tomorrow and the confidence to find it". "To walk in your footsteps is the highest form of respect I can give you".

CERTIFICATE

This is to certify that this dissertation entitled **"SYNTHESIZED VOICE : AN APPROACH FOR VOICE THERAPY"** is the bonafide work in part fulfillment for the second year M.Sc. (Speech and Hearing) of the student with Reg. No. M-9420.

Place : Mysore.

Date : May 1996.

Dr. (MISS). S.NIKAM
DIRECTOR
All India Institute of Speech and Hearing.
Mysore - 570 006.

CERTIFICATE

This is to certify that this dissertation entitled **"SYNTHESIZED VOICE : AN APPROACH FOR VOICE THERAPY"** has been prepared under my supervision and guidance.

Place : Mysore.

Date : May 1996.

GUIDE

Dr. N.P. NATARAJA
Head 'of 'the Department.
Dept. of Speech Sciences
All India Institute of Speech and Hearing.
Mysore - 570 006.

DECLARATION

I hereby declare that this dissertation entitled **"SYNTHESIZED VOICE : AN APPROACH FOR VOICE THERAPY"** is the result of my own study under the guidance of **Dr.N.P.Nataraja,** HOD, Department of Speech Sciences, All India Institute of Speech and Hearing, Mysore and has not been submitted earlier at any University for any other Diploma or Degree.

Place : Mysore

Date : May 1996. Reg. No. M-9420.

ACKNOWLEDGMENTS

1. **Dr.N.P.NATARAJA,** Professor and Head of the Department of speech sciences, All India Institute of Speech and Hearing, Mysore. Sir ... your unfailing guidance, support and patience is what has made this study possible to me, my admiration and respect for you sees no boundaries.

2. Iam grateful to **Dr. (Miss). S.Nikam** Director, All India Institute of Speech and Hearing, Mysore for having given me the permission and oppurtunity to undertake this dissertation.

3. A special thanks to **Sridevi Mam, Lalitha Mam, Rajiv P.** Without whose co-operation, my study couldn't have been completed.

4. **Amma, Achan, Sheeba & Vishu -** Thanks for every thing for loving, listening, caring, understanding, supporting and reassuring. "All that I am or hope to be, I owe to my family".

5. **Divya,** you make me feel so very special, you are the dearest in all the world to me, Its hard to find the words to say, how very much you mean to me.

6. **Kuru & Suchi,** I cherish our friendship and thank you for being such special people in my life.

7. **Deepti, Shobha, Priya, Reetha, Sneha, Aparna -** Thanks for being such good pals.

8. My class mates - time flies memories don't, thank you for these beautiful memories.

9. Some people are close to heart, Dear **Niky & Navan,** that's what you are to me.

10. To my aunts & uncles, I thank god for putting us all in a family.

11. To **Sunil,** thanks is a very small word. . . . . .at times.

12. Last, but not the least, thanks to **Mr. Venu gopal, Mr. Sudhakar** for the neat typing.

TABLE OF CONTENTS

LIST OF TABLES

46. Mean, SD range of speed of fluctuation in phonation of /i/ in normals, dysphonics and synthesized voices. 4.32

47. Mean, SD range of speed of fluctuation in phonation of /u/ in normals, dysphonics and synthesized voices. 4.32

48. Comparison of normals Vs dysphonics, normals Vs synthesized voice, and synthesized voice Vs dysphonic voices in terms of speed of fluctuation in intensity in /a/, /i/ & /u/. 4.32

49. Subjective Ratings for the normal, dysphonics and synthesized voices. (7 - point rating scale), made by 5 Judges. 4.35

LIST OF GRAPHS

INTRODUCTION

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

Voice plays an important role in speech and language. The production of voice depends on the synchrony between the respiratory, the phonatory and the resonatory systems. Any anatomical, physiological or functional deviation in any of these systems would lead to a voice disorder. Therefore voice problems must be treated i.e. help must be providedto the individuals with voice problems to overcome the problems or at least to cope with the problem.

"The treatment of patients suffering from dysphonia depends up to the ability to assess initially the type and degree of voice impairment and also to monitor the patient's subsequent progress throughout treatment" (Kelmen, 1981). "Diagnosis is intended to define the parameters of the problem, determine etiology and outline a logical course of action (Emerick and Katten, 1979).

Management of voice disorders is through medical or surgical and/or through therapeutic intervention. The term 'voice therapy' refers to the training or retraining of voice in terms of pitch, tone, quality, volume, breath support and rate. Voice therapy has truly become a blend of art and science.

Voice therapy is of different forms. It is highly individualized. Attempts have been made to treat different voice disorders such as hoarseness, nasality and pitch problems solely by changing the pitch (Williamson, 1944 and Masnemer, 1952).

Most voice therapy involves the identification and elimination of faulty vocal habits and their replacement by more optimum ones. The basic input modality in developing appropriate phonation is the auditory system, particularly the patient's 'self-hearing'. In voice therapy "we are concerned with making the patient a critical listener (Boone, 1967).

The importance of auditory feedback has been stressed by many authors in speech and voice training programs.

According to Boone (1967), it is the auditory feedback system by which one can actually monitor one's phonation.

Most of the therapies of voice disorders are based on the belief that each person has an optimum pitch at which voice will be of good quality and will have the maximum intensity with least expense of energy. And they concern themselves mainly with altering the habitual pitch level of making the case use his optimum pitch (West et al. 19571; Thurman, 1958; Van Riper, Irwin, 1958; Murphy, 1964, Greene, 1964).

Many clinicians stress on pitch discrimination and ear training in the treatment of voice disorders.

"Voice synthesis" is the process of producing an acoustic signal, by manipulating the parameters in a model of voice production".

Using synthesized voice, artificial pitch changing effects can be used as an approach for vocal rehabilitation.

Hence, the purpose of the present study was to synthesize the dysphonic voices, using synthesis programme, to create voice which would serve as models during therapeutic intervention. The research was designed to obtain synthesized voice which would approximate normal voice, so that it can be given as a feedback, auditorily to the dysphonics.

HYPOTHESIS

There is no significant difference between

(a) normals and dysphonics

(b) normals and synthesized voice groups

(c) synthesized voice groups and dysphonics in terms of these parameters.

(i) Mean, Fundamental Frequency (Hz) in phonation.

(ii) Maximum fundamental frequency in phonation for /a/, /i/, and /u/.

(iii) Minimum fundamental frequency in phonation for /a/, /i/ and /u/.

(iv) Frequency range in phonation.

(v) Fluctuations/sec in frequency,

(vi) Extent of fluctuations in frequency.

(vii) Mean intensity (dB) in phonation for /a/,/i/ and /u/.

(viii) Maximum intensity in phonation for /a/, /i/ and /u/.

(ix) Minimum intensity in phonation for /a/, /i/ and/u/.

(x) Intensity range in phonation.

(xi) Fluctuations/sec in intensity,

(xii) Extent of fluctuations in intensity.

LIMITATIONS OF THE STUDY

1. The sample size is very small. Hence, a large data can be collected in order to confirm the present study.

2. Only limited types of dysphonics have been studied.

3. It was considered that the parameters studied would be sufficient to differentiate between synthesized and dysphonics along with normals. Other parameters were not included.

REVIEW OF LITERATURE

"There is nothing more elemental in all existence than communication in humans we see its ultimate expression in the marvellous vehicle of language". (Van Riper and Emerick 1990). Communication has long been recognized as one of the most fundamental components of human behavior (Peterson 58). The ability of the human beings to use their vocal apparatus with other organs to express their feelings, to describe an event and to establish communication is unique to them. It took millions of years for human beings to develop this faculty. The onset of human era is recognized to have started with the acquisition of the ability to communicate using the vocal apparatus for social interaction. No normal person has failed to develop this faculty and no other species is known to have developed this ability.

Speech is the audible manifestation of language. It is the one form of communication which people use most effectively in interpersonal relationships. Speech is a sophisticated and fine movements of the components of the vocal tract and their complex interaction with one another. The speech results due to fine organisation, co ordination and modulation between the respiratory, phonatory, articulatory and resonatory system. With speech, people give form to their innermost thoughts their dreams, ambitions, sorrows, and joys, without these, they are reduced to animal noises and unintelligible gestures. In real sense, speech is the key to human existence. It bridges the differences of distance and helps to give meaning of purpose to their lives (Fisher 75).

According to Boone (1985), the act of speaking is a very specialized way of using the vocal mechanism demanding a combination or interaction of respiration, phonation, resonance of articulation. .

Voice is the vehicle of speech. It is the musical sound produced by the vibration of vocal cords in the larynx by air from the lungs. The importance of voice in speech is very well depicted, when one considers the cases of laryngectomee's and voice disorders.

"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). Voice is more than a means of communication of verbal messages clearly. Voice constitutes the matrix of human communication, infusing all parameters of human speech and unique self, one presents to the world voice has both linguistic of non - linguistic functions. The deqree of dependence of a language on these functions varies from for example, tone lanugane's rely more upon the voice or pitch, more specifically than other languages.

Voice is the carrier of speech, variations in voice in terms of pitch and loudness provide rhythm and break the monotony. This function of voice draws attention, when there is a disorder of voice, "voicing" (presence of voice) has been found to be a major distinctive feature in almost all languages. Voicing provides more phonemes and makes the language broader. When this information is absent or used abnormally, it would lead to a speech disorder.

At the semantic level also, voice plays an important role. The use of different pitches - high or low with the same string of phonemes would mean different hings. Speech prosody, the tone, the intonation and the stress or the rhythm of language are functions of vocal pitch and loudness as well as of phonetic duration.

Perkins (1971) has identified at least five non linguistic functions of voice. Voice can reveal speaker identity i.e. voice can give information regarding sex, age, height and weight of the speaker. Lass, Brong, Ciccolella, waiters of maxwell (1980) have reported several studies which have shown that it was possible to identify the speaker's age, sex, race, socio - economic status, racial feature, height and weight based on voice.

It is a prevailing notion that there is a relationship between voice and personality i.e. voice reflects the personality of an individual (stark weather 1916, Markel, Meisels and Havck 1964, Rousey and Moriarty, 1965). Fairbanks (1942, 1966) and Huttar 1967, have concluded from their studies that the voice reflects the emotional conditions reliably.

Voice has also been considered to be reflecting the physiological state of an individual for example, a very weak voice may indicate that the individual may not be keeping good health or a denasal voice may indicate that the speaker has common cold. Apart from these, it is a well known fact that voice basically reflects the anatomical and physiological conditions of the respiratory, phonatory and resonatory system, i.e. deviation in any of these systems may lead to voice disorders. Our voice reveals who we are and how we feel, giving considerable insight into the structure and function of certain parts of the body. (Titze, I.R. 1995).

A recently developed aspect in the area of early identification of disorders is infant cry analysis. It has been found by many investigators (illingworth 1981, Indira 1982) that it is possible to identify abnormalities in the neonate by analyzing their cry.

Speaker identification by voice would be of immense value in computer technology (development of machines that will respond to spoken commands), forensic medicine (identification of the speaker by his voice and lie detector) and in defence (availability of classified information).

The quality of voice may become important for certain professionals for example radio/TV announcers, actors and singers. Thus, voice has an important role in communication through speech and there is a need for studying voice.

The term 'voice' has been differently defined by different people. The Random house dictionary lists 25 primary and secondary definitions of voice. The first of which is sound or sounds uttered through the mouth of the human beings in speaking, shouting and singing.

Some definitions of voice restrict the term to the generation of sound at the level of larynx, while others include the influence of the vocal tract upon the generated tone and still others broaden the definition by including aspects of speech like articulation and prosody.

Judson and weaver (1942), define voice as "laryngeal vibrations (phonation) plus resonance" Further they state that phonation is laryngeal generator.

The formula P = S.T has been used by Fant (1960) in which speech sound P is the product of the sources and the transfer function of the vocaltract - T.

"When discussing the production of speech, it should be noted, that the source S, of the formula, P=S.T is an acoustic disturbance, superimposed upon the flow of respiratory air and is caused, by a quasiperiodic modulation of the air flow due to opening and closing movement of the vocal fold" (Fant 1960).

Michel and Wendahl (1971), after reviewing various definitions of voice, define voice as "the laryngeal modulation of the pulmonary air stream which is then further modified by the configuration of the vocal tract.

Thought there are varied definitions of voice. It is a difficult task to define normal voice.

An attempt has been made by Nataraja and Jayarama (1975) to review the definitions of normal voice, critically. They have concluded that each of the available definitions of voice have used subjective terms, which are neither defined nor measureable.

They have suggested the possibility of defining good voice operationally as the good voice is one which has optimum frequency as its fundamental (habitual) frequency.

It is apparent that a good voice is a distinct asset and a poor voice may be a handicap. If a person's voice is deficient enough in some respect, that it is not a resonably adequate vehicle for communication, if it is distracting the listener, one can consider this as a disorder-

In general, the following requirement can be set to consider a voice as adequate as stated by lwata of Von Leden (1978).

1) The voice must be appropriately loud.

2) Pitch level must be appropriate.

The pitch level must be considered in terms of age and sex of the individual. Men of women differ in vocal pitch level.

3) Vocal quality must be reasonably pleasant. This criterion implies the absence of such unpleasant qualities like hoarseness, breathiness, harshness and excessive nasality.

4) Flexibility must be adequate. Flexibility involves the use of pitch and loudness inflections. An adequate voice must have sufficient flexibility to express a range of differences in stress, emphasis and meaning. A voice which has good flexibility is expressive. Flexibility of pitch and flexibility of loudness are not easily seperable, rather they tend to vary together to a considerable extent.

Wilson (1962) is of the opinion that good voice should have the following characteristics.

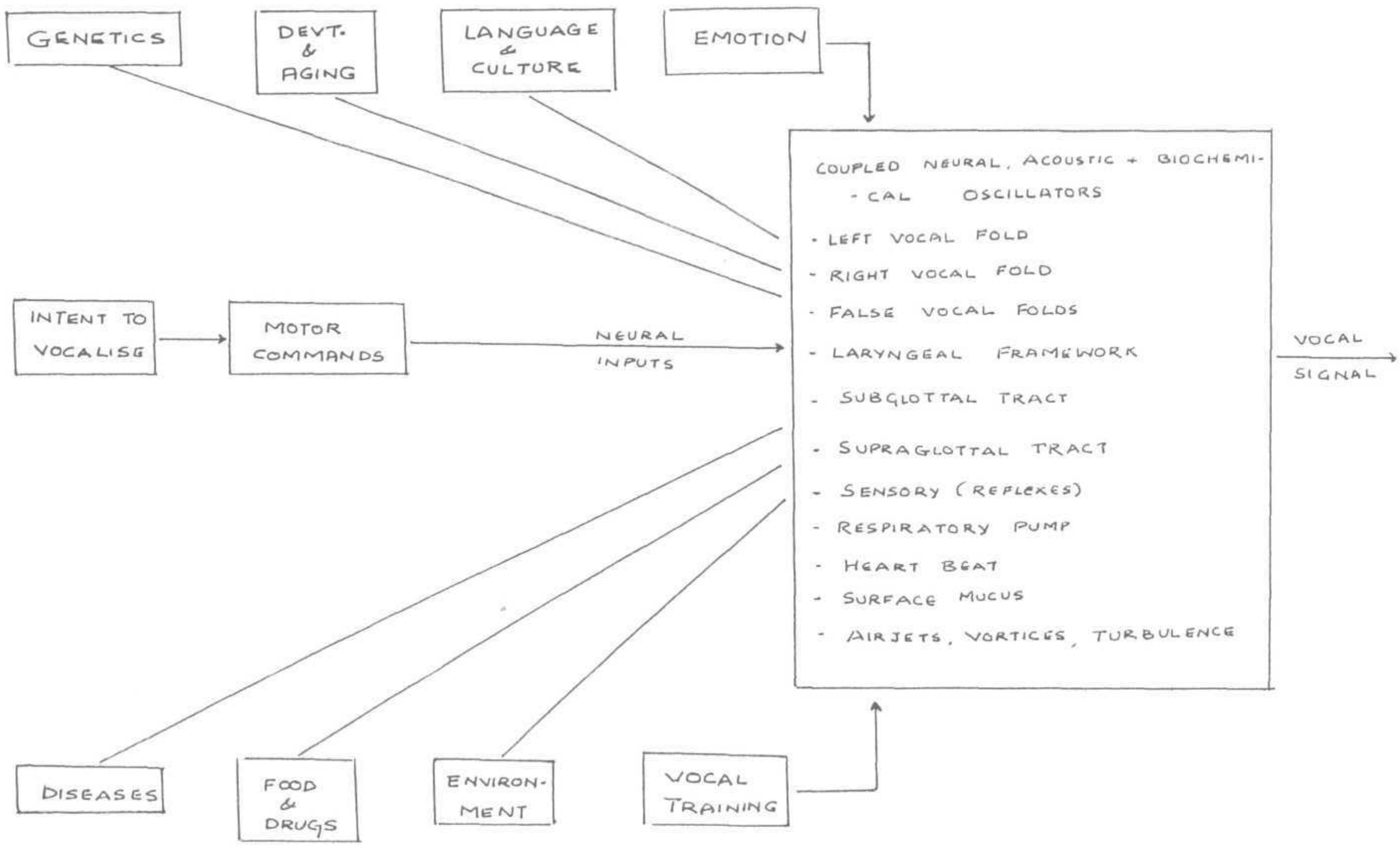
a) Pleasing voice quality.
b) Proper balance of oral and nasal resonance.
c) Appropriate loudness.
d) A model frequency level suitable for his age of sex.
e) Appropriate voice inflections involving pitch and loudness.

The production of voice, depends on the synchrony or the co-ordination between the systems viz, the respiratory, phonatory and resonatory. Voice production involves a complex and precise control by the central nervous system of a series of events in the peripheral phonatory organs. The crucial event essential for voice production is the vibration of the vocal folds it changes DC air stream to AC airstream converting aerodynamic energy into acoustic energy.

Voice production can be thought of as the activation of an entire system of coupled oscillators. The intent to vocalise activates motor commands that are responsible for the neutral inputs to an array of bio-mechanical, neural and acoustic oscillators. The vocal folds are the primary oscillating system that produce what we might call the carrier signal with glottal airflow. All these oscillators can be thought of as modulators of the carrier signal. Some of the modulations are nearly sinusoidal (respiratory, heartbeat) but many ar high dimensional (action potentials of muscles, air vortices, mucus in motion). Yet others are passive oscillators (tracheal resonator, supraglottal vocal various sinuses) that can influence the primary oscillating system.

The system of coupled oscillators contains and releases information about the human body, in particular, about its genetics, development, age, disease, language, culture, food & drug intake, and response to the environment. (Titze)

Fig (1)



A LIST OF BIOLOGICAL OSCILLATORS INVOLVED IN VOICE PRODUCTION AND FACTORS THAT MAY INFLUENCE THEM.

Two broad cateqotries of theories have dominated in dealing with voice production. They are Myooelastic aerodynamic theory and neurochronaxic theory.

Myoelastic aerodynamic theory (Muller 1843) holds that phonation is the result of the balancing of forces of air pressure against tension, elasticity and mass of the vocal folds. Displaced by the air pressure the vocal folds return, to a resting state due to combination of factors, the chief ones being the drop of air pressure at the glottis following the valvular opening of vocal folds and the vocal fold mass and elasticty. They function of the vocal fold themselves is in large part passive. As in respiration, the final movements of the vocal folds are not under specific conscious control.

Neurochronaxic theroy (Husson 1950) hold that vocal fold vibration is an active process. Motor impulses are said to be emitted from cortical centres to the muscles of the folds via the recurrent laryngeal nerves under the regulations of a "cochlear reccurential reflex" vocal fold stimulation of this kind assumes that the recurrent nerve is capable of transmitting high frequency stimuli i.e of the order of 1,000 impulses per second.

The crucial event for voice prooduction is the vibration of vocal folds, it changes DC air stream to AC air stream, converting aerodynamic energy into acoustical energy. From this point of view, the parameters involved in the process of phonation can be divided into three major groups:

1. The parameters which regulate the vibratory pattern of the vocal folds.
2. The parameters which specify the vibratory pattern of the vocal folds.
3. The parameters which specify the nature of sound generated, (cotz, 1961).

Hirano (1981) has further elaborated on this, by stating that "The parameters which regulate the vibratory pattern of the vocal fold can be divided into two groups: Physiological and physical. The physiological factors are those related to the activity of the respiratory, phonatory and articulatory muscles. The physical factors include the expiratory force, the conditions of the vocal folds and the state of the vocal tract.

The vibratory pattern of the vocal folds can be described with respect to various parameters including the fundamental frequency, regularity or periodicity, in successive vibration, symmetry between the two vocal folds, uniformity in the movements of different points with in each vocal fold, glottal closure during vibration, contact area between the two vocal folds and so on.

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

Thus, voice has various functions which are varied too and it lays a major role in speech and hence in communication. Therefore voice needs to be constantly monitored and in the event of abnormal functioning of voice an immediate assessment should be undertaken. This assessment wil lead to diagnosis which not only identifies the voice disorders, but also acts as an indicator for the treatment and the management to be followed.

The purposes of clinical evaluation of voice are :

1. To diagnose the etiological disease.
2. To determine the degree and the extent of the etiological disease.
3. To evaluate the degree and nature of dysphonia.
4. To determine the prognosis and
5. To monitor change.

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

The management of voice problem is through either medical, surgical or therapeutic intervention. After the primary pathology, if any, is treated therapeutic intervention is done, if the voice problem persists or to correct the undesirable habits in producing voice. Again based on the nature, extent and servity of the voice disorder and/or a combination of the intervention strategies are considered. Example a vocal nodule may require all three, while puberphonia with pitch breaks requires only voice therapy. As effective management requires, it becomes necessary to describe and use a wide battery of tests or assessment strategies in order to arrive at an effective diagnosis.

There are various means of analysisng voice, developed by different workers, to note the factors which are responsible for creating an impression of particular 'voice' (Hirano 1981, Nataraja and Jayarama 1979, Rashmi, 1985).

There are various methods of direct or indirect assessment, observations and/or measurement of the parameters involved in the process of production of voice. Some of those selected clinical examinations which are specific or directly related to voice include:

a. Acoustical analysis of voice

b. Aerodynamic measurements

c. PSychoacoustic evalvation of voice

d. Examination of phonatory ability

d. Methods to study vocal fold vibration these include :

1. Stroboscopy
2. Ultra sound glotoography/echoglottgraphy.
3. Ultra high speed photography
4. Inverse filtering
5. Photo electric glottagraphy (P G G)
6. Electroglottography.

ACOUSTIC ANALYSIS OF VOICE

Acoustic analysis has been considered as the basic tool in the inestigation of voice disorder. It has been considered vital in the diagnosis and management of patient with voice disorders.

Hirano (1981) has pointed out that the acoustic analysis of the voice signal may be one of the most attractive methods for assessing phonatory function or laryngeal pathology because it is non invasive and provides objective and quantitative data.

Analysis of acoustic signals of the human voice has many purposes. From a technological stand point, ther is an ever-growing need to store, code, transmit and synthesize voice signals. From a basic science stand point, investigators have traditionally studied the microphone signal to understand speech production and perception, given that the acoustic signal is the common link between them. Finally, from a health science standpoint, the human voice has been shown to carry much information about the general health and wellbeing of an individual. (Titze 1995)

Many voice tests, are, infact necessary for the diagnosis of the etilogic disease. Further, a clinican will not really know what to expect with a medical diagnosis having complete physical description of larynx together with some adjectives like 'hoarse' or 'rough' until be actvally sees the case. (Michael and wendahl 1971) on the other hand, if the cilinican recieves a report which includes measures of frequency ranges respiratory function, jitter, shimmer, their related variation, noise and harmonic components etc. in the form of a voice profile, the clinician can then compare these values to the norms for each one of the parameters and thus have a relatively good idea as to how to proceed with therapy. Moreover, periodic measurement of these parameters during the course of therapy may well provide an useful index so as, the success of the treatment. (Michael and Wendahl[1] 1917).

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

Jayarama (1975), has made an attempt to compare some of the parameters of voice between normals and dysphonics. A siqnificant difference in the habitual freqvency measures were got between the subjects of both groups. Nataraja (1972), Samuel (1973), Shanta (1973) Sheela (1974), Asthana (1977) have used stroboscope with tacho unit and SPL meter to determine fundamental frequency of voice in their studies. The subjects were instructed to phonate a vowel in his normal speaking voice and this phonation was fed to the stroboscope through the SPL meter and Tacho unit. The fundamental frequency was read directly from tacho unit. There are various methods to evaluate these parameters stroboscopic procedure, perdue, pitch meter, high speed cinematography, digipitch, pitch computer, ultrasonic recordings and the high resolution signal analysis.

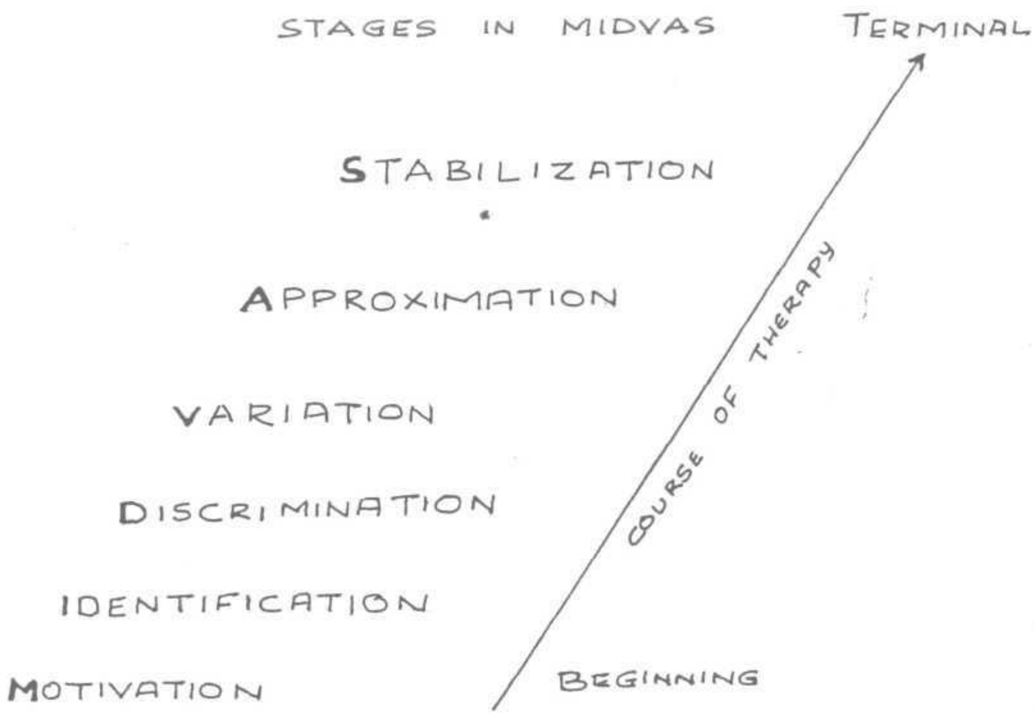
There are various means of analysing voice developed. Some of these being, the long term avetage spectrum which provides information on the spectral distribution of speech signal over a period of time ; spectrographic analysis, pitch perturbations, Harmonic to noise ratio.

VOICE THERAPY

Management of voice disorder is through medical and/ or surgical and/or thevapeutic intervention i.e, in that order that if medical and surgical intervention have taken place therapeutic intervention is done, if the voice problem persists or to correct the undersirable habits in producing voice.

Voice therapy has truly become a blend of art with science. The term 'voice therapy' refers to the training or retraining of voice in terms of pitch, tone focus, quality, volume, breath support and rate. It is often inherent as an integral part in resolving functional and organic dysphonias.

"MIDVAS" is an acronym where each seperate letter of this word refers to the goal of x particular phase of therapy. Briefly, 'M' refers to motivation any case comes to the clinician, for voice therapy has to be motivated. This is the primary phase of rehabilitation. The second phase of therapy in which the basic goal is the identification and evaluation of various factors in the client's problem. This is the period of self study, of self exploratory. Desensitization is the third major phase in the treatment. The major goal in this part of the therapy is to toughen the case to those factors which normally increases his problems. Here the clinician models for the case. It is not enough to motivate, to identify and desensitise, although these bring reductions in the problem. In this phase, change is made to modify the reaction. Here, the patient has to be taught to discover a new voice by varying pitch, intensity and quality. Progressive approximation is the next phase, here the concept of auditory feed back can be used percieving the differences in the voices, the case can approximate his voice. Once the patient learns to use it in therapy sessions consistently, patient will have to master the new voice in extraclinical situations. This phase is called stabilization" (Van riper, IRWIN - 1976)



Voice therapy is of different forms. The kind of therapy given to people, who simply want to improve their voices might vary markedly from that given to a patient with a paralysed vocal fold. Voice therapy is highly individualised according to the physical problem, the length of its existence, the quality of voice, feelings of the patient about the problem. In general, a four point program (Boone (1993)) is followed with cases of voice problems:

1. Indentifying abuse or misuse.
2. Reducing its occurence.
3. Using the dianostic probe.
4. Practicing facilitating approaches.

Attempts have been made to treat different voice disorders such as hoarseness, uasality and pitch problems soley by changing the pitch, Williamson (1944), Masnemer (1952).

Luschsinger, classifies voice therapies under three groups :

1. Physical therapy
2. Treatment of disorder of phonic respiration.
3. Medicinal.

PHYSICAL THERAPY :

Various forms of physiotherapy in the form of heat, infrared light diathermy, electrical stimulation with several types of currents vibrational massage and so forth were claimed to be as effective in paralytic condition of the larynx as they are else where in the body.

a. Larynqeal manipulaltions: two types of compressions are employed :

- frontal and lateral

b. Electrical stimulation

c. Activation of elementary laryngeal functions. Brodnitz (1961) Weiss(1932) have suggested that the pushing excercise may be used an part of therapy, the flexed arms are elevated to the chest and vigorously pushed down. AT the same time the patient phonates single syllables. This method aims at activation of the primitive, protective sphincteric action of laryngeal closure.

From the viewpoint of vocal physiology, the establishment of a vicarious regression to a lower functional level of phonation is not a goal of vocal rehabilitation.

d. Auditory training :

For centuries musical practice has made use of the empirically discovered feed back machanisms of auditory, tactile and proprioceptive monitoring.

Voice is ephemeral, lacking finite acoustic boundaries. A mildly defective voice can become incorporated into the self without much notice especially in children whose concerns over its acceptability are minimal. Even in severe voice disorders, the person may be aware of and upset over the voice but has long forgotten the sound of normal voice and therefore any notion of how it ought to sound again.

For these reasons, auditory training defined as teaching the identification and discrimination among different voices, is the staple technique of all voice therapy. Before a better voice can be achieved, the person has to know how his or her voice sounds. Patients of all ages need to hear the differences between normal and defective voices. After comparing them with their own, they should discuss those differences with the clinician. Of major importance is instantaneous auditory feedback. What is singular about voice is that in the early, critical stages, improved voice will break through suddenly and momentarily milliseconds in duration, although the clinicians may hear and identify these gains, the patient usually does not consequently, the clinican needs to listen carefully and, when the voice changes for the better or worse, communicate that information instanteneously to the patient.

2. Treatment of disorders of phonic respiration

a. Breathing therapy :

One of the advocates of breathing therapy is Hosbaver (1921) + (1948). This exercise aims at the systematic prolongation of phonation time. The ratio between inspiration and expiration was regulated by a timing device. Another method used was walking while humming two steps indicate the time of inspiration while the following ten to twelve steps time the humming expiration.

b. Active relaxation therapy

This was developed by fause (1954). This method is based on the principle that under normal conditions any type of tension is followed by release. Bur tensions stresses may amount without subsequent intervening releases, The respiratory and laryngeal musculature react with a forced manner of breathing and a squeezed type of phonation. Relaxation therapy aims at correction of excessive physical neuromuscular tensions.

c. The chewing method : Originated by froeschles (1952), advocated as a nonspecific relaxation method in the treatment of all those voice disorders in which the functional disturbance predominates.

d. The Yawning method : Proposed by lancau (1952). Yawning represents a prolonged and deepened inspiration with maximal widening of the upper airways.

e. Autogeneous training : Schultz (1953) has developed a psychotherapeutic method based on the psychological and physiological phenomena associated with hypnosis, which he calls autogenous training.

A majority of the experiments end up saying recovery depends upon the severity of the problem. Another factor concerning layngeal manipulations is that we do not yet have a definite idea about the type of cases which would benefit by this technique.

Much of the voice therapy cited above is a process of experimentation with the individuals voice. Ascertaining the habitual pitch level, altering it towards an arbitary level and checking the effects on overall voice usage as a function of changes in intensity and pitch consume most of the early phases of voice therapy. A major focus is centered on the discovery of the new voice.

There are many facilitating techniques. The most frequently used vocal techniques in voice disorders are tabulated by Murphy.

1. Determine and establish optimal pitch range.
2. Alter loudness level
3. Alter loud staccato tones.

-Relax musculature or reduce tension
-Increase muscle tension.
-Develop soft, clear vocal attack.
-Increased balanced resonance.
-Increased size of mouth opeining
-Lower the tonque
-Move the tonque forward.
-Increase pitch range
-Increase vocal variety in pitch
-Increase vocal variety in loudness
-Improve articulation ability
-Loud sign technique
-Singing.

-Humming

-Yawning

-Locating "best vowel" and fanning out

-Eliminate vocal abuse

-Vocal rest

-Increase kinesthetic awareness

-alter rate of speech

-Pushing exercise

-Chewing method

-muscle training

-External manipulation of vocal mechanism

-Velopharyngeal control

-Blowing exercises

-Oral pressure build up and release

-correct abnormal postures'

-Imitating voice of others

-Coughing, throat clearing, grunting

-Negative practice

-Speaking against background noise

-Alter respiratory patterns

-Carry over new voice to life situations

-Psycho therapy.

-Auditory techniques : Self listening sound discrimination, matching and comparing voices in quality, loudnes and pitch, imitating. Perkins (1972) commenting on these vocal techniques says that "Unfortunately such an abundance of procedures bespeaks dissatisfaction with results achieved as much as it atlests to clinical ingenuity. There is no common criteria for selecting a technique from this plethora of tactics which are claimed to be most effective. Chances are that all have worked well with some clinican for some time for some patient. What few reports of clinical results are available offer limited help they are based on small samples of patients with widely varied problems treated with diverse combination of techniques under uncontrolled conditions and evaluated by such disparate criteria as improvement in laryngeal pathology to improvement in vocal tone. Lacking firm evidence, up to this time, we had no alternative but to rely on clinical judgement without a clear rationale for achieving an ambiguously defined goal". Some of the therapy techniques listed are found to be inadequate on the following grounds.

In spite of many techniques available to elicit optimum pitch or desired pitch, some times. Some cases fail to achieve the target for many reasons. Clinical observations of patients with hyper functional voice problems suggest that many of these patients may experience difficulty in singing a tone matching a pitch or dicriminating between pitches (seashore, 1938, Travis and Davis 1928; Hanely 1956, Eisenson kospein (1958) studies suggest clinical groups to be significantly poorer in pitch discrimination than the control groups.

In voice therapy, we are concerned with making the patient a critical listener (Boone 1967).

There are only a few studies available on frequency discrimination ability in dysphonics (Gilkinson 1943, Eisenson et al 1958, Boone et al 1967) even though poor frequency disctimmination has been cited as one of the possible causes of dysphonia.

Gilkinson (1943) and Eisenson et al (1958) indicate a relationship between poor frequency discrimination and dysphonia, while Boone (1967) study indicates no such relationship, though he suggests that individual dysphonics who have good frequency discrimination show better prognosis. However, there seems to be an agreement on the need for ear training in voice therapy.

Eisenson et al (1958) measured frequency discrimination in their groups of dysphonics, using the seashore measures of musical ability, before and after voice therapy and ear training. They found that the scores increased significantly after the training period.

The importance of feed back has been stressed by many authors in speech and voice training programs. Tactual and proprioceptive feeback are the other common modalities, by which one gets some information while speaking. But it is the auditory feedback system according to Boone (1967) by which one can actually monitor one's phonation.

Most of the therapies of voice disorder are based on the belief that each person has an optimum pitch at which the voice will be of good quality and will have the maximum intensity with least expense of energy and they concern themselves mainly altering the habitual pitch level of making the case use his optimum pitch (west et al 1957, Thurman 1958 Van Riper, Irwin 1958, Murphy 1964, Grene 1964).

Many clinicans stress on pitch discrimination and ear training in the treatment of voice patients.

Van Riper (1963), while discussing therapy with voice patients, states that one of the ways of using progressive approximation in voice therapy is the use of a binaural auditory trainer, feeding cases voice into one ear and therapists model voice into the other ear, so that the patient has a simultaneous comparison to make from unison slight changes towards the desired pitch are made, such that the patient unconsciously switches over to the new voice as he percieves it, "The basic development of the input modality in voice therapy, or appropriale phonation, is the appropriate auditory system, especially patients self hearing" (Boone 1967). Boone (1967) further states that many people rarely realise how their voices sound until they hear their recorded sample, hence most clinicians face a problem during therapy with individuals who have lack of voice feeback. Some dysphonics, he continues, like some individuals in the normal population demonstrates poor pitch discrimination and tonal memory and these patients face more difficulty in voice therapy in discriminating between pitches and remembering their own model voice. These patients according to Boone can be given ear training to differentiate between 'good' and 'bad' voices and voice training should include instruction in pitch discrimination. He concludes that clinician should however assess the patients ability in this area first and only if deficient in this aspect will the patient benefit from such training.

Van Riper and Irwin 1968 in explaining MIDVAS as applied in voice therapy state that only after making variations and discriminations, should the patient be instructed to produce the model pitch, by approximation should be finally stabilised.

A Russian voice physician Malutin (1897, 1924) first stated the principle of improving phonation through application to the larynx of a mechanical vibration of the same frequency as that of the vocal tone. This has been exclusively studied by H. Gutzmann (1911-24). Numerous instruments have been devised for this purpose.

Isochronal tone stimulation is based on the principle of applying a machanical stimuli through tactile channel to the larynx in order to correct the pitch. The essential point is to achieve a harmonic relationship between these impulses and laryngeal tone to be produced when two vibrations differ slightly (10 cycles) in phase and frequency, they result in beats. Apart from being audible these beats can be felt as a peculiarily unpleasant sensation as reported by cases or they can be seen on oscilloscope. "As soon as the beats appear, the cerebero laryngeal mechanism begins to adjust the vocal cord vibrations, until union is achieved. (Luschsinger 1905) Shantha (1973) concludes that isochronous tone stimulation was found to be useful in a majority of voice disorders. By changing pitch and by providing optimum frequency, voice problems such as puberphonia nasality, hoarseness spastic dysphonia can be treated.

Most voice therapy involves the identification and elimination of faulty vocal habits and their replacement by 'more optimum ones. The basic input modality in developing appropriate phonation is the auditory system particularly the patients self hearing. No one has much awareness of what he is doing laryngeally ; whether he is approximating his folds or shortening or lengthening them, except as he hears his voice. The surprise nearly always evoked in people at hearing their own voice on recordings is one indication of how gross our self hearing is. This lack of voice feedback has always presented problems to the clinician as the patient literally does not know what he is doing when he phonates. He may need practice in learning to listen to his own voice. David and Boone 1945 report that some voice patients like some people in the normal population, demonstrate difficulty in pitch discrimination and tonal memory as measured by subtests of the seashore tests of musical aptitude. Such patients may have serious problems in voice therapy in making pitch discriminations and in remembering the sound of their own model voices. Through the use of auditory feedback devices such as loop tape recorders, he learns to hear and monitor auditorily his own phonation. For patients who have defective litening skills, voice training must include instruction in making pitch discriminations improving tonal memory and learning to hear ones "good" and "bad" voices. But the clinicians should first, assess the parients listening skills, for many voice patients have no problem in this area for others, just as for some people in the normal population, listening abilities may be surprisingly deficient. It is the latter group that may profit from ear training.

As sommers and Brady (67) have stated "improvements in phonation and resonance are heavily dependent upon the sujects obility to detect desirable changes as a function of specific voice therapy activities".

SPEECH ANALYSIS : Speech analysis can be thought of as that part of voice processing that converts human speech to digital forms suitable for transmission or storages by computers 'speech synthesis' functions are essentially the inverse of speech analysis they reconvert speech data from a digital form to one that similar to the original recoding and suitable for playback.

SPEECH SYNTHESIS

Speech synthesis is the process of producing an acoustic signal by controlling a model of speech production with a set of parameters if the model and parameters are sufficiently accurate then the production of intelligible synthetic speech should be possible. There are two basic approaches in modelling the speech production process. One is direct approach which attempts to model the system in detail. This is commonly referred to as articulatory speech synthesis and attempts to directly model the motion of the speech articulators as well as the generation and propagation of sound inside the vocal tract. This approach is still the subject of research and although it seems to have the potential for producing the most natural sounding speech in the long term it has not as yet been as successful as approaches, that attempt to simply copy the frequency response characteristic of the vocal tract.

Various synthesizers have been designed, over the years. Some of them formant synthesizers, copy synthesizer, linear predictive synthesizers, phoneme synthesizers text to speech synthesizers etc.

Using synthesized voice, artifical pitch changing effect can be used as an approach for vocal rehabilitation.

Anne - Maria Laukkanen (1994) studied the effects of artificial pitch change of the auditory feed back on the fundamentals frequency of normal hearing subjects in text reading. The subjects taken were mainly speech trainers and trained singers, and also subjects using non-optional speaking pitch. The results suggested that changing the pitch of the auditory feedback can make a person change his habitual pitch. Thus, it was suggested that the method might be worth testing in voice training and therapy practise.

METHODOLOGY

The objectives of the study were:

1) To determine the parameters in the dysphonic voice, which are deviating from normal voice.

2) To modify the parameters deviating towards normal using synthesis program.

3) To use the (modified) synthesized voice model in voice therapy.

SUBJECTS:

Five dysphonics (3 males, 2 females) who visited All India Institute of Speech and Hearing, Mysore with complaints of voice problems in the age range of 13-29 years, formed the experimental group. These subjects had been diagnosed as cases of "Voice disorder" after the routine otolaryngological, speech, psychological and audiological evaluations. Ten normal subjects (male and female) in the age range of 13-25 years were also considered for the study. The subjects of this group had no apparent speech, hearing or E.N.T. problems. They had no complaints about their speech, hearing or voice.

PROCEDURE;

TEST ENVIRONMENT:

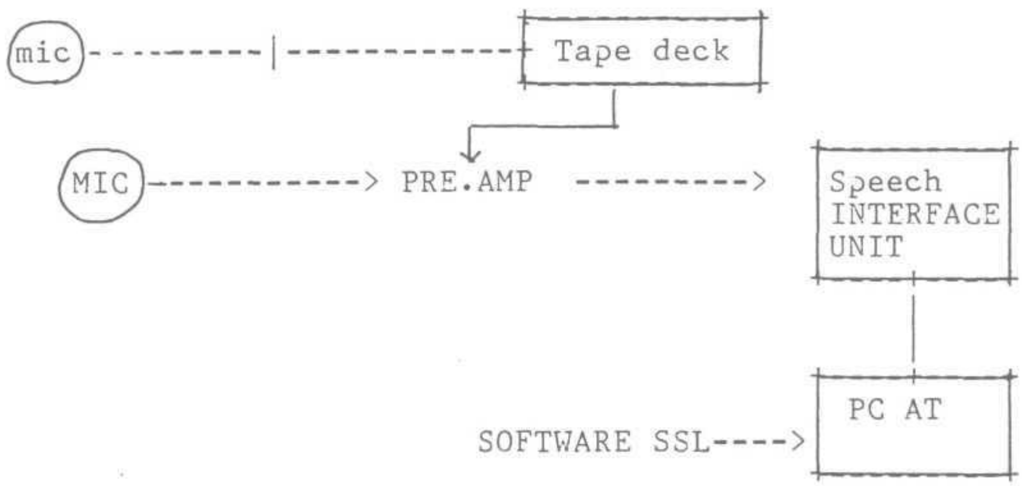
The recording was carried out in a sound treated room of the phoniatrics laboratory of the department of speech sciences. AIISH, Mysore.

INSTRUMENTATION:

The following instruments were used in recording and analysis:

1. Dynamic microphone (AHUJA AUD - 535M)
2. Pre-amplifier (PHILLIPS, PHILLIAMP 60)
3. Sony tape deck (TC FX 170)
4. Speech interface unit (Voice + Speech systems, Bangalore)
5. Vaghmi/SSL software
6. PC - AT.

BLOCK DIAGRAM:



SPEECH SAMPLES:

Sustained phonation of vowels a, i, u were used.

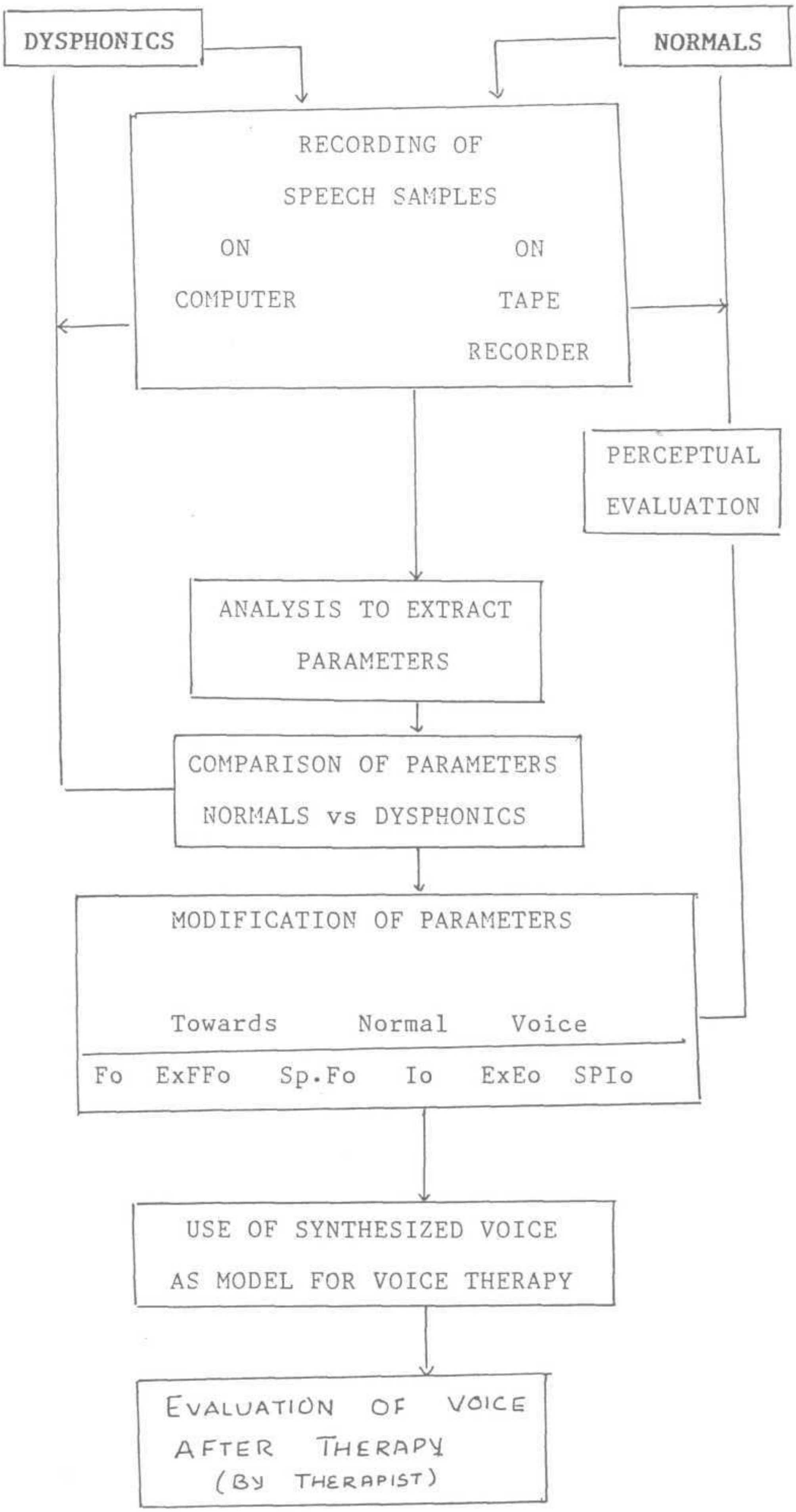
RECORDING OF SAMPLES:

The subjects were seated comfortably in a chair in the sound treated room. The dynamic microphones (AHUJA, AUD - 535M) were kept in front of the subject at a distance of about 15 cms from the mouth. They were instructed to take a deep breath and phonate /a/. They had to maintain a constant intensity and pitch at comfortable level, simultaneously, the output was recorded on the computer and the tape recorder. The tape recorded samples were played back to the input of the speech interface unit for digitization.

Similarly recording of vowels /i/ and /u/ were carried out.

(REFER TO THE FIGURE)

: SCHEMATIC REPRESENTATION OF THE DIFFERENT STAGES CARRIED OUT IN THE STUDY :


DYSPHONICS
NORMALS
RECORDING OF
SPEECH SAMPLES
ON
COMPUTER
ON
TAPE
RECORDER
PERCEPTUAL
EVALUATION
ANALYSIS TO EXTRACT
PARAMETERS
COMPARISON OF PARAMETERS
NORMALS vs DYSPHONICS
MODIFICATION OF PARAMETERS
Towards Normal Voice
Fo ExFFo Sp.Fo Io ExEo SPIo
USE OF SYNTHESIZED VOICE
AS MODEL FOR VOICE THERAPY
EVALUATION OF VOICE
AFTER THERAPY
(BY THERAPIST)

The dysphonic and the normal voices were analyzed using the VAGHMI software to obtain fundamental frequency and related measurements i.e., to find out the parameters in dysphonic voices which were deviating from normal, so that these parameters could be corrected to obtain normal/near normal voice.

These parameters were then compared i.e., between normal voice and the dysphonic voices.

Frequency Parameters.

Mean FO (Hz)

Maximum FO

Minimum FO

Range (FO)

Fluctuations / Sec.

Extent of fluctuations

Intensity parameters:

Mean AO (dB)

Maximum AO

Minimum AO

Range AO

Fluctuations / Sec.

Extent of fluctuations.

After comparison of the parameters the parameters deviating from normal voice were modified using synthesis program.

Experiment No.2:

Using the software program the digitized signal were analyzed to extract fundamental frequency, intensity bandwidth information. The analyzed signal was then synthesized using the steps.

STEPS OF SYNTHESIS:

FBAS (Formant based analysis synthesis model) implemented in SSL was used. In the FBAS sub-module, analysis of speech signal was performed using the autocorrelation method, to obtain the source and filter parameters at a uniform frame rate.

Programs for editing, FOEDIT and TXTTRK were used to edit the source parameters and the formant data respectively.

'FOEDIT' gave a graphic display of various source parameters. FOEDIT was used to edit the source parameters like FO, intensity, and voice source parameters, like open quotient, speed quotient and leak quotient. Editing operations like 'interpolate' and 'change' were used to edit the parameters. Fundamental frequency was changed to the optimum frequency of the individual for e.g. FO of 178Hz was changed to his optimum frequency of 120Hz

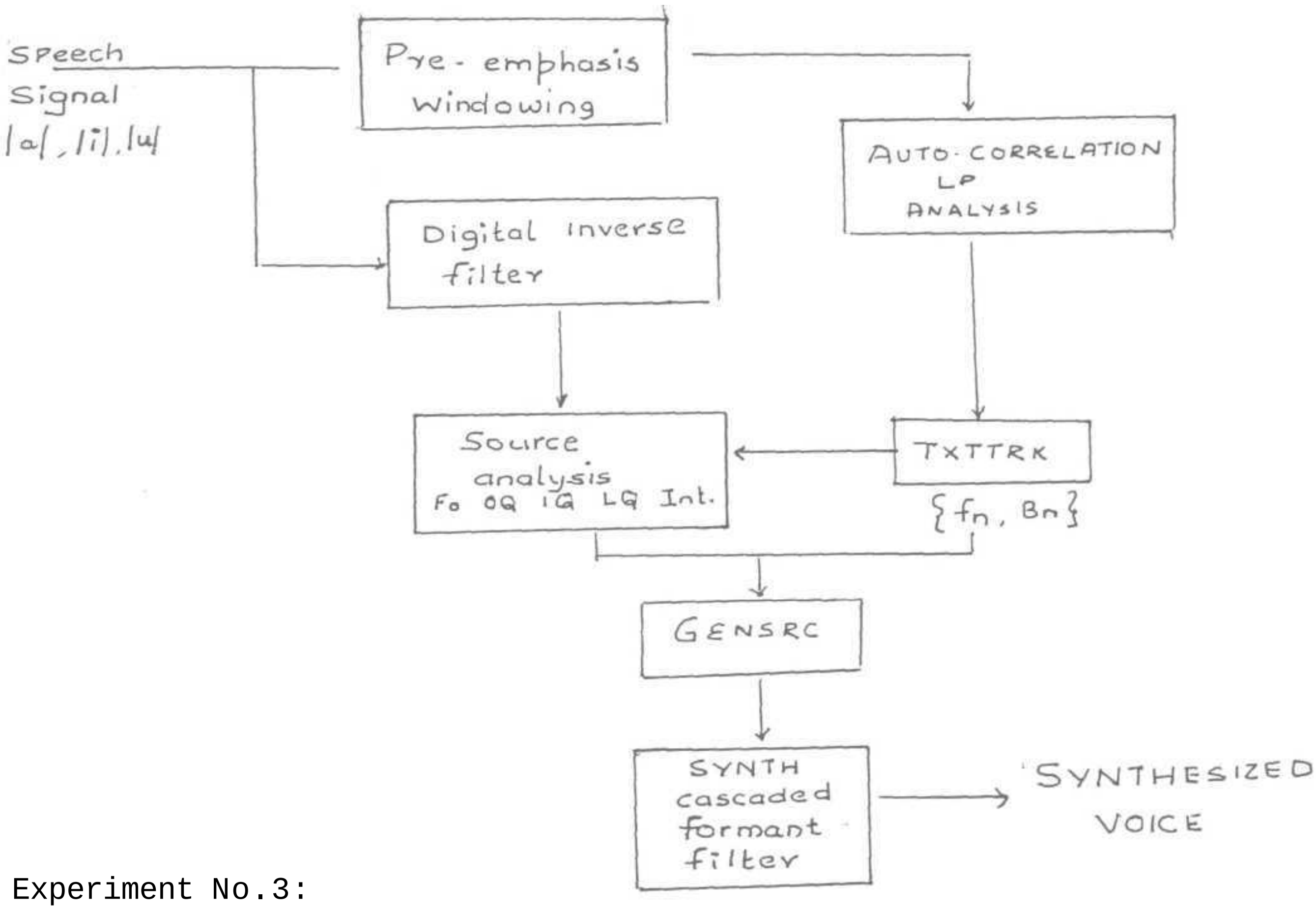
Using the same program intensity parameters were also varied for e.g. Intensity (original) of 48. 49dB was varied to 49 dB

In the next stage of synthesis included the TXTTRK program. This program presented the data i.e., formants and bandwidths in a text mode. This program was used to change and / or delete and / or interpolate bandwidths between two locations (points) (i.e., from Oms frames to 1000ms frames)

After the editing of these parameters the signal parameters were used to generate excitation signal using GENSRC program. The speech signal was then synthesized using 'SYNTH' program. SYNTH program synthesizes speech signal using the source signal created by GENSRC and a cascaded formant network. (As per manual of SSL voice and speech system Bangalore)

Using the Vaghmi program the synthesized voice parameters were analyzed again to obtain frequency related parameters so that the synthesized voice could be compared with that of normal voice to note the similarities and differences between the two. After, comparison the data was subjected to statistical analysis using "EPISTAT" to obtain descriptive as well as inferential statistical information.

FIGURE 3.a : FLOW CHART SHOWING THE STEPS IN SYNTHESIS OF VOICE.



Experiment No.3:

Five experienced voice therapists carried out quality ratings of the synthesized voice, the original voice and normal voices recordings on a 7 point rating scale, where

0 - Very good
1 - gOOd
2 - normal
3 - near normal
4 - mild hoarse
5 - moderate voice.
5 - severe

Experiment No.4:

The final aspect of the experiment included use of the synthesized voice for therapy with dysphonics i.e., presenting synthesized voice auditorily as a model. The dysphonics were made to listen to their abnormal voice and the synthesized voice (which approximated normal voice).

For comparison and discrimination, using the synthesized voice, the dysphonics were asked to vary their voice to approximate the model voice.

Once they had discovered their new voice i.e., by varying pitch and other parameters they were made to phonate the same. This was stabilized and then carried out in extraclinical situations.

RESULTS AND DISCUSSION

The purpose of the study was to find out the feasibility to use synthesized voice as a model, during therapeutic intervention in case of dysphonics. For this purpose it was decided to determine the parameters in the dysphonic voice which are deviating from normal voice and to modify these parameters towards normal voice, using synthesis program and use it as a model for therapy.

Further, Listeners judgements were used to perceptually rate the quality of synthesized and unmodified voice of the dysphonics along with the normal voices.

EXTRACTION AND COMPARISON OF PARAMETERS OF VOICE OF NORMALS AND DYSPHONICS:-

Using the VAGHMI soft ware (INTON analysis), fundamental frequency and related parameters were extracted from the voice samples of dysphonics and the normals. A comparison of these parameters between normals and dysphonics i.e., for vowels /a/ /i/ and /u/ were made. The details of the parameters have been shown in the tables 50, 51, 53 for all three vowels.

	MEAN	MAX.	MIN.	RANGE	FLUC./SEC	EXT. OF FLUC.	MEAN	MAX.	MIN.	RANGE	FLUC/SEC.	EXT OF FLUC.
DYSPHONIC:												
MEAN	207.19	212.1	221.17	8.85	16.09	3.072	48.37	48.54	47.536	0.88	·69	1.92
SD	103.68	105.7	131.02	6.07	12.472	12.472 1.755	6.014	5.953	6.099	0.66	·83	·83
RANGE	98.73 - 338.25	100- 342.7	97.26 - 385.1	2.74- 14.99	·01- 31.37	·01 - 4.45	39 - 62.49	38.34 - 52.69	37.48 - 52.37	·01 - 2.03	·01 - 1.92	1.01 - 1.92
NORMAL:												
MEAN	190.7	199.21	194.4	5.484	2.19	2.52	51.85	52.66	50.92	1.73	·01	·01
SD	64.17	57.69	56.32	1.610	1.63	1.43	2.55	2.53	2.77	·496	·00	·00
RANGE	119.3- 245.19	151.53 - 249.4	119.8 - 235.01	3.27- 6.78	·01 - 4.02	·01 - 3.33	49.26- 55.07	50.76- 56.04	48.12- 54.31	1.09- 2.37	·01	·01

THE ABOVE TABLE SHOWS COMPARISON OF NORMALS VS DYSPHONICS FOR FREQUENCY & INTENSITY PARAMETERS FOR VOWEL /a/.

	MEAN	MAX.	MIN.	RANGE	FLUC/SEC	EXT. OF FLUC.	MEAN	MAX.	MIN.	RANGE	FLUC/SEC	EXT. OF FLUC.
DYSPHONIC:												
MEAN	173.84	178.77	164.95	13.73	6.59	7.67	40.21	47.77	45.88	1.934	1.928	1.078
SD	84	82.86	90.25	20.87	12.66	15.84	6.64	6.57	6.57	1.98	4.31	2.410
RANGE	106.99-301.2	108.8-304.5	105.44-298.24	2.34-51.03	0.01-35.94	0.01-35.9	38.25-54.2	40-52.69	35.85-53.78	0.76-5.46	0.01-1.96	.01-5.39
NORMAL:												
MEAN	205.46	213.05	205.00	8.05	0.01	2.34	50.72	51.86	50.04	2.296	0.01	0.01
SD	54.75	54.55	50.44	6.67	0.00	2.17	4.96	5.8	4.207	1.429	0.00	0.00
RANGE	125.3-247.8	136.27-256.8	131.95-245.03	2.41-19.09	0.01	0.01-4.18	46.8-59.32	47.28-61.9	46.29-57.27	0.99-4.73	0.01	0.01

THE ABOVE TABLE SHOWS COMPARISON OF NORMALS VS DYSPHONICS FOR FREQUENCY AND INTENSITY PARAMETERS FOR VOWEL /u/

	MEAN	MAX.	MIN.	RANGE	FLUC./SEC	EXT. OF FLUC.	MEAN	MAX.	MIN.	RANGE	FLUC./SEC	EXT. OF FLUC.
DYSPHONIC:												
MEAN	171.64	174.7	166.13	8.56	6.95	3.25	43.53	46.05	43.55	2.50	3.64	0.85
SD	88.27	88.84	89.65	6.07	4.44	2.05	2.78	6.39	4.44	1.18	7.112	1.318
RANGE	102.17 -304.9	104.58 -308.49	99.38 -300.7	5.2 -19.18	0.01 – 11.76.	0.01 -5.63	38.34 -52.69	38.86 -52.4	38.45 -50.73	1.15 -3.95	0.01 -3.95	0.01 -1.9
NORMAL : 20												
MEAN	202.72	161.48	157.78	5.104	2.09	2.45	48.33	49.93	48.66	1.27	0.01	0.01
SD	55.6	101.19	98.84	2.57	2.19	2.34	3.12	3.21	3.006	0.49	0.00	0.00
RANGE	100- 249.3	103.2 – 242.4	105– 245.03	3.2 - 8.8	0.01- 4.026	0.01 – 4.18	46.96 -54.55	47.5 -55.3	46.35 -53.63	0.9 -1.88	0.01 -1	0.01.

THE ABOVE TABLE SHOWS COMPARISON OF NORMALS VS DYSPHONICS FOR FREQUENCY AND INTENSITY PARAMETERS FOR VOWEL /i/.

SYNTHESIS OF THE DYSPHONIC VOICES:

Using the SSL program the dysphonic voices were analyzed and then synthesized for all the three vowels /a/ /i/ and /u/ of each subjects.

As the previous part, i.e., comparison of normals with dysphonics had shown significant difference in terms of fundamental frequency bandwidth of formant frequencies and variation in intensity. Hence these parameter were considered for modification. The following parameters were manipulated using synthesis program.

- Fundamental frequency
- Bandwidth
- Intensity

RESULTS:

In the FBAS sub-module, speech signal was analyzed to obtain the source and filter parameters at a uniform frame rate using analysis program. FOEDIT gave a graphic display of the various source.

Fo Edit was used to edit the source parameters like Fo, intensity, and voice source parameters open quotient, speed quotient and leak quotient. Editing operations interpolate and 'change' were used to edit the parameters i.e. to change towards normals.

Fundamental frequency was changed to the optimum frequency of the individual. Optimum frequency has been defined as the frequency of the vocal cord which elicits maximum resonance of the vocal tract.

Nataraja (1986) indicated that normals had fundamental frequency within +/- 40 Hz from the optimum frequency. However, the main objective of voice therapy is to provide the best possible voice to the subject. Therefore, providing optimum pitch would lead to the better voice, therefore the voice therapy aims at providing optimum pitch to the patient during therapy. Hence, the fundamental frequency was changed to the optimum frequency. The optimum frequencies of the dysphonics have been listed:

Cases		Fo	Optimum Frequency
Males	1	178 Hz	120 Hz
	2	131.15	126 Hz
	3	108.8	100 Hz
Females	1	291.2	210 Hz
	2	338.5	220 Hz.

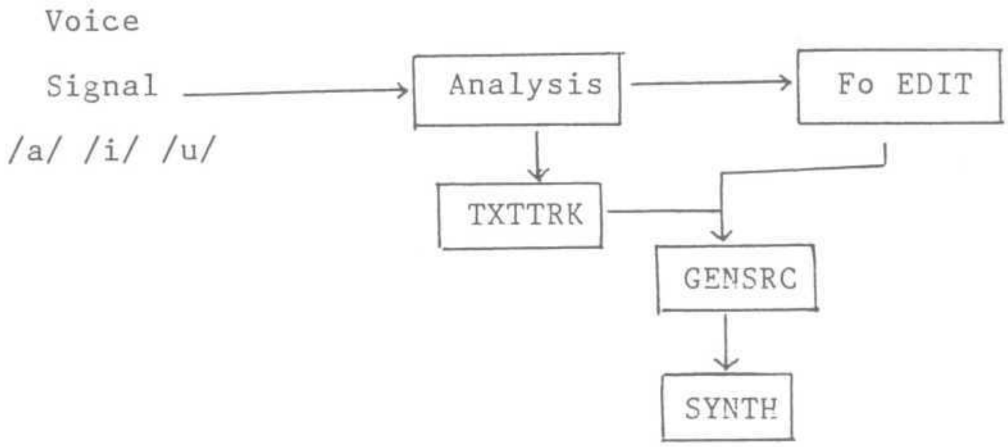
Using the same program, intensity parameters were also varied. The intensity of the signal was normalized throughout the signal. Finally the leak quotient, speed quotient and open quotient were normalized for each of the signal. This was essential to bring the voice towards normal.

The next stage of synthesis was use of TXTTRK program. This program was used to change, delete and interpolate bandwidth between two locations (points) (i.e. from 0 frames to 1000 ms frame).

Bandwidths of formant frequencies were found to play a major contributing factor in synthesis of voice. It was found that when bandwidths were not modified, the synthesized voices was not satisfactory. There fore the band width was varied and synthesized (modified).

The result was obtained in the form of synthesized voice, which was found to approximate normal voice, which was the major purpose of the study.

The modified - synthesized voice was compared with the original dysphonics and the normals.



ORGANIZATION OF FBAS MODULE

Twelve acoustic parameters were extracted using inton analysis, from the voices of five dysphonics their modified - synthesized voices and normals. The results of the extracted parameters are given below:

FUNDAMENTAL FREQUENCY IN PHONATION :

Fundamental frequency was measured for phonation in /a/, /i/ and /u/.

As per tables, 1, 2, 3, 4, 5 and 6 and graphs la, 1b, 1c, the comparison of the normals with the dysphonics showed varied results in /a/, /i/ and /u/. It was observed that there were differences in the fundamental frequency of normals and dysphonics, but, statistically, these differences were not significant; for /a/ and /i/. However, it was significant for /u/.

The fundamental frequency in phonation were different in the dysphonics and synthesized voice. These differences were found to be statistically significant, for the vowels /a/, /i/ and /u/.

A comparison of fundamental frequency between normals and synthesized voices showed very less variation. These differences were found to be not significant. This showed that synthesized voice approximated normals voice.

Group	Mean	SD.	Range
Normals	190.70	64.157	119 245.19
Dysphonic group	207.92	103.682	98.73 - 338.25
Synthesized voice group	153.79	65.37	101.96 - 229.72

Table-1 Mean, SD and range of Fo in phonation for /a/ in dysphonics, normals and synthesized voices.

Group	Mean	SD.	Range
Normals	202.72	55.635	100 - 249.3
Dysphonic group	171.64	88.27	102.17 - 304.9
Synthesized voice group	153.342	50.419	103.19 - 229.56

Table-2: Mean, SD, range of Fo in phonation for /i/ in dysphonics normals and synthesized voices.

Group	Mean	SD.	Range
Normals	205.462	54.748	125.3 - 247.85
Dysphonic group	173.838	84	106.99 - 301.17
Synthesized voice group	153.512	51.345	104.96 - 229.03

Table-3: Mean, SD, and range of Fo in phonation for /u/ in normals, synthesized voices and dysphonics.

Groups	r coefficient	Significance
Normal vs. dysphonics	.6	-
Normal vs. synthesized	.606	-
Dysphonics vs. synthesized	-.03	+

Table 4. Comparison of normals vs synthesized voice and normals vs. dysphonic voices, in terms of fundamental frequency in phonation of /a/.

Group	r	Significance
Normal vs dysphonics	.5151	-
Normal vs synthesized	.616	-
Dysphonics vs synthesized	.3	+

Table-5: Comparison of normals vs. synthesized voices and normals vs. dysphonics, synthesized vs. dysphonics, in terms of Fo in phonation of /i/.

Groups	r	Significance
Normal vs dysphonics	.77	-
Normal vs synthesized	.3	+
Dysphonics vs synthesized	.3	+

Table-6: Comparison of normals vs. synthesized voices and normals vs. dysphonics, synthesized voices vs. dysphonics, in terms of Fo in /u/.

MAXIMUM FUNDAMENTAL FREQUENCY :

A comparison between normals and dysphonics again showed no significant difference in maximum fundamental frequency of /a/, /i/ and/u/.

From the tables 7, 8, 9, 10 and Graphs la, 1b, and lc,it can be observed that there are great differences between the dysphonics and synthesized voices in terms of mean and standard deviation of maximum fundamental frequency. Statistically, the difference was found to be not significant. This indicates that these parameters are not significant to differentiate between synthesized and dysphonic voice.

Normals showed lesser variations, from the synthesized voice. The difference was found to be insignificant, statistically also. From this,it can be inferred, that the synthesized voice approximates normal voice, in terms of this parameter.

Group	Mean	SD.	Range
Normals	199.208	57.693	151.53 - 249.38
Dysphonic group	212.096	105.70	100 - 342.68
Synthesized voice group	156.46	66.627	103.9 - 235.3

Table-7: Mean, SD and range of maximum fundamental frequency in phonation of /a/ in normals, dysphonics and synthesized voices.

Group	Mean	SD.	Range
Normals	161.48	101.92	103.2 - 252.44
Dysphonic group	174.70	88.84	104.58 - 308.49
Synthesized voice group	153.34	54.32	105.22 - 235.29.

Table-8: Mean, SD and range of maximum fundamental frequency in phonation of /i/ in normals, dysphonics and synthesized voices.

Group	Mean	SD.	Range
Normals	213.05	54.559	136.27 - 256.86
Dysphonic group	178.77	82.865	108.84 - 304.51
Synthesized voice group	156.886	52.83	106.78 - 235.29

Table-9: Mean, SD and range of maximum fundamental frequency in phonation of /u/ in normals, dysphonics, and synthesized voices.

Groups	r			Significance		
	/a/	/i/	/u/	/a/	/i/	/u/
Normals vs. Synthesized	.70	.82	.71	-	-	-
Normal vs. dysphonics	.6	.51	.6	-	-	-
Synthesized dysphonics	.7	.9	.9	-	-	-

Table-10: Comparison of normals vs. synthesized voices normals vs. dysphonics and synthesized voices vs.dysphonics, in terms of maximum fundamental frequency in /a/,/i/ and /u/.

MINIMUM FUNDAMENTAL FREQUENCY IN PHONATION :

Minimum frequency in phonation, by the study of tables, 11, 12, 13 and 14, and graphs la, 1b, and 1c it was found that there are differences between synthesized and dysphonic voices. But, statistically, it was observed that there was no significant difference between synthesized voices and dysphonics. Though the mean, and standard deviation showed differences, statistically no significant difference was observed for these groups.

A comparison between normals and dysphonics also showed no significant difference. A comparison between normals and synthesized voices again showed no significant difference. From this, the inference can be drawn that synthesized voice approximates normal voice in terms of minimum frequency.

Group	Mean	SD.	Range
Normals	194.370	56.32	149.82 - 235.61
Dysphonic group	221.178	131.016	97.26 - 385.06
Synthesized voice group	153.272	63.76	100.95 - 227.25

Table-11: Mean, SD and range of minimum fundamental frequency in phonation of /a/ in normals, dysphonics and synthesized voices.

Group	Mean	SD.	Range
Normals	157.78	98.848	105 - 245.03
Dysphonic group	166.136	89.657	99.38 - 300.77
Synthesized voice group	148.74	52.035	103.87 - 226.84

Table-12: Mean, SD, range of minimum fundamental frequency in phonation of /i/ in normals, dysphonics and synthesized voices.

Group	Mean	SD.	Range
Normals	205.002	50.444	131.95 - 245.03
Dysphonic group	164.95	90.25	105.44 - 298.24
Synthesized voice group	152.13	51.034	104.24 - 230.01

Table-13: Mean, SD, range of minimum fundamental frequency in phonation of /u/ in normals, dysphonics and synthesized voices.

Groups	r			Significance		
	/a/	/i/	/u/	/a/	/i/	/u/
Normals vs.dysphonics	.6	.9	.5	–	–	–
Normals vs synthesized	.7	.92	.72	–	–	–
Synthesized vs dysphonics	.7	.9	.7	_ _ _		

Table-14: Comparison of normals vs. synthesized voices, normals vs. dysphonics, synthesized voices vs. dysphonics in terms of minimum fundamental frequency in /a/, /i/ and /u/.

RANGE OF FUNDAMENTAL FREQUENCY IN PHONATION

The frequency range in phonation was defined as the difference between maximum and minimum frequency in phonation.

Range values, in synthesized voices showed lesser variations, when compared to dysphonics, and normals. This can be made out from the study of tables 15 and 16 and graphs la, 1b and 1c It was observed that there was no significant difference between the three groups for /a/, /i/ and /u/.

Groups	Mean			SD			Range		
	/a/	/i/	/u/	/a/	/i/	/u/	/a/	/i/	/u/
Normals	5.89	5.10	8.05	1.61	2.51	6.67	3.27- 6.78	3.2 8.'8	2.41 19.09
Dys phonics	8.85	8.56	13.73	5.5	6.07	20.87	2.74 14.99	5.2 19.18	2.34 51.1
Syn.	2.81	4.59	4.09	0.17	0.87	0.99	1.73 5.28	1.35 8.46	2.54 5.28

Table-15: Mean, SD and ranges of range in phonation of /a/, /i/ and /u/ in normals dysphonics and synthesized voices.

Groups	r			Significance		
	/a/	/i/	/u/	/a/	/i/	/u/
Normals vs.dysphonics	.5	.9	.7	—	—	—
Normals vs synthesized	.8	.9	.9	-	-	-
Synthesized vs dysphonics	.7	.8	.8	---		

Table-16: Comparison of normals vs. synthesized voice, normals vs.dysphonics and dysphonics vs.synthesized voices in terms of fundamental frequency range of vowels /a/ /i/ /u/ in phonation.

EXTENT OF FLUCTUATIONS :

The extent of fluctuation in frequency in phonation was defined as the means of fluctuations in frequency in a phonation of one second. As per tables, 17, 18, 19 and 20, and graphs la, 1b and 1c, comparison between normals and dysphonics showed significant difference, on this parameter, in phonation of /a/, /i/ and /u/.

The results and statistical analysis of this parameter also showed that dysphonic voices and synthesized voices were significantly different. The means and SDs were high for the dysphonics than the normals.

Group	Mean	SD.	Range
Normals	2.522	1.432	.01 - 3.33
Dysphonic group	3.072	1.755	.01 - 4.45
Synthesized voice group	1.652	2.344	.01 - 3.26

Table-17: Mean, SD and range of extent of fluctuations in phonation of /a/ in normals, dysphonics and synthesized voices.

Group	Mean	SD.	Range
Normals	2.45	2.341	.01 - 4.18
Dysphonic group	3.25	2.05	.01 - 5.63
Synthesized voice group	1.582	2.258	.01 - 3.05

Table-18: Mean, SD and range of extent of fluctuations in phonation of /i/ in normals, dysphonics and synthesized voices.

Group	Mean	SD.	Range
Normals	2.38	2.177	.01 - 4.18
Dysphonic group	7.67	15.84	.01 - 35.94
Synthesized voice group	2.098	3.12	.01 - 3.48

Table-19: Mean, SD and range of extent of fluctuations in phonation of /u/ in normals, dysphonics and synthesized voiced.

Groups	r			Significance		
	/a/	/i/	/u/	/a/	/i/	/u/
Normals vs.dysphonics	.1	.3	.1	+	+	+
Normals vs synthesized	.6	.51	.72	-	-	-
Synthesized vs dysphonics	-.7	.1	.4	+	+	+

Table-20: Comparison of normals vs. synthesized voices, normals vs. dysphonics, synthesized voices vs. dysphonics, in term of extent of fluctuations for' /a/, /i/ and /u/.

Extent of fluctuation in frequency was considered as a major parameter which differntiate synthesized voice from dysphonic voices and thus contributing for abnormal voice.

SPEED OF FLUCTUATIONS IN FREQUENCY :

Speed of fluctuations in frequency can be defined as the . number of fluctuations in frequency in a phonation of one second.

Fluctuations, Speed of fluctuations in frequency was also considered as a major parameter which differentiate between synthesized voices and dysphonic voices.

From the study of tables 21, 22, 23 and 24, it was observed that there was significant difference between normals and dysphonic voices. This may be due to the irregular variations in fundamental frequency in case of dysphonics. This irregular variations, which is considered as a rough measure of jitter has been reported to be contributing for dysphonics (Nataraja.,1986).

A comparison between dysphonic voice and synthesized voice showed significant difference in terms of speed of fluctions for /a/, /i/ and /u/. Whereas, there was no significant diference between normals and synthesized voices.

Groups	Mean	S.D	Range
Normal	2.194	1.639	.01 - 4.02
Dys	16.09	12.472	.01 - 31.37
Syn.	1.784	2.472	.01

Table 21: Mean, SD, range of speed of fluctuations in phonation of /a/ in normals, dysphonics and synthesized voices.

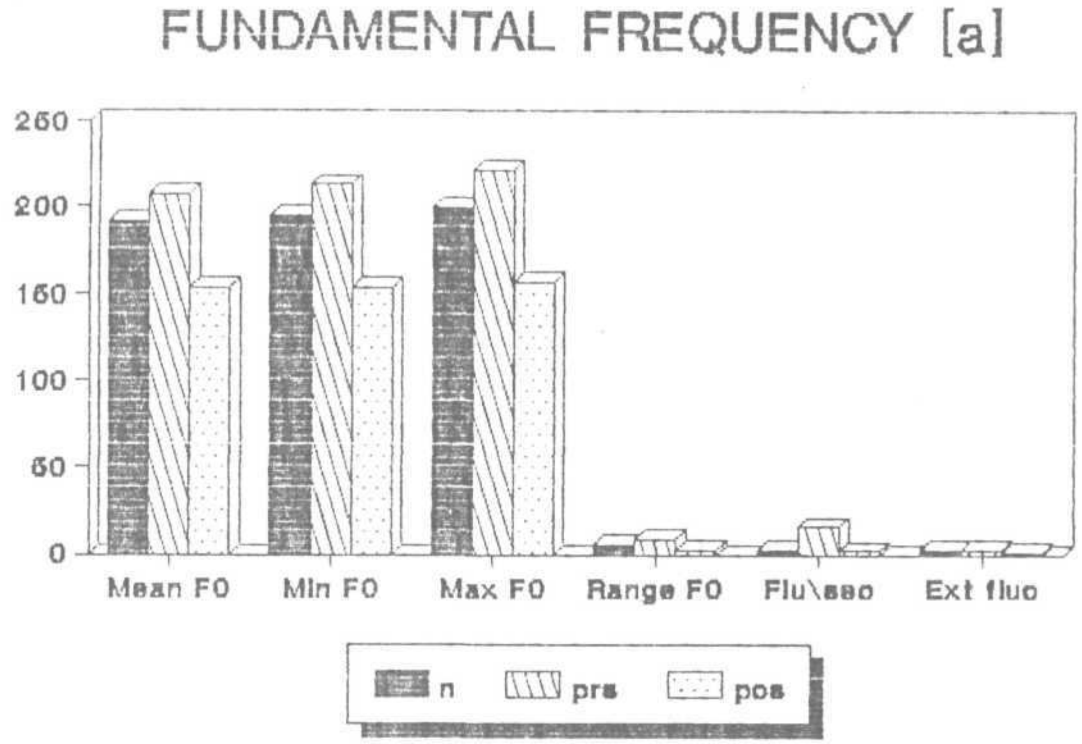
Groups	Mean	S.D	Range
Normal	2.096	2.19	.01 - 4.026
Dys	6.95	4.44	.01 - 11.76
Syn.	1.45	2.308	.01

Table 22: Mean, SD, range of speed of fluctuations in phonation of /i/ in normals, dysphonics and synthesized voices.

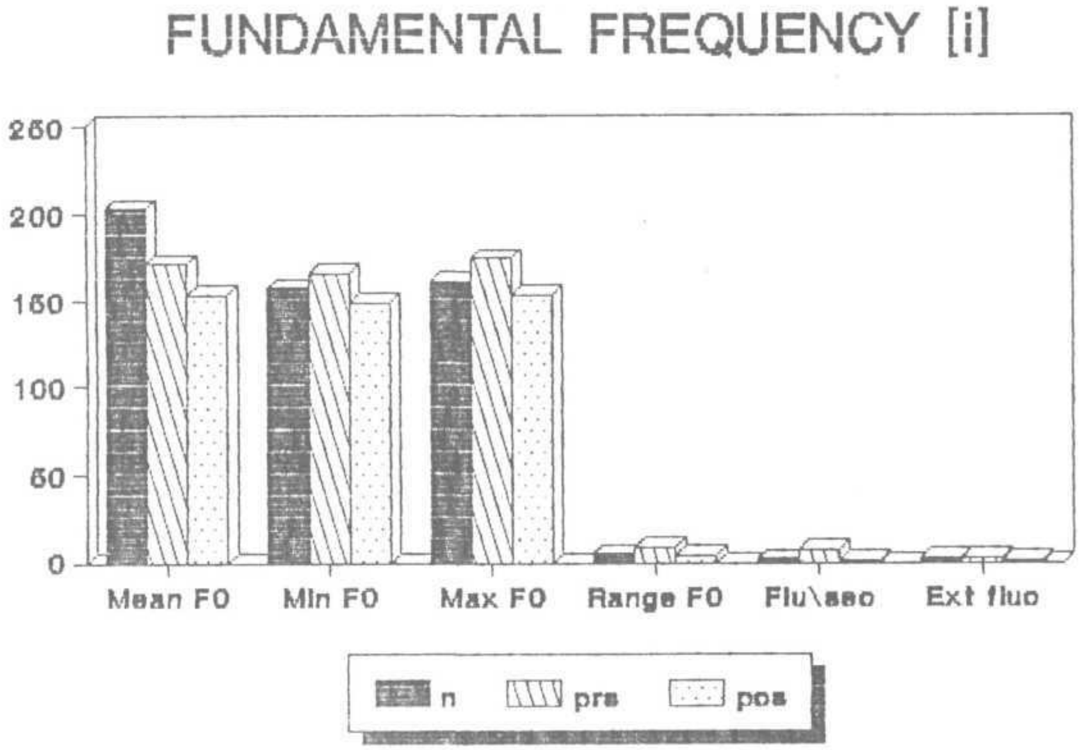
Groups	Mean	S.D	Range
Normal	.01	.01	.01
Dys	6.59	12.66	.01 - 35.94
Syn.	.84	1.72	.01

Group	r			Significance		
	/a/	/i/	/u/	/a/	/i/	/u/
N Vs dys	.1	.3	.1	+	+	+
N Vs syn.	.7	.6	.8	-	-	-
Syn Vs dys	.3	.5	.4	+	+	+

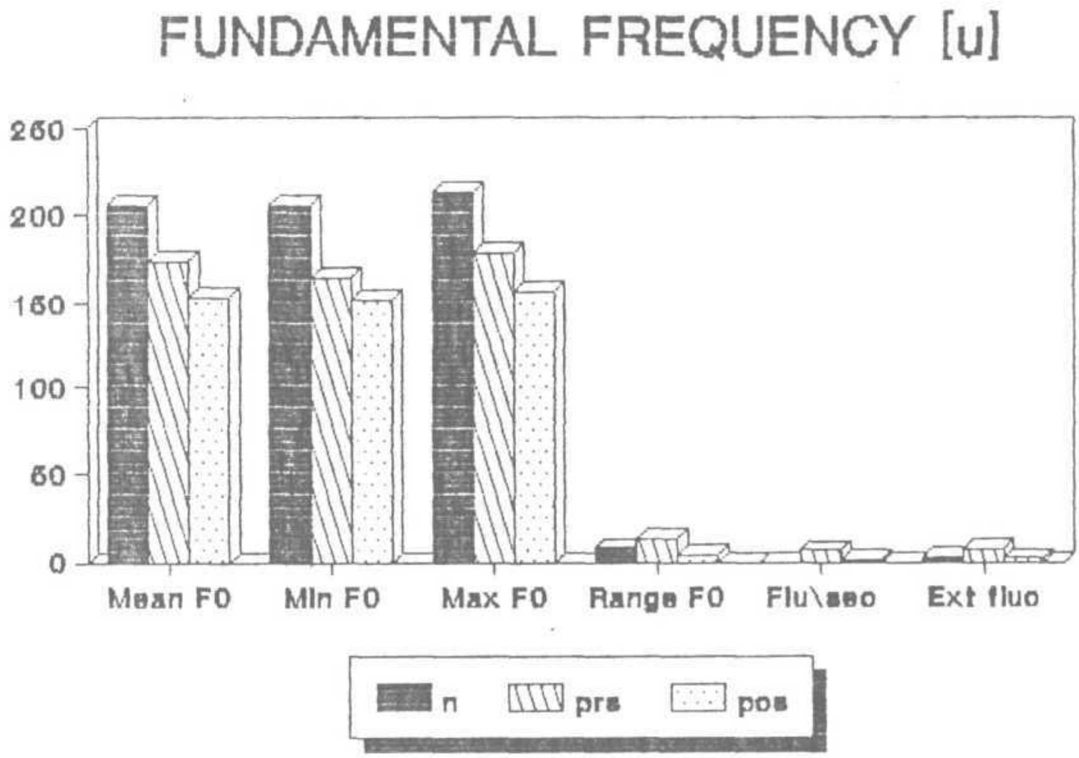
Table 24: Comparison of normals Vs synthesized voices and normals Vs dysphonics, synthesized voices Vs dysphonics, in terms of speed of fluctuations in /a/, /i/ and /u/.



GRAPHS !a. Means of the frequency parameters in normals, dysphonic and synthesized voices of phonation /a/.



GRAPH 1b. Means of the frequency parameters in normal, dysphonic and synthesized voices of phonation /i/.



GRAPH 1.c. Means of the frequency parameters in norr dysphonic and synthesized voices of phona -ion, /u/.

INTENSITY PARAMETERS

Intensity in phonation:

Mean Intensity in phonation was considered as the mean intensity of the steady portion of phonation.

From the study of tables 25, 26 and 27, 28 and graphs 2a, 2b and 2c it can be made out that there is no significant difference between synthesized voices and dysphonic voices, in phonation for /a/, /i/ & /u/. Comparison between normals and dysphonics, showed no significant difference in intensity. The same was true for the third group also i.e., when a comparison was made between normals Vs synthesized voices.

Groups	Mean	S.D	Range
Normal	51.85	2.557	49.21 - 55.07
Dys	48.374	6.014	38 - 52.49
Syn.	47.842	4.79	40.4 - 52.6

Table 25: Mean, SD range of intensity in phonation of /a/ for dysphonics and synthesized voices.

Groups	Mean	S.D	Range
Normals	49.332	3.124	46.96 - 54.55
dys	43.53	2.784	38.34 - 52.69
Syn.	49.704	5.001	40.58 - 47.72

Table 26: Mean, SD range of intensity in phonation of /i/ for dysphonics and synthesized voices.

Groups	Mean	S.D	Range
Normals	50.726	4.965	46.81 - 59.32
dys	46.202	6.642	38.25 - 54.18
Syn.	48.86	5.196	40.4 - 53.77

Table 27: Mean, SD range of intensity in phonation of /u/ for dysphonics and synthesized voices.

Group	r			Significance		
	/a/	/i/	/u/	/a/	/i/	/u/
N Vs dys	.8	.6	.51	- - -		
N Vs syn.	.9	.8	.6	-	- -	
Syn Vs dys	.7	.6	.8	-	- -	

Table 28: Comparison of normals Vs synthesized voices and normals Vs dysphonics, synthesized voices Vs dysphonics, in terms of mean intensity in /a/, /i/ and /u/.

MAXIMUM INTENSITY IN PHONATION:

The maximum intensity measured in a steady portion of phonation was considered as the maximum intensity.

The results obtained in the study in the normals, dysphonics and synthesized voice, are presented in tables 29, 30 & 31, 32, and graphs 2a, 2b, 2c.

The results, showed no significant difference between the dysphonics and synthesized voices. This was same when comparisons between normals and dysphonics and normals and synthesized voices were made i.e., and significant differences were found between the voice & normals & synthesized and normal & dysphonic.

Groups	Mean	S.D	Range
Normal	52.66	2.53	50.76 - 56.04
Dys	48.54	5.953	38.34 - 52.69
Syn.	48.18	4.779	40.58 - 52.94

Table 29: Mean, SD range of maximum intensity in phonation for /a/ in normals, dysphonics and synthesized voices.

Groups	Mean	S.D	Range
Normal	49.93	3.218	47.5 - 55.31
Dys	46.056	5.396	38.89 - 54.35
Syn.	50.246	4.712	41.62 - 54.23

Table 30 : Mean, SD range of maximum intensity in phonation for /i/ in normals, dysphonics and synthesized voices.

Groups	Mean	S.D	Range
Normal	. 51.896	5.808	47.28 - 61.99
Dys	47.774	6.57	40 - 54.69
Syn	49.966	5.300	42.82 - 54.73

Table 31 : Mean, SD range of maximum intensity in phonation for /u/ in normals, dysphonics and synthesized voices.

Groups	r			Significance		
	/a/	/i/	/u/	/a/	/i/	/u/
N Vs dys	.6	.6	.7	-	-	-
N Vs syn.	.6	.7	.6	-	-	-
Syn Vs dys	.9	.6	.9	-	-	-

Table 32: Comparison of normals Vs dysphonics normals Vs synthesized, synthesized Vs dysphonics in /a/, /i/ & /u/.

Therefore, maximum intensity was treated as a non-significant parameter in dysphonia in the present series of cases.

MINIMUM INTENSITY IN PHONATION:

The minimum intensity measured in a steady portion of phonation, was considered as the minimum intensity.

The results obtained in the study in the normals, dysphonics and synthesised voices, are presented in tables 33, 34, & 35, 36 and graphs 2a, 2b 2c.

The results of the statistical analysis significant difference between the dysphonics and the synthesized voices. The means standard deviation of synthesized voice showed lesser variations than the dysphonics hence statistically, no significant difference between the dysphonics and synthesized voices, were observed.

The same was true for the comparison normals & dysphonics and normals between thus the minimum intensity not a significant parameter for the purpose and synthesized voices. Thus the minimum intensity is not a significant parameter for the purpose of synthesis of voice.

Groups	Mean	S.D	Range
Normal	50.92	2.77	48.12 - 54.31
Dys	47.536	6.099	37.48 - 52.37
Syn.	47.38	5.024	40.11 - 52.2

Table 33 : Mean, SD range of maximum intensity in phonation of /a/ for normals, dysphonics and synthesized voices.

Groups	Mean	S.D	Range
Normal	48.66	3.006	46.35 - 53.63
Dys	43.548	4.448	38.45 - 50.73
Syn.	49.188	5.275	40.32 - 54.3

Table 34: Mean, SD range of maximum intensity in phonation of /i/ for normals, dysphonics and synthesized voices.

Groups	Mean	S.D	Range
Normal	50.004	4.207	46.29 - 57.27
Dys	45.886	6.578	37.85 - 53.78
Syn.	48.302	5.67	38.89 - 53.44

Table 35 : Mean, SD range of maximum intensity in phonation of /u/ for normals, dysphonics and synthesized voices.

Groups	r			Significance		
	/a/	/i/	/u/	/a/	/i/	/u/
N Vs dys	.9	.7	.6	-	-	-
N Vs syn.	.9	.8	.6	-	-	-
Syn Vs dys	.9	.6	.6	-	-	-

Table 36: Comparison of normals Vs dysphonics normals Vs synthesized voices and synthesized Vs dysphonics in terms of minimum intensity in /a/, /i/ & /u/.

INTENSITY RANGE IN PHONATION:

Groups	Mean	S.D	Range
Normal	1.738	0.496	1.09 - 2.27
Dys	0.888	0.66	.01 - 2.03
Syn.	1.162	0.56	1.16 - 2.03

Table 37: Mean, SD range of maximum intensity range in phonation /a/ for all three groups.

Groups	Mean	S.D	Range
Normal	1.272	0.493	.9 - 1.88
Dys	2.504	1.118	1.55 - 3.95
Syn.	1.05	0.823	.43 - 2.5

Table 38 : Mean, SD range of maximum intensity range in phonation of /i/ for all three groups.

Groups	Mean	S.D	Range
Normal	2.296	1.429	.99 - 4.73
Dys	1.934	1.980	.76 - 5.46
Syn.	1.66	1.494	.79 - 1.29

Table 39: Mean, SD range of maximum intensity range in phonation of /u/ for all three groups.

Groups	r			Significance		
	/a/	/i/	/u/	/a/	/i/	/u/
N Vs dys	.1	.3.	.4	+	+	+
N Vs syn.	.7	.6	.8	-	-	-
Syn Vs dys	0	-.515	-.6	+	+	+

Table 40: Comparison of normals Vs dysphonics normals Vs synthesized voices and synthesized Vs dysphonics in terms of intensity range in phonation of /a/, /i/ & /u/.

The results obtained in the present study in the normals, dysphonics and synthesized voice groups are presented in tables 36, 37, 38, 39 & 40 and graphs 2a, 2b 2c.

The results showed statistically significant difference between synthesised voice groups and dysphonics. the dysphonics showed inability to maintain the intensity at a steady level like normals. Hence, a statistically significant difference was observed between normals and dysphonic groups.

A comparison between normals and synthesised voices showed no significant difference in intensity range for phonation of /a/, /i/ & /u/. Therefore it can be stated that the dysphonic voice had approximated the normal voice after synthesis in terms of intensity range.

EXTENT OF FLUCTUATIONS IN INTENSITY:

The extent of fluctuations in intensity has been considered as indicating the regualrity of uibration of the vocal cords.

Groups	Mean	S.D	Range
Normal	.01	.00	.01 -
Dys	1.92	.83	.01 - 1.92
Syn.	.01	.00	.01

Table 41: Mean, SD range of extent of fluctuations in intensity for phonation for /a/ for 3 groups.

Groups	Mean	S.D	Range
Normal	.01	.00	.01
Dys	.85	1.318	.01 - 1.9
Syn.	.01	.00	.01

Table 42 : Mean, SD range of extent of fluctuations in intensity for phonation for /i/ for 3 groups.

Groups	Mean	S.D	Range
Normal	.01	.00	.01
Dys	1.078	2.410	.01 - 5.39
Syn.	.01	.00	.01

Table 43 : Mean, SD range of extent of fluctuations in intensity for phonation for /a/ for 3 groups.

Groups	r			Significance		
	/a/	/i/	/u/	/a/	/i/	/u/
N Vs dys	.1	.2	-.3	+	+	+
N Vs syn.	.7	.9-	.8	-	- -	
Syn Vs dys	.1	.3	0	+	+	+

Table 44: Comparison of normals Vs dysphonics normals Vs synthesized voices and synthesized Vs dysphonics in terms of extent of fluctuation in intensity in phonation of /a/, /i/ & /u/.

The results, regarding extent of fluctuations in terms of intensity in phonation have been tabulated in table 41, 42, 43 and 44 and graphs 2a, 2b 2c. A significant difference was oberved between normals and dysphonics. It also revealed that there was significant difference between synthesized voice and normals, in terms of extent of fluctuations in intensity in phonation. Further no significant difference was observed for synthesized voices & normals.

SPEED OF FLUCTUATIONS:

This parameter along with the other parameter extent of fluctuations in intensity has been considered to provide information about the condition and functioning of vocal cords. Tables 45, 46, 47 and 48 and graphs 2a, 2b and 2c showed the results obtained for normals, dysphonics and synthesized voices.

Similar to the extent of fluctuation in intensity, the speed of fluctuation in intensity showed significant difference between the dysphonics and the synthesized voices. Further normals showed significantly different values than the values for the dysphonic group. The comparison between the normals and the syntesized voice groups showed no significant difference, statistically.

Groups	Mean	S.D	Range
Normal	.01	.01	.01
Dys	.69	.83	.01 - 1.92
Syn.	.01	.00	.01

Table 45 : Mean, SD range of speed of fluctuations in intensity for /a/ for 3 groups.

Groups	Mean	S.D	Range
Normal	.01	.00	.01
Dys	3.648	7.112	.01 - 3.95
Syn.	.01	.00	.01

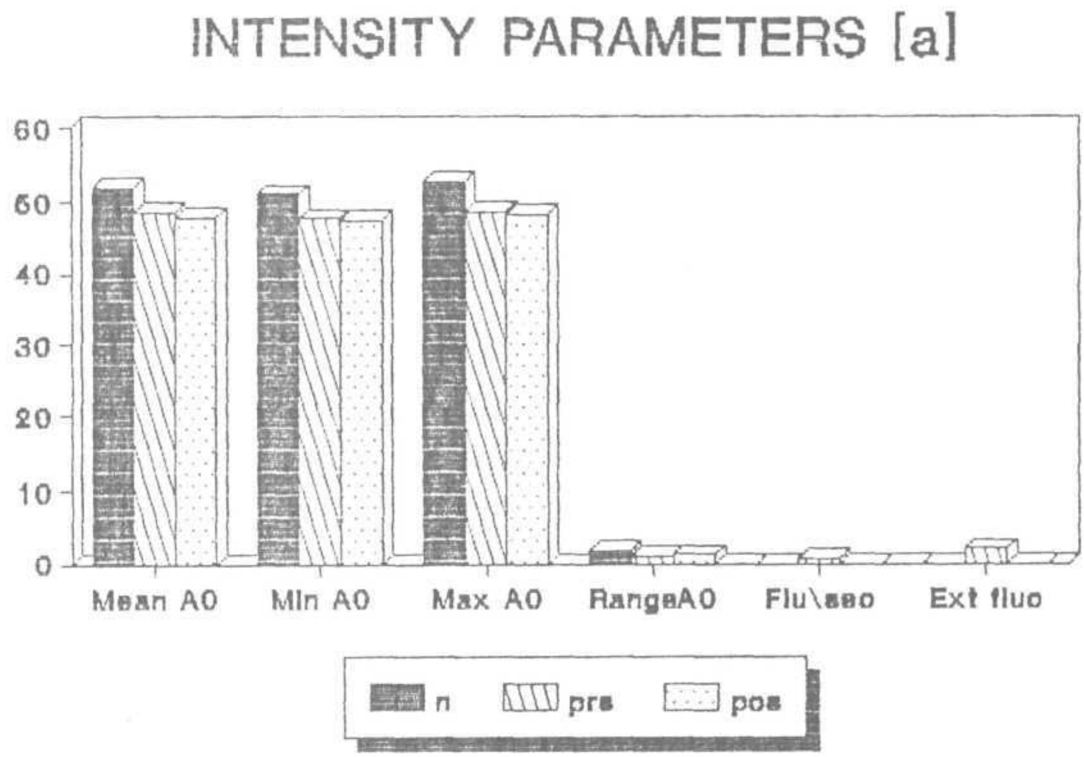
Table 46 : Mean, SD range of speed of fluctuations in intensity for /i/ for 3 groups.

Groups	Mean	S.D	Range
Normal	.01	.00	.01
Dys	1.928	4.311	.01 - 1.96
Syn.	.01	.00	.01

Table 47 : Mean, SD range of speed of fluctuations in intensity for /a/ for 3 groups.

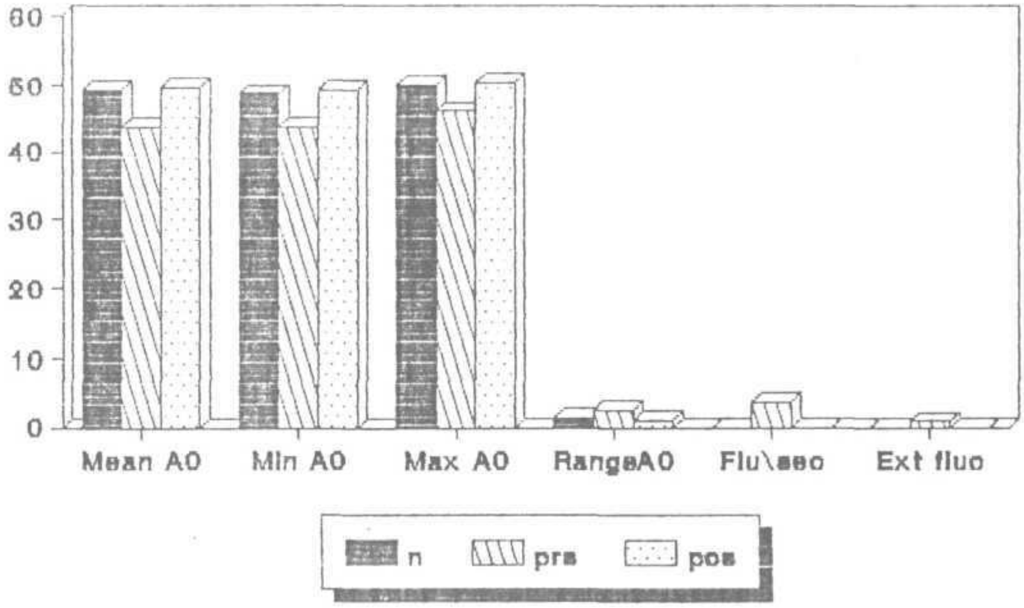
Groups	r			Significance		
	/a/	/i/	/u/	/a/	/i/	/u/
N Vs dys	-.2	.1	0	+	+	+
N Vs syn.	.9	.7	.8	-	-	-
Syn Vs dys	.1	.2	-.3	+	+	+

Table 48: Comparison of normals Vs dysphonics normals Vs synthesized voices and synthesized Vs dysphonics in terms of speed of fluctuation in intensity in phonation of /a/, /i/ & /u/.



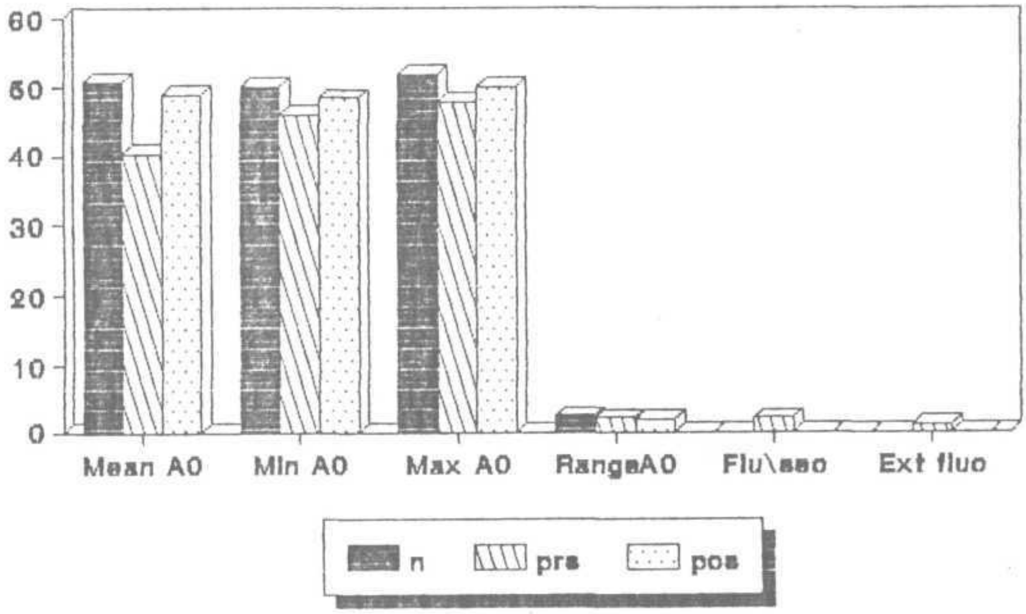
GRAPH 2a. Means of the intensity parameters in normals, dysphonic and synthesized voices of phonation /a/.

INTENSITY PARAMETERS [i]



GRAPH 2b. Means of the intensity parameters in normals dysphonic and synthesized voices of phonation /i/.

INTENSITY PARAMETERS [u]



GRAPH 2c. Means of the intensity parameters in normals dysphonic and synthesized voices of phonation /u/.

The results of the comparisons show that the extent of fluctuations and speed of fluctuations in intensity, which are considered to be rough measures of shimmer, have been reported to contribute to the poor voice quality of the dysphonic. A correction of the parameter significantly improve the voice. The parameters which differentiated synthesized and dysphonics and between normals & dysphonics are listed as follows:

a) Mean fundamental frequency
b) Extent of fluctuations in frequency
c) Speed of fluctuations in frequency
d) Intensity range
e) Extent of fluctuations in intensity
f) Speed of fluctuations in intensity.

The hypothesis that there is no significant difference between normals and dysphonics can be rejected. Similarly the hypothesis that there is no significant difference between synthesized voice and dysphonic voice can be rejected.

Normal and synthesized voices showed no significant difference in all 12 parameters. Hence, the hypothesis that there is no significant difference between normals and synthesised voice was accepted.

QUALITY JUDGEMENTS:

Five experienced voice therapists carried out quality ratings of the normal synthesized and dysphonic voice recordings. 15 normal, 15 synthesised voice samples and 15 dysphonics voices (/a/, /i/ and /u/) were presented randomly to the voices therapists. They were asked to rate on 7 - point rating scale, which was as follows:

6 - severe
5 - moderate abnormal
4 - mild
3 - near normal
2 - normal
1 - good
0 - very good.

The dysphonics were grouped under severe, moderate and mild hoarse voices. The synthesized voices were taken as the rest of voices, voice therapists were instructed to identify and rate the dysphonic voices, normal and synthesized voices.

Results have been shown in table 49. This table shows the rating for each voice, which were presented. The rating were taken into account, when 3 or more voice therapists rated the same.

	DYSPHONICS /a/	/i/	/u/	SYNTHESIZED VOICE /a/	/i/	/u/
	6	5	4	2	3	3
	4	4	7	3	3	3
RATING	5	4	4	3	2	3
GIVEN	4	5	5	3	2	3
	4	4	5	3	3	2

From the above table 49, the quality rating can be made. (3 out of 5 voice therapists rating were taken as final scale). Under three main categories i.e., severe, moderate and mild dysphonia. These dysphonics included both males and females.

It can be noted that the subjective evaluation of these dysphonics correlated with that of the objective evaluation.

The synthesized voices of these dysphonics were rated mainly under "near normal" and "normal" i.e. in the 7- point rating scale, synthesized voices were given a rating of "3" and "2". Only one particular synthesized voice i.e. /i/ of a particular case was rated as being i.e., "mild" abnormality.

All normals were rated as either 2 or 1 i.e. 'normal' or 'good'. From the above results, it can be shown that the synthesized voice has approximated normal voice, as rated subjectively. Thus the synthesized voices of the dysphonics was good enough to therapeutic intervention.

The hypothesis that the synthesized voice of dysphonics, without using the characteristics of voice which were not altered in process of synthesis, could be used as a model for voice therapy for dysphonics, was accepted as the synthesized voice was found to be normal, near normal, by the judges on the perceptual evaluation.

SYNTHESIZED VOICE AS A MODEL IN VOICE THERAPY

All the five dysphonics, were advised voice therapy at AIISH clinic.

They were motivated towards therapeutic intervention by exposing them to the latest software packages available for voice and emphasising the importance of the programmes. As described earlier their dysphonic voices were synthesized, by manipulating various parameters. Using this as basis, further therapeutic intervention was carried out, systematically.

i. The dysphonics were presented with the synthesized voice of that particular subject (their own voice), auditorily using the software program.

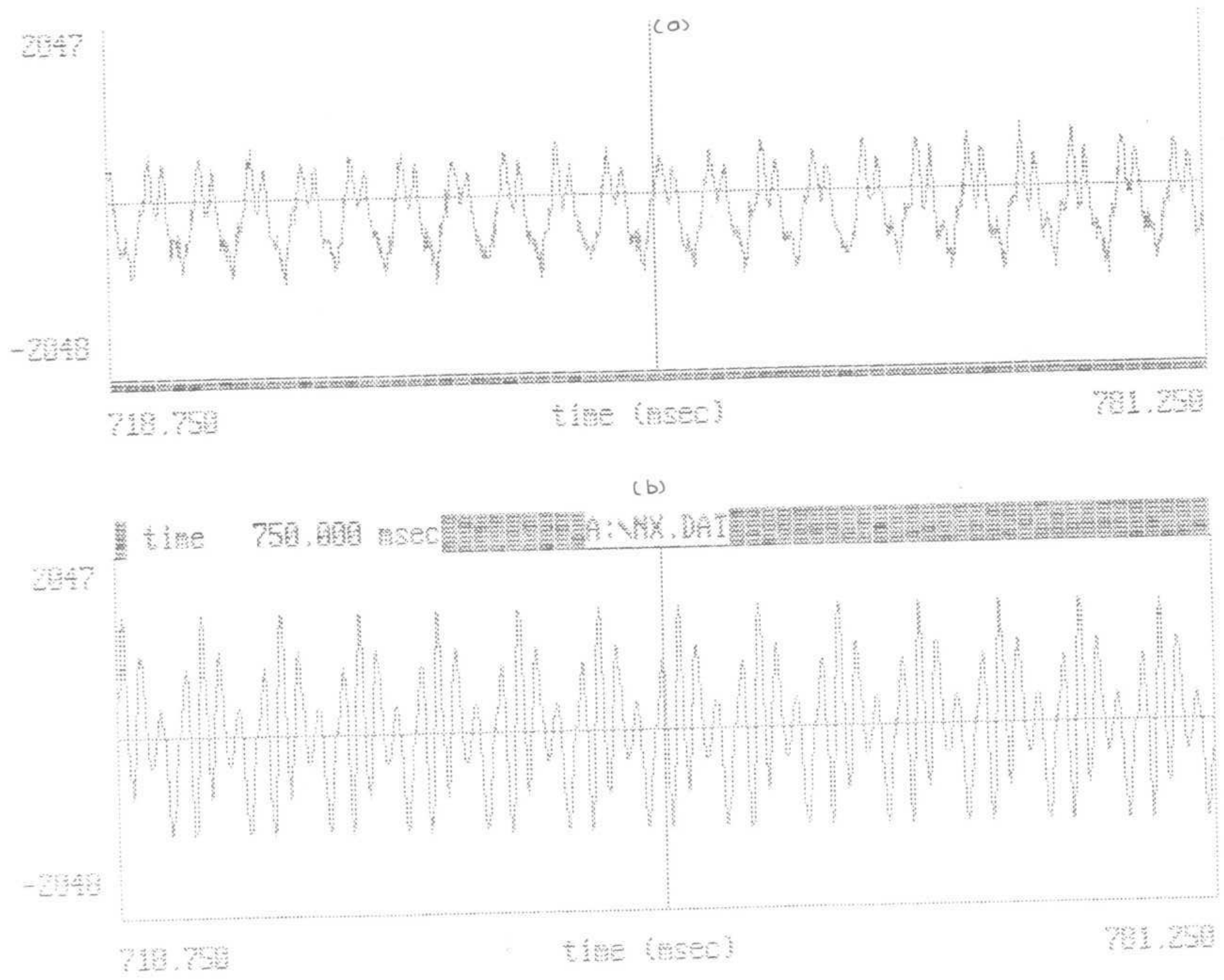
ii. They were made to identify the differences in synthesized voice, by matching their own voice, and also made to discriminate their voice from the synthesized voice.

iii. The dysphonics were motivated to vary the voice to approximate the model provided i.e. synthesized of that particular subject.

iv. Once the therapist was sure that the subject had approximated the synthesized voice, the final step in therapy was carried out i.e. stabilization of the voice established through this approach.

This was used with all the 5 dysphonics, during the initial stages five of them underwent this kind of approach.

v. All the dysphonics showed improvement in their voice i.e. they were able to discover a new voice on the basis of the synthesized voice. The dysphonics also showed motivation, in this kind of approach which there by facilitated his improvement in his voice.



GRAPH 3.b : WAVEFORM DISPLAY OF DYSPHONIC VOICE (a) AND SYNTHESIZED VOICE (b) in phonation, /a/ (of a particular subject)

SUMMARY & CONCLUSIONS

This study aimed to synthesize the voices of dysphonics using synthesis program, to create voice which approximated normal voice and would serve as a model during therapeutic intervention. Five dysphonics (3 males & 2 females) were taken for the study. 5 normals were also taken as controls.

Using the SSL program, the dysphonic voices were synthesized. Three major parameters i.e. Fundamental frequency, Bandwidths and intensity were manipulated. The synthesized voice was found to approximate normal voice subjectively. Subjective evaluation was done by 5 experienced voice therapists, on a 7 - point rating scale.

Objectively, 12 parameters were acquired, analysed and measured for the dysphonic voice, synthesized voice and normals. The parameters measured were :

1) Frequency parameters:

- Mean fundamental frequency (Hz)
- Maximum (Hz)
- Minimum (Hz)
- Range (Hz)
- Fluctuations/sec
- Extent of fluctuations.

Intensity parameters:

- Mean AO (dB).
- Maximum AO (dB).
- Minimum AO (dB).
- Range AO (dB).
- Fluctuations/sec
- Extent of fluctuations.

The results obtained were then subjected to statistical analysis using rank correlation coefficient in the "Epistat" program. The following conclusions have been drawn, from the statistical analysis:

In the mean fundamental freq. (Hz), extent of fluctuations in frequency, speed of fluctuations in frequency, intensity range, extent of fluctuations in intensity, speed of fluctuations in intensity, significant difference was observed for the synthesized and the dysphonic groups. This was same for the dysphonic group and the normal groups.

Significant difference was not observed for the synthesized voice and normals. This statistical results correlated with the subjective results.

Synthesised and dysphonic groups showed no significant difference for the maximum fundamental frequency, minimum fundamental frequency, frequency range in phonation, mean intensity in phonation, maximum intensity and minimum intensity in phonation.

As a part of this study, synthesized voice were presented to the dysphonics, auditorily, during therapeutic intervention. The dysphonics matched their voices with the synthesized voice and varied their fundamental frequency. In this manner, they discovered a new voice, which they approximated and stabilised.

IMPLICATIONS:

1) This approach of voice synthesis can be used to create models for the dysphonics during therapeutic intervention.

2) Using synthesized voice, the dysphonics listens to their own voice as normal voice; self hearing can show better improvement in their voices.

RECOMMENDATIONS FOR FURTHER STUDY:

1) These parameters may be studied with different groups i.e., (male & females separately) for particular age groups.

2) More number of parameters can be included to study the difference between the synthesized voice and the dysphonic voices.

3) Pre and Post therapy, using synthesized voice program can be studied.

BIBLIOGRAPHY

ANDERSON, V. (1961) : Training the speaking voice, Oxford University, New York.

ARNOLD, G.E. (1955) : Vocal rehabilitation of paralytic dysphonia II, acoustic analysis of vocal fold function analysis of vocal fold function; Archives, Otolargngology, 62 (593-601).

AITKINSON, J. (1978) : Correlation analysis of the physiologyical factors controlling fundamental voice frequency, Journal of acoustic society of America 63, (211-222).

BAKEN, R.J. (1990) : Consideration of the relationship between fundamental frequency of phonation and vocal jitter, Folia phoniatrica vol. 42, Pg(31).

BERRY, R.J., EPSTEIN, R., FOURCIN, A.J., FREEMAN, M., MAC CURTAIN, F. and NOSCOE N.J. (1982) : An objective analysis of voice disorder part I & II : British journal of communication disorders, 17 (67-83).

BOOHE, D.E. (1993) : The voice and voice therapy ( ) prentice Hall Inc. Eagle wood Clifs, N.J.

BORDEN, G.J. and HARRIS, K.S. (ED) 1980 : speech science primer : Physicology, acoustics, and perception of speech. The Williams and wilkins company, Baltimore.

BOWLER, N.W. (1964) : A fundamental frequency analysis of harsh vocal quality, speech monograph, 31 (128-134).

BECKETT, I.P. (1971) : Parameters of voice quality in TRAVIS, L.E. (1971) hand book of speech pathology & Audiology, Appletoncentury - CROFTS, N.Y.

BRIESS (1959) : Voice therapy : Part 2 : Essential treatment phases of specific laryngeal muscle dysfunction Arch, of otolaryngology, 69 (61-69).

BRODNITZ, F.S. (1995) : Vocal rehabilitation, rochesler, Minnesota whiting press.

BUNCH, M. (1982) : Dynamics of singing voice, springer-VERLAG, N.Y.

CARLSON, B.G. (1991) : Experiments with voice modelling in speech synthesis, speech commonication, vol:10 Pg:481.

DAVIS, S.B. (1976) : Computer evaluation of laryngeal pathology based on inverse filtering of speech. SCRL monograph, 13, California.

DAVIS. (1979) : "Acoustic characteristics of normal and Pathologyical voice" in LASS, N.J. (ED) - speech and language advances in basic reserach and practice, vol.1. Academic Press, N.Y.

DRAKE, O. (1937) : Towards an improved vocal quality, Quartely journal of speech, 23 (620-626).

DREHER, J. and BRAGG, V. (1953) : Evaluation of voice normality speech monograph, 20 (74-78).

EMANUEL, F.W. & WHITE HEAD R.L. (1979) : Harmonic levies and vowel roughness. JSHR, 22 (4) (829-840).

ESKENAGI, L., CHILDRES, D.G. and HICKS, D.N. (1990) : Acoustic corretales of voice quality, JSHR, vol.33, (298).

FAIRBANKS, G. (1970) : Recent experimental investigations of vocal experimental speech. JASA, 11 (457-466).

FANT, G. (1970) : Acoustic theory of speech prodution : (2nd ed) Mouton, the Haque.

FISHER, H.B. (1966) : Improving voice and articulation, Hougho Miffin co, BOSTON, N.Y.

FITCH, J.L. & HOLBROOK, A. (1970) : Modal vocal fundamental frequency of young adults. Arch otolaryngology, 1970, 92 (379-382).

FLANAGAN, J.L. (1965) : Speech analysis synthesis and perception Berlin - Heidelberg - N.Y. springer.

FLANAGAN (1972) : Speech analysis perception and synthesis (2nd edn) springer - VERLAG, N.Y.

FROESCHELS (1960): Remarks on some pathologic and physiologic conditions of voice. Archives of otolaryngology, 71, (787-788.

FOX, D. (1978) : Evaluation of voice problems. INSING, S and LYNCH, J. (EDS) : Diagnostic procedurs in hearing, speech and language, Baltimore, University, Park press.

GEETHA, P. (1974) : "Relation between fundamental frequency and resonant frequencies of vocal tract". Master's Dissertation, University of Mysore.

GEORGE, S. (1973(: A study of fundamental frequency of voice and natural frequency of vocal tract on an ranges. Master's Dissertation, University of Mysore.

GORDON, M.T., MORTON. P.M. and SIMPSON. I.C. (1960): Air flow measurements in diagnosis assessement and treatment of mechanical dysphonia. Folia Phoriat., 30, (161-174).

GOULD, W.J. (1975): Quantitative assessment of voice function in microlaryngology. Folia Phoniat., 27(3). (190-200).

GRY, G.W., and WISE. C.M. (1959): "The Bases of Speech". Harper & Row Publishers, N.Y.

GREEN D.S. (1956): The fundamental frequency characteristics of the speech of profoundly deaf individuals. Doctoral Dissertation Purdue Univ.

GREEN. M.C.L. (1964): The voice and its disorders. Mitman Medical. London.

HAMMARBERG, B. (1980): "Perceptual and Acoustic Correlates of abnormal voice qualities". Acta Otolaryngol., 90, (441-451).

HANSON. R.J. (1980): A two state model of Fo control. J.A.S.A., 64, (534-544).

HIBERGER. V.L. & HORII. Y. (1982): "Jitter and Shimmer in sustained Phonation" in Lass. N.J., (ed)Speech and Language. Vol.7 Academic Press, New York.

HIKI, S., IMAISUMI, S., KORANO, M., MATSUSHITA, H., & KAKITA, Y. (1975 A): Acoustical analysis for voice disorders: An attempt towards clinical application. 'J.A.S.A., 58, (111-).

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

HIRANO, M., OHALA, J., & VENNARD, W. (1969) : The function of larnygeal muscles in regulating fundamental frequency and intensity of phonatior. JSHR., 12, (616-628).

HIXON, T.J., (1973) : Respiratory function in Speech. In MINIFIE., Hixon, T., and WILLIAMS, F. (Eds) : Normal aspects of speech, hearing and language. Englewood cliffs, N.J., Prentice-Hall Inc., (73-125).

HOLLIEN, H., MICHEL, J., & DOHERTY, E.T. (1973) : A method for analysing vocal jitter in sustained phonation, J. of Phonetics, 1, (85-91).

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

HORII, Y. (1980) : "Vocal shimmer in sustained phonation". JSHR. 23(1), (202-209).

IMAIZUMI, S., KIKI, S., HIRANO, M., & MASUSHITA, H. (1980) : Analysis of pathologicla voices with a sound spectrograph. J. Account. Soc. Jpn. 36, (9-16).

ISSHIKI, N. (1965) : Vocal intensity and airflow rate. Folia Phoniat.,

__________, (1964) : Regulatory mechanisms of voice intensity variation. JSHR., 7(1), (17-29).

__________, & RINGEL. R.L., (1964) : Airflow during the production of selected consonants. J. of Speech Hear., 7: (233-244).

__________, & VON LEDEN, H. (1964) : Hoarseness : Aerodynamic studies. Archieves of otolarynogology, 80, (206-213).

__________,__________, TRANABE, M., MORIMOTO, M. (1969) : Differential diagnosis of horseness. Folia Phoniat., 21, (9-19).

JAYARAMA, K., (1975) : An attempt at Differential Diagnosis of Dysphonia. Master's Dissertation, Univ, of Mysore.

______________, GORDON, M.T., MORTON, F.M., & SIMPSON, I.C., (1981) : "Comparison of Methods of assessing vocal function". Folia Phoniat 33(1), 951-65).

______________, & WENDAHL, R., (1971) : "Correlates of voice production" in TRAVIS, L.E. (ed) Hand book of Speech Pathology and Audiology, Prentice - hall, Inc., Englewood cliffs, N.J. (465-480).

KOUFMAN and ISAACSON (1991) Cited by Joseph C.S. Stemple "Voice therapy: clinical studies" Mozby - year Book Inc, St.Louis, 1993.

KLATT. D., KLATT, L.C., (1990) Analysis synthesis and perception of voice quality variations among female and male talkers. Journal of acoustical society of America, (850-854) 1990.

LAUKKANEN ANNE - MARIA, (1994) Artificial pitch changing in auditory feedback as a possible method in voice training and therapy, Folia phoniatrica, 48 (86-96) 1994.

MURPHY, A.J. (1964) : Functional Vocie Disorders. Englewood cliffs Prenntice-Hall, Inc., N.J.

MURY, T., (1978) : "Speaking fundamental frequency characteristics associated with voice pathologies, Jr. Sp. Hg. Dis., 43(3).

NATARAJA, N.P. (1972) : "Objective method of Locating optimum pitch". Master's Dissertation Univ. of Mysore.

NATARAJA, N.P. (l984(a)) : Phonation duration and optimum frequency. J.A.I.I.S.H., 15, (117-122).

NATARAJA, N.P. (1984(b)) : Mean air flow rate and optimum frequency. JAIISH, 15, (133-136).

NATARAJA, N.P., & JAGADISH, A. (1984) : "Vowel Duration & Fundamental Frequency". J.A.I.I.S.H., 15,( ).

NATARAJA, N.P., & JAYARAM, M. (1982) : A new approach to the classification of voice disorders. J. of All India Institute of Sp.

NATARAJA, N.P, and SAVITHRI, S.R. (1990) cited by Rashmi, M. "Acoustic aspects of the speech of children" unpublished Master's dissertation submitted to University of Mysore. 1995.

NATARAJA N.P. , (1986): Differential diagnosis of dysphonias, Doctoral thesis, University of Mysore, 1986.

PANNSACKER, M. (1984) : Classification of voice disorders: A review of Literature. Lang. Sp. and Hg. Services in Schools, 15, (169-174).

SHOUP, J.E., LASS, N.J., & KUEHN, D.P. (1982) : "Acoustics of Speech", in LASS, N.J. et. al. (ed) "Speech, Language and Hearing. Vol.1, W.B.Sannders Company, Philadelphia.

SORENSEN, D., & HORII, Y. (1984) : "Directional perturbation factors for jitter and for shimmer". Jr. Commun. Dis., 17. (143-151).

STARE, R.E. (1977) : Clinical assessment of laryngeal function: a symposium. Annu. Conv. Am. Sp. Hg. Ass. Chicago.

THURMAM, W.L. (1958) : "Frequency, intensity, relationship and optimum pitch". JSHR. Vol.1. (177-129).

TITZE, I.R. (1980): Comments on the nyoelastic-aerodynamic theory of phonation. JSHR., 23, (495-510).

TITZE. I.R. & STRONG. W. (1975): Normal modes in vocal cord tissues. J.acoust. Socm am., 57, (736-744).

TITZE, I.R., & TALKIN, D. (1979): A theoritical study of the effects of the various laryngeal configurations on the acoustics of phonation. JASM., 66 No.l (60-74).

TITZE, I.R., 1995. Workshop on acoustic voice analysis, summary statement, National centre for voice and speech, (1-36).

USHA, A.A. (1978): "A study of fundamental frequency in Indian population". Master's Dissertation. AIISH. Univ. of Mysore.

VANAJA, C.S., 91986):Acoustic parameters of normal voice. Master's dissertation, Univ., of Mysore.

VAN 00RDT, H.W., & DROST, H.A. (1963): "Development of frequency range of the voice in children". Folia phonia. 15. (289-298).

VERSTRAETE, J, 1993. The effect of sustained phontion at high and low pitch on vocal jitter and shimmer, JSHR, 45 (223), 1993.

WEST, R., ANSBERRY, M., & CARR, A. (1957): The rehabilitation of speech., (III edn) harper & Row., New York.

ZEE, E. (1978): "Duration and Intensity as correlates of FO". Jr.Phonate., 6, (213-220).

ZEMLIN. W.R. (1968): Speech and Hearing Science. Englewood cliffs. N.J., Prentice-Hall, Inc..,

______(1981): Speech and Hearing Science. Prentice-Hall, Inc., Englewood Cliffs, N.J.