SPEECH ANALYSIS USING MODERN TECHNIQUES OF NONLINEAR DYNAMICS

A Thesis

submitted in partial fulfillment of the degree of

DOCTOR OF PHILOSOPHY

by

P M RADHAKRISHNAN

INTERNATIONAL SCHOOL OF PHOTONICS,
COCHIN UNIVERSITY OF SCIENCE AND TECHNOLOGY, KERALA
INDIA

NOVEMBER 2009

DECLARATION

I hereby declare that the work presented in this thesis entitled "Speech Analysis using Modern Techniques of Nonlinear Dynamics" is based on the original work done by me under the supervision and guidance of Dr. Narayanan Nampoori, Professor, International School of Photonics, Cochin University of Science and Technology. No part of this thesis has been presented for any other degree from any other institution.

P.M. Radhakrishnan

Thrikkakara

10-11-2009

CERTIFICATE

This is to certify that the thesis entitled "Speech Analysis using Modern Techniques of Nonlinear Dynamics" is a report of the original work done by P.M. Radhakrishnan under my supervision and guidance in the International School of Photonics, Cochin University of Science and Technology. No part of this thesis has been presented for any other degree from any other institution.

Thrikkakara
10.11.2009

Prof. Dr. V.P. Narayanan Nampoori
Supervising Guide,
International School of Photonics
Cochin University of Science and Technology

ACKNOWLEDGEMENTS

I am extremely grateful to Prof. V P Narayanan Nampoori, my supervising guide for his invaluable guidance and encouragement throughout the course of my research. Working with him has been a great experience and I am indebted to him for his valuable support and fruitful discussions extended at all stages of this work. He is an eminent personality with vast experience and profound insight. He is a very great human and humble person. I express my sincere thanks to him.

I would like to gratefully acknowledge Mr. Sunilkrishnan, Stuttering foundation, Ernakulam , Dr. Madhumita Kumar, Amrita Institute of Medical Sciences and Dr. Saju , ENT, General hospital, Ernakulam for the valuable assistance rendered in data acquisition required for completion of this thesis.

I am extremely grateful to Prof. Nandakumar and Prof. P. Radhakrishnan, Faculties, International School of Photonics for their support and fruitful discussions extended at all stages of this work.

I am extremely thankful to Dr. Usha Nair, Reader, Division of Electrical Engineering for her inspiration and timely guidance. I am indebted to her for her co-operation, go getting nature and invaluable assistance and inspiration without which this work would not have reached completion. I express my sincere thanks to her and her family.

I am extremely thankful to Dr. Bindu M Krishna, Young Scientist, and an eminent personality with vast experience and profound insight from the Sophisticated Test and Instrumentation Center, Cochin University of Science and Technology for her immense support during this tenure. Without her inspiration and timely guidance, this work would not have been completed. I express my sincere thanks to her.

I am grateful to Dr. Narayanan Nambhoothiri, Head, Division of Mechanical Engineering for his inspiration and timely guidance. I express my sincere thanks to him .

I am grateful to members of the Research Committee of International School of Photonics , for their kind suggestions at various stages of this work.

I would like to gratefully acknowledge Sri. A. Sreekumar, Department of Computer Applications for his valuable help in completing this work.

I am equally thankful to Sri. Russell Ravindranath, former faculty member of Civil Engineering, Cochin University of Science and Technology (CUSAT), Prof. R. Pratap, an emeritus Professor at CUSAT, Dr. Indic, former faculty member of Division of Electrical Engineering, who pursued me into this field, Sri. Ashokan, Smt. Suseela Ashokan, My School mate Dr. (Ortho) Joseph Kailath, Dr. Rajesh V.G , faculty , Model Engineering College, Thrikkakara , for their assistance at various stages of this work.

I am grateful to Sri. Joseph Alexander, Sri. Mohammed saber, Sri. Tide P. S., Sri Biju, Sri. Gireesh Kumar, Sri.Jacob Alias, faculty, Division of Mechanical Engineering, School of Engineering for supporting me during this tenure.

I sincerely thank all my friends and colleagues in School of Engineering where I am working particularly Sri. Arun Kumar, Dr. Pramod, Dr. Kannan for their support and co-operation.

My special thanks goes to my loving wife and also to my daughter for their special tips and love.

My respect and prayers are also to my relatives who prayed for me and motivated me during this work.

A special thanks to my brother and his team for the typing work.

Finally I dedicate this thesis to my mother, memories of my loving father and the lord who serves everyone as in Milton's lines "They also serve who stand and wait..."

ABSTRACT

Medical fields requires fast, simple and noninvasive methods of diagnostic techniques. Several methods are available and possible because of the growth of technology that provides the necessary means of collecting and processing signals. The present thesis details the work done in the field of voice signals. New methods of analysis have been developed to understand the complexity of voice signals, such as nonlinear dynamics aiming at the exploration of voice signals dynamic nature. The purpose of this thesis is to characterize complexities of pathological voice from healthy signals and to differentiate stuttering signals from healthy signals. Efficiency of various acoustic as well as non linear time series methods are analysed. Three groups of samples are used, one from healthy individuals, subjects with vocal pathologies and stuttering subjects. Individual vowels /അ/, /ഇ/, and /ഉ/ and a continuous speech data for the utterance of the sentence "ഇരുവരും വലിയ ചങ്ങാതിമാരാണ്" the meaning in English is "Both are good friends" from Malayalam language are recorded using a microphone . The recorded audio are converted to digital signals and are subjected to analysis.

Acoustic perturbation methods like fundamental frequency (F0), jitter, shimmer, Zero Crossing Rate(ZCR) were carried out and non linear measures like maximum

lyapunov exponent((λ_{max}), correlation dimension (D_2), Kolmogorov exponent(K_2), and a new measure of entropy viz., Permutation entropy (PE) are evaluated for all three groups of the subjects. Permutation Entropy is a nonlinear complexity measure which can efficiently distinguish regular and complex nature of any signal and extract information about the change in dynamics of the process by indicating sudden change in its value.

The results shows that nonlinear dynamical methods seem to be a suitable technique for voice signal analysis, due to the chaotic component of the human voice. Permutation entropy is well suited due to its sensitivity to uncertainties, since the pathologies are characterized by an increase in the signal complexity and unpredictability. Pathological groups have higher entropy values compared to the normal group. The stuttering signals have lower entropy values compared to the normal signals.

PE is effective in charaterising the level of improvement after two weeks of speech therapy in the case of stuttering subjects. PE is also effective in characterizing the dynamical difference between healthy and pathological subjects. This suggests that PE can improve and complement the recent voice analysis methods available for clinicians.

The work establishes the application of the simple, inexpensive and fast algorithm of PE for diagnosis in vocal disorders and stuttering subjects.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS

EX	Coefficient of excess
PA	Pitch Amplitude
HNR	Harmonic to Noise Ratio
LE	Lyapunov Exponent
K-S entropy	Kolmogorov-Sinai entropy
FNN	False Nearest Neighbour
FFT	Fast Fourier Transform
PE	Permutation Entropy
ApEn	Approximate Entropy
SampEn	Sample Entropy
STFT	Short-time Fourier Transform
PWS	Person With Stammering
D_2	Correlation Dimension
K_2	Kolmogorov entropy
LPC	Linear Predictive Coding
ZCR	Zero Crossing Rate
% jit	Percentage jitter
% shm	Percentage shimmer

CHAPTER 1-INTRODUCTION

Speech is the basic mode of communication and a unique feature of human beings. Human speech results from the complex interaction between phonation and articulation. Phonation or primary voice signal is generated in the larynx by the vibratory structures like vocal folds, ventricular folds and epilarynx. These primary sound signals are filtered in the respiratory cavities by vocal tract articulators like the tongue, soft palette, lips pharynx and nasal cavity. These articulations are produced by variation of the geometry of the filter cavities and thus the recognizable words are produced. If the vocal tract contains no constrictions then vowels are produced. In general, speech can be considered as a series of voiced and unvoiced sounds in a meaningful temporal sequence [1].

According to their mode of excitation, speech sound can be classified into unvoiced sounds or fricative and plosives [2]. Speech is basically the acoustic wave produced when air is expelled from the subglottal system comprising of lungs, bronchia and trachea. When air expelled from the lungs propagates through the nasal tract and the vocal tract, which are tubes of non-uniform crossectional area, these tubes modify the frequency spectrum by their resonance effect. The geometry and dimension of these tubes decides their frequency selectivity. The resonance frequencies of the vocal tract tube are called the formant frequencies. The uniqueness of speech of any individual is decided by its spectral characteristics. Complex interplay of frequencies produces characteristic features of sound. This emphasizes the importance of acoustic analysis of human voice or vocal disorders. Thus sound spectrograph was the principle tool for speech analysis for many years. In addition several acoustic perturbation methods were also in use for measurement diagnosis and voice treatment. The conventional acoustic analysis tools include fundamental frequency (F0), Harmonic to Noise Ratio (HNR), Pitch Amplitude (PA), jitter, shimmer, Zero Crossing Rate (ZCR). However, all these traditional analysis techniques were based on the assumption that speech production can be modeled as a linear process according to the standard assumptions of linear acoustics and 1D plane wave propagation of sound in the vocal tract [3]. Eventhough this model has been applied to speech coding [2, 4]; recognition and synthesis, there are several experimental and theoretical evidences for nonlinear aerodynamic phenomenon during speech production [5 - 8]. Therefore, recently the application of nonlinear dynamical analysis has received much scientific interest and has successfully characterized laryngeal pathologies [9 - 15].

The sensitiveness of traditional perturbation methods to signal length sampling, noise and its appropriateness to aperiodic voice signal [15 - 18] has added to the significant shift of interest to non-linear methods. Nonlinear analysis methods are generally based on the theory of dynamical systems. Based on this theory, the dynamics of any system can be represented in the state space constructed from its dynamical variables. According to nonlinear time series methods, the dynamics of any system with many degrees of freedom can be investigated using time series of a single scalar observable. From a dynamical system perspective a speech signal can be regarded as an observable of speech production system, which can be used to uncover its dynamics.

Considering the voice signal as a time series data necessary information about the underlying system dynamics can be extracted by reconstruction of phase space behaviour. Recently such method has gained much attention by their successful voice characterization [19, 20]. Although such phase space reconstruction methods are found to be successful to an extent [12, 14,15, 18, 21] they are sensitive to non-stationarity and noise contamination of the signals. Conventional nonlinear method of correlation dimension is found to be affected by non-stationarity, noise and finite signal length in most of the biological data including electroglotagraphic signal [21]. Under such circumstances new method of entropy based on the asssement of the predictability of the system are found to be a better choice [22 - 26].

The aim of this thesis is to compare nonlinear dynamical analysis and perturbation analysis methods for assessing their applicability in real time and online setup. One of the main problems still unsolved in the domain of speech fluency disorders is an objective and an automatic way of judgments of patient performance before and after speech therapy sessions and an assessment of gains made after intervention. For an effective evaluation and comparison of speech quality improvement clinicians and therapist should have quantitative information about the improvement that could be provided by specific methods.

The appropriate application of perturbation analysis and nonlinear dynamic analysis can provide different but complementary information and may potentially improve our ability to objectively assess voice disorders and evaluate the effects of treatment of laryngeal pathologies.

Most of the existing methods of signal analysis give significant results when the time series is simulated from low dimensional dynamical systems and fails or misleads in the presence of noise. Hence real world time series analysis of the data requires preprocessing for noise elimination. Furthermore embedding dimension and time delay are critical parameters in reconstruction of state space and computation is time consuming, which restricts its application on real time basis. Hence it is essential to have a very fast algorithm, which can process the data at the same rate at which it is acquired.
Our goal is primarily to detect dynamical changes using technique from real world data sets where there is no time for preprocessing and fine-tuning of data. Here we aim to make use of nonlinear parameters like maximum Lyapanov exponent (λ_{max}), Correlation Dimension (D_2), Kolmogorov entropy (K_2) and a new entropy measure Permutation Entropy (**PE**) for this purpose by establishing a link with the conventional state space approach.

We also aim to study the effectiveness of the invariant parameters and PE in extracting the change in dynamics caused by abnormalities in the vocal tract. This may be of great advantage in preliminary clinical diagnosis in identifying the vocal disorders before proceeding for further expensive treatment strategy.

Aim of the thesis

(1) To study the nonlinear characteristics of stuttering subjects vs. normal Subjects.
(2) To study the nonlinear properties of normal and abnormal speech Processes.
(3) To directly compare the efficiency of traditional perturbation and nonlinear time series analysis methods in characterizing speech signals of vocal disorder and stuttering subjects.
(4) To estimate the efficacy of speech therapy rendered to the stuttering subjects using a fast and robust entropy called Permutation Entropy (PE).
(5) To investigate the effectiveness of different methods in detecting change in dynamics in the above processes.

1.1 THESIS ORGANISATION

Chapter 1: Introduces the problem and defines the aim of the thesis.

It explains that speech production is not a linear phenomenon but it is a complex mechanism and acoustic studies and linear methods are not sufficient to completely study the dynamics. It is better to go for a more general method such as an entropy measure to differentiate between two kinds of signals, the healthy and pathological cases.

Chapter 2: Speech production and evidence of Nonlinear Behaviour.

It deals with the anatomy of speech production and explains the various frequencies that arises due to the oscillations of the vocal tract which appears to be quasi periodic in nature and due to the modulations of muscle parts in the nasal tract like tongue, lips , jaw, teeth etc different resonances of sounds are produced. The evidences of nonlinear dynamics in speech production are also explained.

Chapter 3: Current Techniques in Speech Analysis

It includes various linear and nonlinear methods used in the present work and in general. Various linear measures to characterize speech process are spectrogram, jitter, shimmer, signal to noise ratio, linear predictive coding (LPC), Pitch amplitude, zero crossing. Power spectral analysis, wavelet analysis and statistical methods used to describe the complex nature of the system. Nonlinear measures used for characterizing the dynamics of speech are Lyapunov exponent, Correlation Dimension, cross-correlation sum, pseudo entropy, pseudo correlation dimension, fractal dimension etc.

Chapter 4: Discusses the results of the analysis of stuttering.

Chapter 4 deals with the data recording and methodology of analysis adopted in stuttering. The present study evaluated speech of 10 stuttering subjects of age groups 16 - 40 and 10 age matched control subjects speaking Malayalam –their native language(a south Indian language). The intension of present research is to find the difference in speech parameters between stutterers and fluent speakers. Vocal signals are recorded for Malayalam vowels('അ' 'ഇ' 'ഉ' and a sentence from a children's story "ഇരുവരും വലിയ ചങ്ങാതിമാരാണ്" the meaning in English is "Both are good friends" using a microphone and a multimedia Computer in a sound proof room before a group of audience with a sampling rate of 11khz.

The results achieved were to develop an objective and automatic method for charactersing the stuttered and fluent speech. The present work focuses on the robust and fast processing of the data using Permutation entropy (PE) even in the presence of noise which were not possible by other nonlinear methods.

The results indicate that PE is an efficient measure for characterising the audio signal and establishing quantitatively the improvement in the progress of post treatment rendered by speech therapist and clinicians to Persons With Stammering (**PWS**) time to time. Specifically, the evolution of PE, average phonation number, average PE, maximum Lyapunov exponent (**λ_{max}**), Correlation Dimension (D_2), Kolmogorov entropy (K_2) calculated proves this. At every level of treatment the PE data can be stored in the database of the patient, which can be compared to the PE data of patients before treatment, and the database of PE of the fluent speaker. This can also give the first hand information in the diagnosis about the treatment and thus help the therapist and clinicians. The result confirms that the stuttering dynamics exhibits lowered complexity when compared to that of normal speech signals. No noise filtering or reconstruction of attractor is needed unlike other nonlinear methods, which requires lots of computation time. PE is robust to noise and allow fast evaluation.

Chapter 5: Discusses the results of the analysis of vocal disorders.

The Permutation Entropy technique allows visualizing the differential dynamics between healthy voices and voices with some vocal fold pathologies. The evolution of PE, Average PE, Correlation dimension, Lyapunov exponent and Kolmogorov entropy calculated differentiates the healthy signals from the pathological signals. The results show that using PE an overall increase is indicated in PE and this can be measured very fast. No noise filtering or reconstruction of attractor is needed unlike other nonlinear methods, which requires lots of computation time. PE is robust to noise and allow fast evaluation.

Two groups of samples were used, one from healthy individuals and the others from people with nodule in the vocal fold, Reinke's edema, paralysis, polyps, cancer and papilloma . Vowels /അ/, /ഇ/ and /ഉ/ from Malayalam language are used for analysis.

Permutation entropy is well suited due to its sensibility to uncertainties, since the pathologies are characterized by an increase in the signal complexity and unpredictability. The results showed that the pathological groups had higher entropy values in accordance with other vocal acoustic parameters presented. This suggests that these techniques may improve and complement the recent voice analysis methods available for clinicians. Effectiveness of PE to identify the vocal pathologies is verified on clinically characterized vocal sound data from 10 patients suffering from six different cases. Time series of sound signal for each vowel from 10 pathological cases are used for PE analysis. PE of order 5 is calculated for sliding window of 512 samples.

Chapter 6: Contains Summary, conclusions and scope for future work.

This deals with the overall summary, conclusion and scope for future work.

1.2 REFERENCES:

[1] M. Petursson and J. Neppert., Elementarbuch der Phonetik. Buske, Hamburg, 1991.

[2] L.R.Rabiner and R.W.Schafer, Digital Processing of Speech Signals, Englewood Cliffs, NJ, Prentice-Hall, 1978.

[3] Vassilis Pitsikalis and Petros Maragos, Speech Analysis and feature extraction using chaotic models , IEEE, 2002.

[4] L. R. Rabiner and B. H. Juang, Fundamentals of Speech Recognition, Prentice-Hall, 1993.

[5] J. F. Kaiser, Some observations on vocal Tract Operation from a Fluid Flow Point of View, in Vocal Fold Physiology: Biomechanics, Acoustic, and Phonatory Control, I. R. Titze and R. C. Scherer(Eds.,), Denver Centre for Performing arts, Denver, CO, pp. 358-386, 1983.

[6] H. M. Teager and S. M. Teager, Evidence of Nonlinear Sound Production Mechanisms in the Vocal Tract in speech Production and Speech Modelling, W. J. Hardcastle and A. Marchal, Eds., NATO Advanced study Institute Series D, vol.55, Bonas, France, July 1989.

[7] T.J.Thomas, A finite element model of fluid flow in the vocal tract, Comput. Speech Lang. vol. 1, 131-151, 1986.

[8] D. J. Tritton, Physical Fluid Dynamics, 2nd edition, Oxford Univ. Press, New York, 1988.

[9] H. Herzel, D. Berry, I. R. Titze, and M. Saleh, Analysis of vocal disorders with methods from nonlinear dynamics, J. Speech Hear. Res. 37, 1008–1019, 1994.

[10] I. R. Titze, R. Baken, and H. Herzel, Evidence of chaos in vocal fold vibration, in Vocal Fold Physiology: New Frontier in Basic Science, edited by I. R. Titze, pp. 143–188, Singular, San Diego, 1993.

[11] A. Behrman and R. J. Baken, Correlation dimension of electroglottographic data from healthy and pathologic subjects, J. Acoust. Soc. Am.102, 2371–2379, 1997.

[12] Y. Zhang and J. J. Jiang, Nonlinear dynamic analysis of signal typing of pathological human voices, Electron. Lett. 39, 1021–1023, 2003.

[13] A. Giovanni, M. Ouaknine, and J.-M. Triglia, Determination of largest Lyapunov exponents of vocal signal: application to unilateral laryngeal paralysis, J. Voice 13, 341, 1998.

[14] J. J. Jiang and Y. Zhang, Nonlinear dynamic analysis of speech from pathologic subjects, Electron. Lett. 38, 294–295, 2002.

[15] Y. Zhang, C. McGilligan, L. Zhou, M. Vig, and J. J. Jiang, Nonlinear dynamic analysis of voices before and after surgical excision of vocal polyps, J. Acoust. Soc. Am. 115, 2270–2277, 2004.

[16] Kay Elemetrics Corp, Multi-Dimensional Voice Program: Software Instruction Manual, Pine Brook, NJ, 1993.

[17] I. R. Titze, Workshop on acoustic voice analysis; Summary statement, National Center for Voice and Speech, Denver, CO (1995), p. 36. 5M. P. Karnell, A. Chang, A. Smith, and H. Hoffman, Impact of signal type of validity of voice perturbation measures, NCVS Status Prog. Rep. 11, 91–94, 1997.

[18] J. J. Jiang, Y. Zhang, and C. N. Ford, Nonlinear dynamic of phonations in excised larynx experiments, J. Acoust. Soc. Am. 114, 2198–2205, 2003.

[19] I.R. Titze, R.J. Baken, H. Herzel, Vocal fold physiology – frontiers in basic science, Evidence of Chaos in Vocal Fold Vibration, Singular Publishing Group, San Diego, 1993.

[20] Marwan Al-Akaidi, Fractal Speech Processing, Cambridge University Press, Cambridge, 2004.

[21] Behrman and R. J. Baken, Correlation dimension of electroglottographic data from healthy and pathologic subjects, J. Acoust. Soc. Am.102, 2371–2379, 1997.

[22] Paulo Roge´rio Scalassara , Marı´a Eugenia Dajer, Carlos Dias Maciel, Rodrigo Capobianco Guido, Jose´ Carlos Pereira, Relative entropy measures applied to healthy and pathological voice characterization, Applied Mathematics and Computation 207, 95–108, ELSEVIER, 2009.

[23] R. Moddemeijer, On estimation of entropy and mutual information of continuous distributions, Signal Process. 16 (3), 233–248, 1989.

[24] R. Kleeman, Measuring dynamical prediction utility using relative entropy, J. Acoust. Soc. Am. 59 (13), 2057–2072, 2002.

[25] R.E. Hillman, E.B. Holmberg, J.S. Perkell, M. Walsh, C. Vaughan, Phonatory function associated with hyper functionally related vocal fold lesions, J. Voice 4 (1), 52–63, 1990.

[26] B. Hammarberg, Perception and acoustics of voice disorders: a combined approach, in: Proceedings of VOICEDATA98, Symposium on Databases in Voice Quality Research and Education, Utrecht Institute of Linguistics, Utrecht, The Netherlands, pp. 1–6, 1998.

CHAPTER 2- SPEECH PRODUCTION AND EVIDENCES OF NONLINEAR BEHAVIOUR

2.1 INTRODUCTION

Speech production is a complex feedback process in which also hearing, perception and information processing in the nervous system and the brain is involved. In this chapter the human organs which are contributing to the speech production process are explained in brief. For the different categories of speech sounds the mechanism of production is explained with respect to sound excitation and the different articulators. On the basis of proper tube models, the physics and acoustics of speech sound generation, propagation within the speech production system and radiation is discussed.

Voice production can be thought of as the activation of an entire system of coupled oscillators. The intent to vocalize activates motor commands that are responsible for the neural inputs to an array of biomechanical, neural, and acoustic oscillators (large box in Figure 2.1).

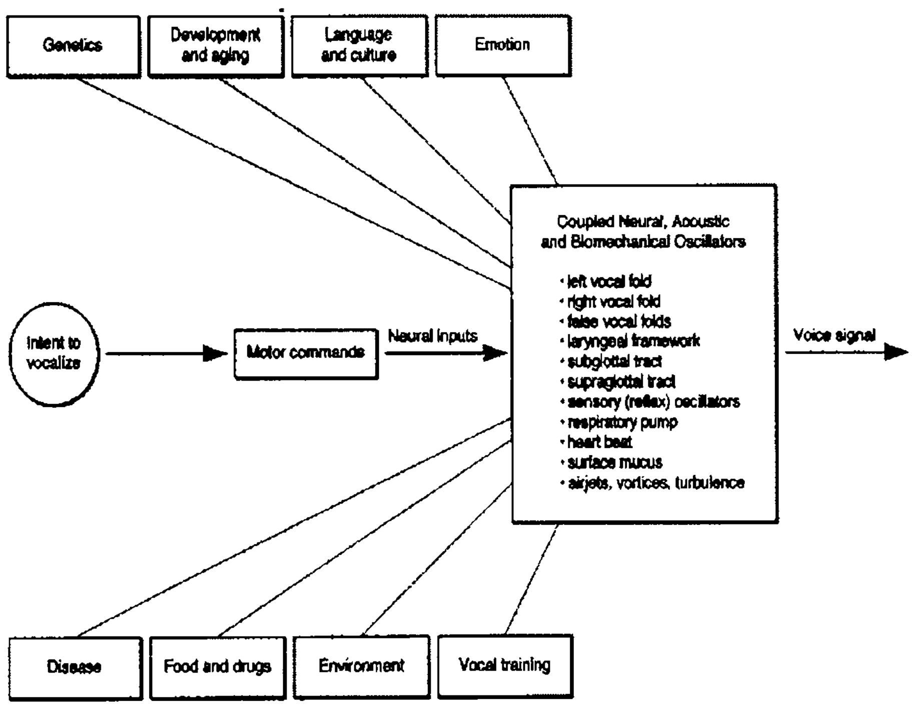


Figure 2.1. A list of biological oscillators involved in voice production and factors that may influence them.

The vocal folds are the primary oscillating system that produce the carrier signal (the glottal airflow). All other oscillators can then be thought of as modulators of the carrier signal. Some of the modulations are nearly sinusoidal (respiratory, heart beat) but many are high dimensional (action potentials of muscles, air vortices, mucus in motion). Yet others are passive oscillators (tracheal resonator, supraglottal vocal tract, various sinuses) that can influence the primary oscillating system. We can assume that the system of coupled oscillators contains and releases information about the human body; in particular, about its genetics, development, age, disease, language, culture, food and drug intake, and response to the environment (Figure 2.1).

Voice perturbation analysis has the goal of extracting some of this information from the voice signal. In all cases, the procedure is extremely difficult and usually requires considerable a priori knowledge about the modulations to be extracted. Therein lies the primary problem of voice perturbation analysis in its present state. Many studies are needed to isolate the individual contributions of each oscillator. Some of these studies are underway (J. van den Berg [1]).

Before discussing the acoustic theory and modern techniques in nonlinear dynamics for speech analysis, it is important to consider the various types of sounds that make up human speech. Speech can be broken down in to small segments called phonemes, each of which is unambiguously distinguishable and can be represented by any of a number of different phonetic alphabets. The figure 2.2 shows a schematic diagram of the vocal system .

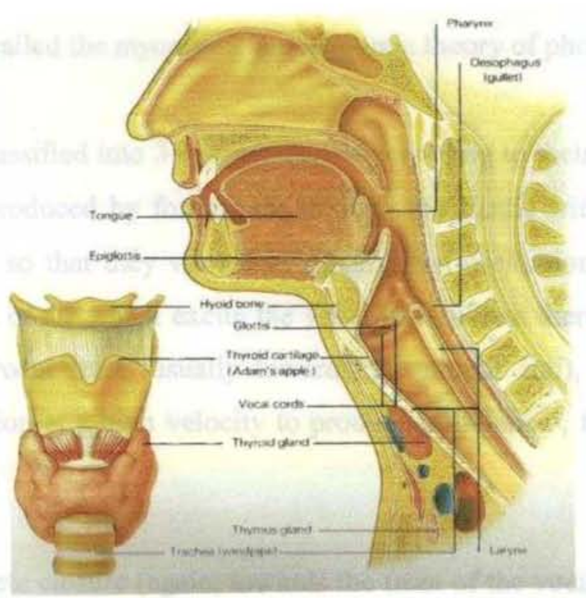


Figure 2.2. Schematic diagram of the vocal system shows components in the airway system in the head, neck, and chest (Titze, 1994a)[2]: The lungs produce pressure that drives the subglottal airstream. The airstream is fed to the larynx via bronchi and trachea. The primary function of the larynx is to protect the airway system from foreign material (such as food) which passes into the esophagus during swallowing. The vocal tract, comprised of pharynx and oral and nasal cavities, filters the primary sound signal generated by airstream-driven oscillations of the vocal folds in the larynx.

The basic mechanism for producing any sound is to expel air from the lungs, using muscular action. The actual mechanism by which we create a phoneme can be split into two main categories, voiced and unvoiced that can be further split in to vowels, fricatives or plosives.

Voiced speech or phonation is produced by oscillating the fleshy membranes inside the larynx which are known as the vocal folds. The oscillation is set up by forcing the vocal folds closed which causes pressure to build up below the folds, gradually forcing them to open again allowing the air to flow from the sub glottal region in to the mouth. This rapid air flow creates a Bernoulli force which coupled with the muscular action of the vocal muscles produces the sound.

Thus an oscillation is set up with the fundamental frequency being a function of the vocal fold tension which is controlled by the vocalis muscles. This theory of how the oscillations occur is called the myoelastic/aerodynamic theory of phonation [3].

Speech sounds are classified into 3 distinct classes according to their mode of excitation. Voiced sounds are produced by forcing air through the glottis with the tension of the vocal cords adjusted so that they vibrate in a relaxation oscillation, thereby producing quasi-periodic pulses of air which excite the vocal tract. When there is a constriction at some point in the vocal tract (usually towards the mouth end), thereby forcing air through the constriction at a high velocity to produce turbulence , fricative or unvoiced sounds are generated.

When there is complete closure (again, towards the front of the vocal tract), building up pressure behind the closure , and abruptly releasing it plosive sounds is produced.

2.2 PRODUCTION OF SPEECH

As sound generated as mentioned above, propagates down the non uniform cross-sectional tubes of the vocal tract and nasal tract shown as in figure 2.3.

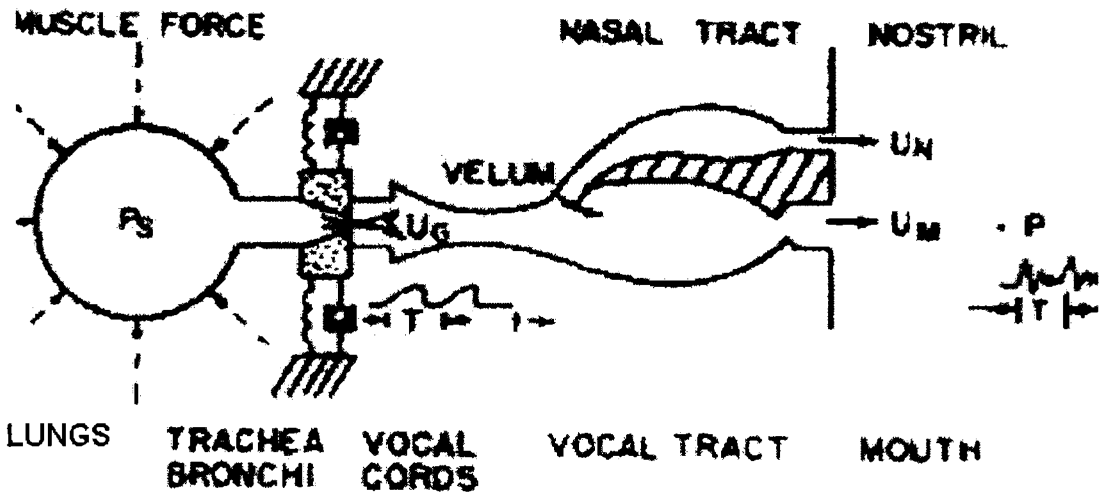


Figure 2.3.–Schematized diagram of the vocal apparatus (After Flanagan et al. [4])

The frequency spectrum is shaped by the frequencies selectivity of the tube as shown in figure 2.3. The resonance frequencies of the vocal tract tube are called formants. The formant frequencies depend upon the shape and dimension of the vocal tract. Each shape is characterized by the formants frequencies. Different sounds are performed by varying the shape of the vocal tract. Thus spectral properties of the speech signal vary with time as the vocal tract shape varies.

One of the basic method of analysis is spectrogram. The time varying spectral characteristics of speech signal can be graphically displayed through the use of sound spectrograph. It produces a 2 dimensional graph called spectrogram in which the vertical dimension corresponds to frequency and horizontal dimension to time. The darkness of the pattern is proportion to the signal energy. Thus the resonance frequencies of the vocal tract show up as dark bands in the spectrogram.

Due to the complex interplay of frequencies produce the characteristic features of sound. This emphasizes the importance of acoustic analysis of human voice or human disorders. For an extended introduction to the anatomy and physiology of the human voice production system see, e.g., [4] and [5].

2.3 EVIDENCE OF NONLINEAR BEHAVIOUR

There are a number of areas that show evidence of nonlinear behaviour in speech generation:

1. Vocal folds
2. Turbulent air flow
3. Non-plane wave propagation
4. Higher order statistics
5. Chaotic behaviour

The following sections look at each of these areas.

1. Vocal folds

There are a number of features about the oscillation of the vocal folds that show they are nonlinear: In a linear model the output is proportional to the input and yet the waveform generated from vocal fold oscillation actually changes shape under different amplitude levels. Detailed studies [6, 7] of this effect show that not only does the spectral content of the pulse alter with amplitude but also that the spectral envelope changes with fundamental frequency. The vocal chords display bifurcations [8–10]. The clearest example of this is the passage from unvoiced to voiced speech where the oscillations move from an equilibrium state, i.e. not moving, to a pseudo-periodic motion. Bifurcations are a trait of nonlinear systems but in this case there has to be a question as to whether the bifurcations are caused by the driving force passing a threshold, as in classic bifurcation systems, or whether there is some higher muscular force that is controlling the transition. Models such as multiple masses routinely include nonlinear coupling between the mass elements [11, 10]. This is based on the knowledge that the cartilage and flesh constructing the larynx have nonlinear stretching qualities. A number of works have suggested that chaotic modes of operation can be found for vocal fold oscillation. These works should be viewed very carefully since chaos is an extremely difficult phenomenon to quantify, as will become clear throughout this thesis, and can very often cause very misleading results.

2. Turbulence

When unvoiced speech is created there is a point of constriction in the vocal tract which causes turbulence to occur. Turbulence is a nonlinear effect which occurs because of an interaction between the air flow and the acoustic sound field. As already discussed there is plenty of evidence to show that cavitation noise, which is turbulence, is chaotic and there are a number of works that suggest the same is true of fricative sounds.

3. Non-plane wave propagation

The usual model of the vocal tract is that the sound travels along the tract as plane wave propagation. Recently this view has been challenged by Teager and Teager [12] who suggest that the flow consists of a number of vortices. This work is based on examining the air flow at a range of points within the vocal tract using hot wire anemometers. If this is indeed the case then it throws into question the whole acoustic model which is based on the idea that the vocal tract can be considered as a number of acoustic tubes which have well defined reflection and standing wave properties. A good example that shows this effect is given by Kubin [13]: what is the mechanism for human whistling since no part of the vocal tract is in oscillation? The explanation given is, in summary, that an unstable jet of air is created which gives rise to vortices, when the travelling time through the vocal tract matches the frequency of the vortices then periodic vortex shedding occurs at the lips giving rise to the narrow band whistle.

4. Higher order statistics

Higher order statistics (HOS) can be used to identify the underlying nonlinearities present in a system. Unfortunately the application of HOS theories to noisy signals is very difficult and consequently the application of HOS to speech has not produced conclusive results. However what results have been published [14, 15] suggest that there is strong evidence of quadratic phase coupling, which would indicate nonlinearity.

5. Chaotic behavior

Several times in this chapter the possible existence of chaotic behaviour has been suggested. This section gives a quick overview of the work that has been conducted and some discussion of the possible shortcomings that may give rise to a number of misleading results. Most of the work in this field seems to have been inspired by the work of Teager and Teager [12] which gave clear indications that speech was nonlinear; if it is nonlinear then could it be chaotic or fractal in nature? Maragos [16] suggests that fricatives have a fractal dimension of as low as 1.7 whilst vowels may have a fractal dimension of nearer 1.2. The calculation of this dimension is through the box counting technique [17] which is restricted to a 2 dimensional plane and explains why the figures are so low, and in disparity with the dimension measures given by other authors. It should be noted that this is not saying that the measurements are wrong it is merely pointing out that the box counting dimension looks at the dimension of a waveform not of the generating system itself. The paper also shows calculation of dimension using very small data sets and showing no form of noise cancellation, these are shortcomings that are consistent with many other papers in the field. Both Boshoff [18] and McDowell and Datta [19] give similar analyses suggesting box counting dimensions of between 1 and 2 although McDowell and Datta [19] point out that the accuracy of these results is questionable. Pickover and Khorasani [20] attempt a similar analysis but on full sentences. This raises the spectre of stationarity; speech is constructed from many small segments that individually may be viewed as stationary, the normal size of these sections is about 10ms which is based on the relatively slow movement of the articulators, but a complete sentence includes many different modes of operation and indeed periods of complete silence. As a diagnostic tool this approach may have some use if it is used to compare the characteristics of different speakers saying the same sentences, but should not be used to give a definition of the fractal dimension of speech as a whole.

Marcato and Mumolo [21] show that fractal theory can be applied to the LPC to give an efficient coding of the residual signal. Fractals are similarly applied to image coding [22] and speech recognition [23].

McLaughlin and Lowry [24] use the correlation dimension to investigate a range of vowels with the conclusion that although they do seem to show low dimensional properties, the correlation dimension fails to give an accurate measure. These results are consistent with the general disenchantment with correlation dimension when applied to real world signals. Tishby [25] again examines the correlation dimension giving similar vague reports of dimensions ranging from 3 to 5 for voiced speech. He also looks at the possibility of forming a local nonlinear predictor using neural networks to enhance current predictor based systems. A similar work by Moakes and Beet [26, 27] suggests that speech is low dimensional and they apply Radial Basis Functions (RBF) to both recognition and predictive problems. Berhard and Kubin [28, 29] give preliminary evidence for low dimensional behaviour, of the order of 1 to 2, for vowels. In a very recent paper Narayanan and Alwan [30] look at fricatives showing the difficulties of convergence for the correlation dimension but suggesting low dimensions for vowels and high, around 4 to 7, dimensions for fricatives. They also examine the Lyapunov spectra suggesting that vowels have a non-chaotic structure whilst fricatives may have a single positive exponent. Bohez, Senevirathne and VanWinden [31] give a very clear application of fractal theory to recognition of vowels. Again they do not attempt to infer the actual underlying system's dimension from the fractal dimension but rather use it as a discriminatory tool. In another paper by the same authors they present an analysis of speech using an alternative box counting technique called the amplitude-scale method. Unfortunately this technique seems to give wildly different results from the box counting technique and again is limited to a 2 dimensional space. Townshend [32] gives a very full overview of the possible uses of nonlinear predictors in speech along with presenting correlation dimension results of just less than 3. These results again do not appear to be for stationary segments of speech and must be considered with care. As should be clear there has been considerable work presented in the field although on the whole the problems of noise contamination, data set size and stationarity have not been addressed fully.

2.4 REFERENCES:

1. J. Van den Berg, Myoelastic -aerodynamic theory of voice production, Journal of Speech and Hearing Reasearch , 111 2227-44, 1958 .

2. I. R. Titze, Principles of voice production, Prentice-Hall, 1994a.

3. A. Breen, Speech synthesis models: a review, Electronics and Communication Engineering Journal, pp. 19–31, February 1992.

4. J. L. Flanagan, C. H. Coker, L. R. Rabiner, R. W. Schafer, and N. Umeda, Synthetic Voices for Computers, IEEE Spectrum, Vo. 7, No. 10, pp. 22-45, October 1970.

5. J. Wendler, W. Seidner, G. Kittel, and U. Eysholdt. Lehrbuch der Phoniatrie und P¨adaudiologie. Georg Thieme Verlag, Stuttgart, New York, 1996.

6. I. R. Titze, The physics of small-amplitude oscillation of the vocal folds, J Acoust Soc Am, vol. 83, pp. 1536–1552, April 1988.

7. I. R. Titze, Phonation threshold pressure: A missing link in glottal aerodynamics, J Acoust Soc Am, vol. 91, pp. 2926–2935, May 1992.

8. J. Awrejcew, Bifurcation portrait of human vocal oscillations, Journal of Sound and Vibration, vol. 136, no. 1, pp. 151–156, 1990.

9. H. Herzel and J. Wendler, Evidence of chaos in phonatory samples, in EUROSPEECH 91̓, vol. 1, pp. 263–266, 1991.

10. I. Steinecke and H. Herzel, Bifurcations in an asymmetric vocal-fold model, Acoustical Society of America, vol. 97, pp. 1874–1884, March 1995.

11. T. Koizumi, S. Taniguchi, and S. Hiromitsu, Two-mass models of the vocal cords for natural sounding voice synthesis, Journal of the Acoustical Society of America, vol. 8, pp. 1179–1192, October 1987.

12. H. M. Teager and S. M. Teager, Evidence for nonlinear sound production mechanisms in the vocal tract, in Proc NATO ASI on Speech Production and Speech Modelling, pp. 241–261, 1990.

13. G. Kubin, Nonlinear processing of speech, in Speech Coding and Synthesis (W. B. Kleijn and K. K. Paliwal, eds.), pp. 557–610, Amsterdam: Elsevier, 1995.

14. S. McLaughlin, S. Hovel, and A. Lowry, Identification of nonlinearities in vowel generation, in EUSIPCO 94, vol. 2, pp. 1133–1137, Elsevier Science, 1994.

15. J. W. A. Fackrell and S. McLaughlin, The higher order statistics of speech signals, in IEE Colloquium on Techniques for speech processing and their application, no. 1994/138, pp. 7/1–7/6, IEE, June 1994.

16. P. Maragos, Fractal aspects of speech signals: dimension and interpolation, in ICASSP 91, pp. 417–420, IEEE, 1991.

17. L. S. Liebovitch and T. Toth, A fast algorithm to determine fractal dimension by box counting, Physics Letters A, vol. 141, pp. 386–390, November 1989.

18. H. F. V. Boshoff and M. Grotepass, The fractal dimension of fricative speech sounds," in COSMIG '91, pp. 12–16, IEEE, 1991.

19. P. S. McDowell and S. Datta, The fractal characterisation of isolated human speech, Proceedings of the Institute of Acoustics, vol. 16, no. 5, pp. 247–253, 1994

20. C. A. Pickover and A. Khorasani, Fractal characterization of speech waveform graphs, Comput and Graphics, vol. 10, no. 1, pp. 51–61, 1986.

21. L. Marcato and E. Mumolo, Coding of speech signal by fractal techniques, in EUROSPEECH '93, pp. 745–748, 1993.

22. M. F. Barnsley and A. D. Sloan, A better way to compress images, BYTE, pp. 215–223, January 1989.

23. E. L. J. Bohez, T. R. Senevirathne, and J. A. VanWinden, Fractal dimension and iterated function system (ifs) for speech recognition, Electronic Letters, vol. 28, pp. 1382–1384, July 1992.

24. S. McLaughlin and A. Lowry, Nonlinear dynamical systems concepts in speech analysis, in EUROSPEECH '93, pp. 377–380, 1993.

25. N. Tishby, A dynamical systems approach to speech processing, in ICASSP '90, pp. 365–368, IEEE, 1990.

26. P. A. Moakes and S.W. Beet, Recurrent radial basis functions for speech period detection, Proceedings of the Institute of Acoustics, vol. 16, no. 5, pp. 271–278, 1994.

27. P. A. Moakes and S.W. Beet, Analysis of non-linear generating dynamics, in ICSLP 94, pp. 1039–1042, 1994.

28. H. P. Bernhard and G. Kubin, Detection of chaotic behaviour in speech signals using fraser's mutual information algorithm, in 13th GRETSI symposium on signal and image processing, 1991.

29. H. P. Bernhard and G. Kubin, Speech production and chaos, in 13th GRETSI symposium on signal and image processing, 1991.

30. S. S. Narayanan and A. A. Alwan, A nonlinear dynamical system analysis of fricative consonants, The Journal of the Acoustical Society of America, vol. 97, pp. 2511–2524, April 1995.

31. T. R. Senvirathne, E. L. J. Bohez, and J. A. VanWinden, Amplitude scale method: Newand efficient approach to measure fractal dimension of speech waveforms, Electronics Letters, vol. 28, pp. 420–422, February 1992.

32. B. Townshend, Nonlinear prediction of speech, in IACSSP '91, pp. 425–428, IEEE, 1991.

Chapter 3-CURRENT TECHNIQUES IN SPEECH ANALYSIS

This chapter gives a detail account of various linear and nonlinear techniques used for speech analysis. Special emphasis is given to methods used for characterizing pathological conditions like vocal disorders and stuttering in speech process. The vocal and voice diseases should to be diagnosed in the early stages. It is well known that most of these diseases cause changes in the acoustic voice signal. Therefore, the voice signal can be a useful tool to diagnose them, In the bibliography, there are many algorithms to calculate acoustic parameters and it was shown that there is a great correlation between these parameters and the pathologies. For example, it was demonstrated that some acoustic parameters are deviated from the mean in the case of pathological voices.

Furthermore, the diagnosis of pathologies from the voice signal presents several advantages with regard to other methods:

1. It is a non-invasive tool and easy to use, the patient only has to speak.
2. It makes possible to build an automatic computer based diagnosis system.
3. The diagnostic is objective because it is based on the value of acoustic parameters.
4. It can also be used for the evaluation of surgical and pharmacological treatments and rehabilitation processes

3.1 LINEAR TECHNIQUES

The analysis of voice signal is usually performed by the extraction of acoustic parameters using digital signal processing techniques. After that, these parameters are analyzed to determine the characteristic of the voice: nonpathologic, pathologic and the type of pathology.

In the bibliography, there are a great number of acoustic parameters that can be extracted and analyzed, however it is not absolutely clear its usefulness for solving the problem. The selection of the most appropriate parameters is still an open problem.

NEEDS FOR DATABASE

There is a need for at least three types of databases, which should be recorded under standard acoustic conditions, on standard linguistic material within a language and comparable linguistic material across languages, on standard equipment, under standard procedures, with standard biographical documentation of the speakers involved. The first is a database of voice samples of patients known to be suffering from specific voice disorders, both pathological and functional. This database should include graded samples of a wide range of disorders at different degrees of severity, from men, women, and children. The second is a database of speakers suspected to be suffering from laryngeal pathology or dysfunction, recorded before diagnosis, with subsequent information available about the diagnoses. The third is a database of a control group of speakers, matched in sex, age, socioeconomic status, general physique, and general state of health to the speakers of the first two databases. This control database should be divided into two halves: smokers and nonsmokers. Neither group of control speakers should have a reported history of voice disorders.

A major part of the study of perturbation has examined the speech waveforms of young, healthy adults (usually university students), together with the phonation of speakers with abnormal laryngeal conditions. In the latter group, the speakers are often older than the healthy speakers, and predominantly male in sex. For a control group, data should be collected from a wider age and sex range of speakers to provide more representative normative information. For the pathological group, data should be collected from as wide a range of disorders as possible. The voice samples in the databases should include not only the production of sustained vowels, but also a sample of continuous speech.

MAJOR ISSUES

The major issue in perturbation is determining the correlation between the acoustic characteristics of perturbed phonation and their anatomical, physiological, and aerodynamic counterparts. Once a better understanding is achieved of the mechanical consequences for phonation of different types of disruption of laminar tissue relationships, the search for an automatic system that is capable of providing diagnostic support to clinicians will be able to proceed on sounder footing. An important part of this development will be the optimization of the choice of perturbation parameters for different disorder types. Practically, a crucial objective is the development of a microprocessor-based version of perturbation analysis for low-cost use in otolaryngology clinics and speech pathology clinics, to provide quantified support for remedial and monitoring functions. In the meantime, progress in this area is likely to be focused on correlational studies of the connection between types and degrees of perturbation and the typology of laryngeal disorder that a combination of acoustic analysis, fiberoptic laryngoscopic inspection and histological examination can show. The potential usefulness of perturbation measures for assessing laryngeal pathology seems beyond doubt. The nature of the approaches to perturbation that will be the most profitable for patients suffering from laryngeal pathology remains to be established.

Scientist from various fields has developed a variety of methods and tools for the detection of normal as well as characteristic features in various vocal pathologies . This has provided the physician and speech therapist with several tradition tools for acoustic analysis. Various cases of acoustic analysis of normal and pathological voices are reported [1-4]. Acoustic methods have the potential to provide quantitative techniques for clinical assessment of laryngeal and vocal tract function. Though several methods like laryngoscopy, glottography, electromyography, stroboscopy and acoustic analysis [5] currently exists for laryngeal and vocal tract research, acoustic analysis have added advantage over other methods because of its nonintrusive nature. Diagnosis of voice pathologies is mainly done using either subjective technique like evaluation of voice quality by the clinician or invasive methods like laryngeoscopical techniques. Several quantitative measures of voice quality assessment [6 - 9] are proposed in the recent years which help in the documentation of evolution of the pathological condition. Such measures can prove to be useful for application in fields like preventive medicine and telemedicine.

Acoustic tools are commonly used by speech clinicians, such as surgeons and speech therapists for recording changes in acoustic pressure at the lips or inside the vocal tract. These tools [10], amongst others, can provide potentially objective measures of voice function. Although acoustic examination is only one tool in the complete assessment of voice function, such objective measurement has many practical uses in clinical settings, augmenting the subjective judgement of voice function by clinicians. These measures find uses, for example, in the evaluation of surgical procedures, therapy, differential diagnosis and screening [10, 11], and often augment subjective voice quality measurements, for example the GRB (Grade, Roughness and Breathiness) scale [12]. These objective measures can be used to portray a 'hoarseness" diagram for clinical applications [13], and there also exists a variety of techniques for automatically screening for voice disorders using these measures [14 - 16].

Traditionally voice signals has been modeled as a linear process and therefore, most of the tools are based on linear system theory. Most conventional linear time series analysis methods [17, 18] implicitly assume that the data come from a linear dynamical system, perhaps with many degrees of freedom and some added noise. Thus the variation is assumed to be a superposition of sine waves or exponentials that grow or decay in time. Most commonly used linear methods to characterise the system dynamics are autocorrelations, Fourier analysis and power spectrum representation. For stationary data with inherent periodicities, Fourier analysis [19] turned out to be extremely useful and this lead to the development of signal processing era in all experimental data. Signal processing continued to gain importance with the growth of electronic industry and became extremely useful with the invention of Fast Fourier Transform computer program [20]. Spectral analysis saw another fantastic leap with the introduction of wavelets in the mid 1980s [21]. With the invention of information theory by Shanon and Weaver [21] time series could be understood in terms of symbolic dynamics.

In short linear methods interpret all regular structure in a data set as a dominant frequency, as linear correlations. This means that the intrinsic dynamics of the system are governed by the linear paradigm that small causes lead to small effects. Since linear equations can only lead to exponentially growing or periodically oscillating solutions, all irregular behaviour of the system has to be attributed to some random external input to the system. [22].

The most important vocal acoustic parameters for clinical use are measurements of acoustic spectrography, fundamental frequency (F0), Harmonic-to-noise ratio(HNR), vocal extension profile, and perturbation index - jitter and shimmer [23] , zero crossing rate , wavelets.

1. **Spectrogram**

According to [24] fundamental frequency is determined physiologically by the number of cycles that the vocal folds make in a second, and they are the natural result of the length of these structures. During sound production the characteristic features of sound are produced by the interaction and complex interplay of frequencies coming from the vocal tract and the nasal tract.

The common acoustic method for speech analysis is the sound spectrogram which gives a graphical display of the time varying spectral characteristics of the speech signal [25, 26] The spectrogram is fundamental to the analysis of speech sounds [27, 28]. The particular arrangement of the frequency components in a phoneme is a strong indicator of the associated phonetic category [27]. Under certain restrictions, similar and related analysis of speech sounds produced by patients can be a valuable aid to the diagnosis and progress monitoring in the course of medical treatment for voice disorders [28]. This device produces a two-dimensional pattern called a **spectrogram** in which the vertical dimension corresponds to frequency and the horizontal dimension to time. The darkness of the pattern is proportional to signal energy. Therefore, the resonance frequencies of the vocal tract show up as dark bands in the spectrogram.

Voiced regions are characterized by a striated appearance due to the periodicity of the time wave form, while unvoiced intervals are more solidly filled in. The spectrogram is basically a series of power spectra for small, consecutive time sections of the speech signal. The width of the time window for the Fourier transform has a great effect on the characteristics of the spectrogram produced. A long time window produces a narrow band-pass filter which allows the harmonic structure to be seen in spite of blurring the time definition. The spectrogram can be used to identify individual segments of speech, such as phonemes, each of which has its own spectral structure that can become clear to the trained eye. Furthermore the spectrogram can also be used to give information on the frequency of the glottal closure of the signal seen on the wide-band spectrogram as the reciprocal of the time period seen between the vertical lines. The spectrogram can also be used to identify the formant frequencies which are the dominant frequencies in the frequency spectrum.

3.2 PERTURBATION ANALYSIS:

The attempt to discover objective acoustic and physiological ways of characterizing laryngeal waveforms in terms of perturbation parameters has a history of over 30 years, beginning with the pioneering work of researchers such as Moore and von Leden [29], Moore [29], Lieberman [30], and Michel [31]. From this base, the topic of perturbation has attracted many contributions from speech science, signal processing, laryngology, and speech pathology [32 - 63]

A somewhat comparable approach is taken by Ludlow et al. [64], where perturbation is represented as a deviation from a locally calculated trend in terms of the difference between a given period and the average of the periods two cycles to the left and right. Alternative characterizations of perturbation have tended to focus on the relationship between adjacent phonatory cycles; for instance, in terms of how often the difference in value of adjacent cycles tends to change its algebraic sign (e.g., Hecker and Kreul's directional perturbation factor) [65]. Another approach is Horii's [47] jitter ratio, which calculates the average magnitude of the differences between adjacent periods. A more computationally intensive approach is to examine the serial correlation between periods as done by Baken [66]; Iwata and von Leden [67] and Iwata [68] explore the interesting possibility that different types of serial correlation may discriminate between different types of pathology.

The analysis of period and amplitude values from continuous speech is more difficult than from sustained monotone vowels because the detecting algorithm must examine a variety of signal structures produced by interactive segmental effects of the dynamic movements of the articulators, together with multiple voicing onset and offsets [69]. It is probably also thereby more subject to the risk of artifactual distortion. If continuous speech data are used for perturbation analysis, then a minimum duration of speech material is required to stabilize long-term measurements of perturbation (e.g., the mean and standard deviation of each perturbatory parameter). In an experiment reported by Hiller [70], it was found that a 40-second sample of read speech provided relatively stable long-term speaker-characterizing parameters of perturbation for healthy male and female speakers. This finding is in general agreement with the results of previous studies of the long-term features of the voice (in particular for long-term characteristics of fundamental frequency) studies described by Hiller [70] and Baken [66] provide the technical basis for the discussion that follows on the applications of automatic systems for quantifying the perturbatory characteristics of phonation. Emphasis in the discussion will be placed on noninvasive acoustic methods, but some comment will also be made on selected physiological methods.

Jitter refers to a short term(cycle-to-cycle) perturbation in fundamental frequency of voice. Acoustically, perturbation can characterize the laryngeal waveform in both the time domain and the frequency domain. In the time domain, dysperiodic perturbation of the cycle-to-cycle variation in fundamental frequency is called jitter. Some of the early investigators (e.g., Lieberman [30], 1961) [69] displayed speech waveforms oscillographically and saw that no two periods were exactly alike. Perturbation can acoustically characterize the laryngeal waveform in both the time domain and the frequency domain. [69]. The fundamental frequency appeared jittery; hence, the term jitter. Jitter is affected mainly because of lack of control of vocal fold vibration, which are related with presence of noise at emission and breathiness .

Small and apparently random perturbations of the vocal waveform's period and amplitude (referred to as jitter and shimmer, respectively) have received a great deal of attention [69]. John Laver et. al. [69] suggests that automatic extraction methods make assessment of vocal perturbation a potential part of routine clinical assessments [69]. Careful attention will have to be paid to several technological issues, particularly to the quality of recorded voice samples used for this kind of analysis. Better databases must be developed, and the underlying voice physiology requires clarification.

Haydee et. al [71] carried out analysed vocal characteristics related to intensity and fundamental frequency and their perturbations indices - jitter and shimmer, in children with phonological disorder [71]. The Computer Speech Lab was used to record and perform acoustic analyses of the vowels /a/, /e/, /i/, through the vocal parameters: fundamental frequency, intensity, jitter and shimmer. The results found were that frequency F0 for vowel /e / was smaller, on an average , in the phonological disorder group and it was higher Hz in the Control group. No differences between the groups were found regarding the averages of jitter and shimmer. The results found in the present study point to the fact that these children with phonological disorders compared to children without disorders do not present any abnormality that affects the vocal folds, either muscle or neural activity involved with phonation, either lesions that may cause increase in aperiodicity of vocal fold vibration, which reflect the increased values of jitter [71].

The study also indicated that the characteristics such as reduction of glottic resistance, vocal fold mass lesions and greater noise at production, factors that could lead to affections of shimmer values, were not necessarily found.

Fundamental frequency FO and Harmonic to noise ratio

The spectral shape and level of the noise in the voice signals was estimated with a ***cepstrum***-based technique described by De Krom for laryngeal voices [72]. Various features of a voice spectrum are disentangled in the corresponding cepstrum (the inverse Fourier transform of the log spectrum) and can then be manipulated separately. Overall spectral shape is represented at the low end of cepstrum. The harmonic structure gives rise to a few equidistant peaks in the cepstrum. The noise contributes to various cepstral parts: the envelope to the low end and the fine structure mainly to the higher regions. By application of a comb filter to the cepstrum, energy related to the harmonic structure in the spectrum can be removed. After a reverse transformation and a level correction, an estimated noise spectrum can be obtained. An example is an acoustical tracheoesophageal (TE)-voice spectrum and its estimated noise spectrum are shown Fig. 3.1. From these spectra the harmonics-to-noise ratio is calculated in the regions of first (HNR *F1*) and second formant (HNR *F2*), by subtraction of noise and signal intensity below 700 Hz and between 700 and 2300 Hz, respectively.

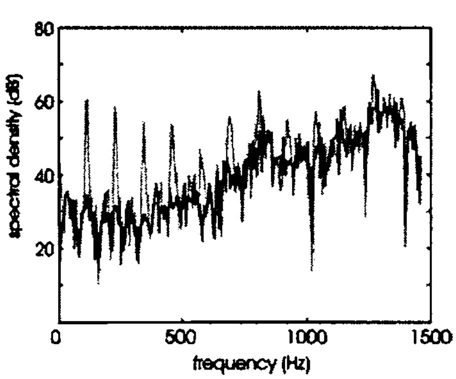


Fig. 3.1 Spectrum of a TE voice (thin, dotted line) with its estimated noise spectrum (heavy line). The lower harmonics are clearly visible and show the periodicity of the

voice but for increasing frequency the level of the noise increases and above 1 kHz the noise spectral density is larger than the level of the harmonics.

When measuring fundamental frequency in signals with a low harmonics-to-noise ratio the result may become very unstable due to random fluctuations of the noise. To reduce the influence of noise the calculations were conducted in the spectral domain: not only the fundamental but also the positions of the higher harmonics were used by calculating the largest common divider of their frequencies.

This method was first described by Schroeder (1968) [72], applied by Festen et al. (1996) [72], and is illustrated in Fig. 3.2. The spectrum of the signal was added to versions of the same spectrum compressed along the frequency axis by small integer numbers. Compression by a factor of two brings the second harmonic at the ***F0***-position, compression by a factor of three the third harmonic, etc. The sum of these compressed spectra is called harmonic product spectrum and it generally shows a clear peak at the fundamental frequency, even for noisy signals. The vocal intensity is determined as the sound pressure level in dB.

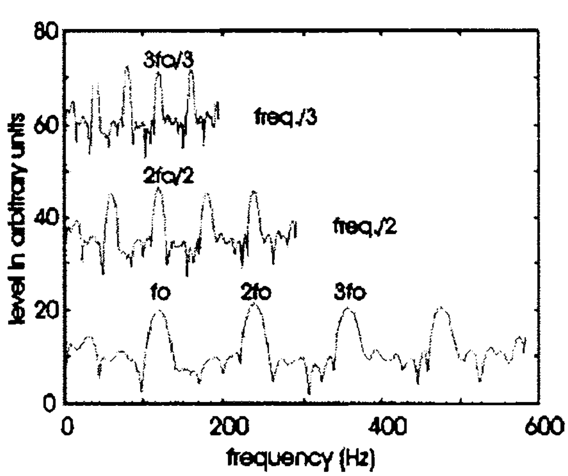


Fig. 3.2 The lower trace shows an arbitrary harmonic voice spectrum. The middle and upper trace show the same spectrum but compressed by factors two and three, respectively. This compression of the frequency scale brings successive higher harmonics

at the position of F0. An addition of these spectra gives a strong peak at the fundamental frequency because here peaks in the various spectra coincide.

Pitch Amplitude and coefficient of excess

Two parameters used to determine the voice signal noise quantity are the deterministic ***harmonic to noise ratio (HNR)*** and the ***coefficient of excess (EX)*** that evaluate the noise from a statistical point of view [73]. The latter measures differences in shape of the distribution of the residue signals.

Pitch Amplitude gives a normalized measure of the amplitude of the pitch period peak of the residue signal autocorrelation function. This parameter is high for signals with clearly defined pitch period such as voiced sounds (vowels). For breathy vowels of pathological speakers, the PA is low because the signals have weak periodicity [73].

Parameters like jitter and shimmer evaluate perturbation or noise content in the voice signal.

Shimmer

Shimmer refers to the variation of the amplitude of sound wave, or intensity of vocal emission and it is the Perturbation of the cycle-to-cycle amplitude of successive laryngeal pulses [69]. Shimmer is affected mainly with reduction of glottic resistance and mass lesions in the vocal folds, which are related with presence of noise at emission and breathiness [69]. In the frequency domain, a frictional element can add spectral noise to the laryngeal waveform. Jitter and shimmer contribute to the rough perceptual effect usually called harshness in acoustory perspective. Spectral noise alone is associated with the auditory effect called whisperiness, and voices showing this effect can be called whispery voices. When spectral noise is added to jitter and/or shimmer, this has the auditory effect of adding whisperiness to harshness, and the composite quality is usually called hoarseness, or hoarse voice. Physiologically, the frictional element in whispery and hoarse voices is caused by incomplete glottal closure [69]. The incomplete closure can be the result either of individual, habitual adjustments of the phonatory muscle system, or of mechanical intrusion into the glottis of obstructions such as vocal nodules, polyps, and other types of growths, or of paralysis of one or both vocal folds. Severe perturbations are almost always signs of either pathological or functional disorder, but slight perturbations are evident in all speaking voices, especially at the borders of voicing episodes. An important consideration in the ability of acoustic techniques to register perturbation data is therefore the sensitivity of the acoustic technique itself.

[69] showed that a simple bivariate plot for a directional perturbation factor in a shimmer parameter against mean fundamental frequency for a control group of 63 male speakers versus a group of 55 male speakers known to suffer from a variety of laryngeal pathologies. The results for the control group are expressed partly in terms of an ellipse surrounding two standard deviations of the data on the two parameters. The boundary of the ellipse can be used as a practical screening threshold for the detection of vocal pathology. This can be seen from the fact that only 9.5% of the control speakers fall outside the ellipse, and therefore register as false positives, whereas 90.1% of the known pathological speakers fall outside and are therefore successfully detected as probable pathological [69]. For a similar treatment of a group of 54 pathological female speakers versus a control group of 58 female speakers presumed healthy, 79.6% of the pathological speakers were successfully detected by this method, at a cost of 10.3% false positives.

Studies [74] used to analyze sustained vowels generated by patients before and after surgical excision of vocal polyps by Yu Zhanga and Clancy McGilligan showed that jitter shows a significant post surgical decrease. Uloza [75] found that jitter and shimmer are useful for studying the effects of surgery on vocal function; however, Zeitels et al. found that there was not a significant difference in jitter measurements after surgical excision of vocal fold lesions, although there was a significant postsurgical decrease in shimmer measurements. Methodological issues have been raised regarding jitter and shimmer because of the sensitivity of these two parameters to variations in recording systems, analysis systems, and extraction algorithms [75]. In addition, systems employed for these perturbation parameters cannot reliably analyze strongly aperiodic signals [76, 77]. It has been suggested that jitter and shimmer measurements are reliable for voice analysis only when their values are less than 5% [78, 76] and may be unreliable when large instabilities in voice waveforms are observed [77]. Since jitter and shimmer only represent reliable parameters for nearly periodic voice signals under small perturbation conditions, seeking complementary objective measures capable of analyzing aperiodic voices and assessing the effects of surgery is important.

On the other hand, Zeitels et al. [79] found that there was not a significant difference in jitter measurements after excision of benign vocal-fold lesions, although there was a significant decrease in shimmer measurements. The study of Zeitels et al. and other studies suggest that these perturbation measures should be applied with caution, particularly when a periodic or chaotic voice signals are being analyzed. A similar conclusion was arrived at via their recent experiments with excised larynges [80]. Jitter and shimmer are recognized as inadequate for analysis of far from-periodic voices. The voices from patients showed a statistically significant postsurgical decrease in jitter, but not in shimmer.

[81] studied vocal jitter and shimmer in stuttering. The purpose of this study was to test for the presence of significantly different jitter and shimmer during vowel production in the fluent phonations of stutterers compared to nonstuttering matched controls.

Vocal jitter and shimmer measures of the fluent phonations of 14 stutterers, 12 male and two female, were compared with jitter and shimmer measures of a group of nonstutterers matched for age and sex. Each subject phonated four vowels nine times in random order. Each phonation was sustained for at least 5 sec and was tape-recorded. The mid-3-set portion of each recorded vowel phonation was subjected to jitter and shimmer analyses. Measures for stutterers were larger in both instances. Significant differences between stutterers and nonstutterers were obtained for shimmer measures. Differences on jitter measures were not significant. High variability in the stuttering group accounted for the nonsignificant finding in jitter measures and, in general, indicated heterogeneity among the stutterers. Findings led to the tentative conclusion that in fluent sustained phonation, stutterers demonstrate less stable control of respiratory-laryngeal dynamics than nonstutterers.

Vocal jitter and shimmer are cycle-to-cycle variations in frequency and amplitude, respectively, of the quasiperiodic glottal tone. Jitter and shimmer are acoustic measures of these vocal perturbations obtained from sustained vowel phonations, which have proved useful in describing voice characteristics of normal and pathologic speakers (Deal and Emanuel, [82] Horii, [82 - 85] Smith, et al., [86]. Lower indices of magnitude on either jitter or shimmer indicate less vocal perturbation and greater stability in the tine motor control of phonatory behavior. If the magnitude of vocal perturbation, either jitter or shimmer, in the fluent phonatory behaviors of stutterers was shown to be significantly greater than that of nonstutterers, this would provide additional support to the hypothesis that stutterers may demonstrate generally less competent neurophysiologic regulatory control over their peripheral mechanisms of phonation and respiration [87].

The jitter and shimmer analysis of the recordings was performed using a computer program SEARP developed by Horii [88, 83, 84]. The mid-3-see segments of accelerometric signals of vowel phonations recorded on magnetic tape were subjected to the SEARP analysis to obtain the jitter and shimmer data. The program used a peak-picking method to identify individual periods. Jitter is reported in percent. Jitter in percent was obtained by averaging absolute differences between periods in milliseconds from one cycle to the next, dividing the result by the average period and multiplying this value by 100. Shimmer in dB was defined as the average decibel difference between peak amplitudes of consecutive cycles.

A significant difference was found between nonstutterers and stutterers on measures of shimmer. However, differences in jitter between the two groups were not found to be statistically significant. This finding was unexpected in that Horii (1980) [84] had reported a significant, though modest, correlation between jitter and shimmer ($r = 0.47$, $p < 0.001$).Because of the supposed relation between the two phenomena, it was expected that either significant differences between stutterers and nonstutterers would be found for both jitter and shimmer measures, or that no significant difference would be found for either measure. Means of both jitter and shimmer measures of the stutterers were larger than those of the nonstutterers, indicating that the sustained phonations of the stutterers were less stable than those of the nonstutterers in terms of both vocal frequency and intensity. The significant difference between stutterers and nonstutterers with respect to shimmer in dB and the differences observed with regard to jitter, coupled with the relation between the two phenomena, led to the probable conclusion that steady-state phonations of stutterers are different from those of nonstutterers. The direction of this difference suggests that stutterers have less stable neuromuscular control over the events regulating the aerodynamics of the laryngeal and respiratory systems during sustained fluent vowel articulations than nonstutterers [81]. Steady-state, sustained phonation involves an even maintenance of such forces as vocal fold tension, mass, length, and subglottic pressures, while it also maintains the supralaryngeal articulatory adjustments required for production of the vowel. Differences observed between the two groups suggest that nonstutterers are better able to control these forces than stutterers. The large standard deviation of jitter measures of stutterers accounts for the lack of significance between the two groups. In general, the higher variability revealed by the standard deviations of both jitter and shimmer measures of the stutterers suggests a greater heterogeneity in this population.

In order to arrive at a conclusive result this type of study on larger homogeneous samples of stuttering speakers are to be carried out.

Mean jitter measures of each of the four vowels were very similar within groups. This was also true for shimmer measures. Given that main effects for vowels and group interaction effects with vowels were not significant, it would seem appropriate to suggest that future studies need not include production of a variety of vowels for analysis.
This is desirable, because the amount of effort and time required to obtain jitter and shimmer measures is considerable. This finding also indicates that the significant shimmer differences were due primarily to generalized factors underlying the

mechanisms of phonation and are relatively independent of the accompanying articulatory adjustments necessary to produce the English vowels /i a u /.

Zero Crossing Rate

A zero crossing is a point where an audio waveform crosses the zero-line from negative to positive. Zero crossing graph displays the count of zero crossing over a brief time interval. The oscillation around the zero line is a measure of high frequency sound context. The higher the zero crossing value, the greater the higher frequency component in the audio-dates. Fricatives generally have a high zero-crossing value [89]. Zero Crossing Rate (ZCR) is defined as the number of time-domain zero-crossings within a defined region of signal, divided by the number of samples of that region.

The zero-crossing rate (ZCR) is obtained by counting the number of times the signal changes sign during one frame interval as given in the equation below:

$$ZCR_t = \sum_{n=t}^{t+N-1} \left| Sgn\{x(n)\} - Sgn\{x(n-1)\} \right| w(n-t)$$

ZCR gives a rough estimate of the frequency content of the speech signals, i.e. a sinusoidal signal of frequency F gives an average ZCR of 2F s^{-1}. ZCR has been used by Rabiner [90] to improve the end-point detection algorithm. It has also been used to distinguish between voiced and unvoiced sounds, i.e. mean ZCR for voiced sounds is less than 1500 s^{-1} while for unvoiced sound it is more than 5000 s^{-1} [90].

Speech signals are broadband signals and the interpretation of average zero-crossing rate is therefore much less precise. However, rough estimates of spectral properties can be obtained using a representation based on the short-time average .

The ZCR of the time domain waveform is one of the most indicative and robust measures to discern voiced speech. It has widely used in practice as a strong measure to discern fricatives from voiced speech [89]. The ZCR is simply the count of crossing the zero throw fixed window size. It is said to occur if successive samples have different algebraic signs.

Zero crossing graph displays the count of zero crossing over a brieftime interval. The oscillation around the zero line is a measure of high frequency sound context. The higher the zero crossing value, the greater the higher frequency component in the audio-dates. Fricatives generally have a high zero-crossing value. The zerocrossing rate of speech signals is a parameter for many speech recognition purposes. [90] studied the speech signal as an analytic signal taking its zerocrossing rate into consideration [90]. The zero crossing rate is one of the characteristic parameters in speech signal analysis. Varada. S and Sankar. R proposed a method for finding the end point of the speech signal using zero crossing rate and energy [92]. Digit recognition using energy envelope and zerocrossing rate is proposed by [93]. [94] proposed a new feature extraction method based on zerocrossings with peak amplitudes for robust speech recognition in noisy environments [94] . [95] present the result of experimental investigations that provides an interesting perspective on the relative importance of zerocrossing interval for speech perception [95].

The zero crossing parameters of the speech signal are important for speech modeling and recognition purposes. Zerocrossing rate is a reliable parameter for many speech recognition purposes. But the variation of zero crossing intervals of the speech signal and its uses in parmetric estimation for recognition purpose is not yet carried out. Also short time energy of speech signal is utilized for boundary value detection and other analysis. But a steady average energy in the zero crossing interval for speech recognition has not yet been carried out.

Zero crossing information of the speech signal is a perceptually meaningful parameter because param eters like formant frequency can be extracted using zerocrossing interval information of noise corrupted speech. T.V. Sreenivas and Niederjohn proposed a new method of finding the formant frequency of the noise corrupted speech signal using statistical properties of zerocrossing intervals [95]. The method is compared with currently popular spectral analysis technique based on singular value decomposition and found to provide generally better resolution and lower variance at low signal to noise ratio (SNR) .

In the context of discrete-time signals, a zero crossing is said to occur if successive samples have different algebraic signs. The rate at which zero crossings occur is a simple measure of the frequency content of a signal. Zero-crossing rate is a measure of number of times in a given time interval/frame that the amplitude of the speech signals passes through a value of zero, Fig.3.3 and Fig.3.4 . Speech signals are broadband signals and interpretation of average zero-crossing rate is therefore much less precise.

However, rough estimates of spectral properties can be obtained using a representation based on the shorttime average zero-crossing rate [89].

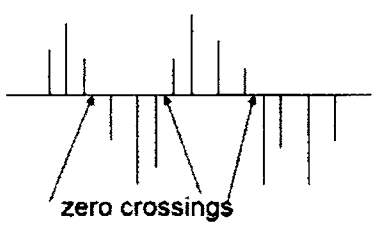


Fig. 3.3. Definition of zero-crossings rate

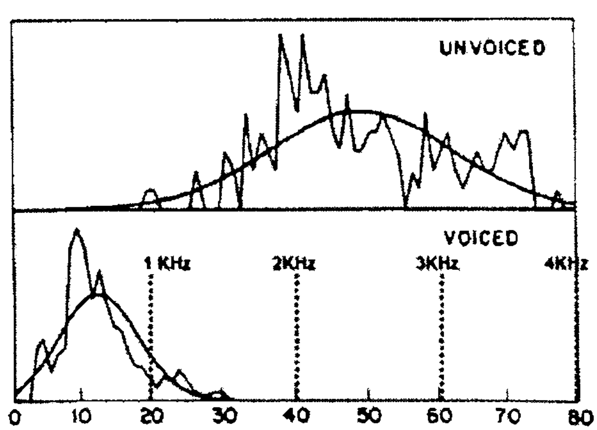


Fig.3.4. Distribution of zero-crossings for unvoiced and voiced speech [89]

The model for speech production suggests that the energy of voiced speech is concentrated below about 3 kHz because of the spectrum fall of introduced by the glottal wave, whereas for unvoiced speech, most of the energy is found at higher frequencies.

Since high frequencies imply high zero crossing rates, and low frequencies imply low zero-crossing rates, there is a strong correlation between zero-crossing rate and energy distribution with frequency. A reasonable generalization is that if the zero-crossing rate is high, the speech signal is unvoiced, while if the zero-crossing rate is low, the speech signal is voiced [89].

Phonation Number: It is a qualitative parameter .The phonation number is defined as the number of single speech fragments separated by silent pauses. The phonation number was equal to the number of pauses. Kuniszyk-Jo'z'kowiak [96] have found that the total phonation time (i.e., the sum of all the phonation times in an utterance) is longer in fluent utterances than in nonfluent ones. The total phonation time depends on the sound level at which it is measured.

The aim of the work [96] focuses to find acoustic (objectively measurable) dimensions that characterize utterances of fluent speakers and of stutterers. Their earlier work [96] had stated that there were significant differences in courses of dependence of phonation number on sound level and in the statistical distributions of the times of phonations and pauses between them. The work focuses on acoustic characteristics of utterances that distinguishes stuttered and fluent speech. The interdependence of phonation number (the number of times a given sound level was crossed) and sound intensity level was determined in this study. Also the average distributions of phonation times and pause durations as a function of sound intensity level were examined. The data suggest that differences in speech envelopes of stutterers and non-stutterers may be used to evaluate the degree of speech non-fluency. The result support the hypothesis that the cause of stuttering is a lack of synchronization between laryngeal functioning and vocalizing activities. Their result states that transitional states between vowels and consonants are either prolonged or more abrupt than normal.

In this study on the basis of the course of speech envelopes the number of times the signal amplitude exceeded the preset criterion was calculated called the number of phonation.

Wavelet Analysis

Wavelet Analysis is the analysis based on time frequency localization of the signal which can be advantageous in nonstationary problems. Wavelet expansions often provide very concise signal representation and thereby can simplify subsequent nonlinear analysis and processing. Wavelet analysis is a specialist tool for deriving an indepth understanding of underlying components of complex curvilinear data. Wavelet analysis is based on a windowing technique employing regions of varying size and can provide the time and frequency information inspite of overcoming the resolution limitations of the Fourier transform(STFT). Long-time intervals are used to provide more precise information at low frequencies; shorter intervals are used to extract characteristics of high-frequency components of a signal. The wavelet transform is a powerful signal-processing tool that has been successfully used in the analysis of non-stationary data from a wide range of physical process, for example, from engineering systems and in meteorology, that exhibit multiscale features [97 - 99]. It has also been used extensively for speech compression and recognition [100, 101] and for a number of studies on speech waveform analysis [102 - 104].

One of the most well-known tools for signal processing is the Fourier transform. This breaks down a signal into its constituent sinusoids of different frequencies. One of the drawbacks of Fourier analysis is that in transforming to the frequency domain, time information is lost. For stationary signals, ie., those that do not change much over time, this is not a serious drawback. However, for nonstationary data, spectral techniques that retain information from the time domain are more appropriate. Nonstationary waveforms are commonly encountered in experimental acoustics. Nodel prize winner Dennis Gabor adapted the Fourier transform in order to analyse only a small section of the signal, known as the spectral window, at a time. This so-called short-time Fourier transform (STFT) maps a signal into a two-dimensional fuction of time and frequency, so one can determine approximately when an event of a particular frequency occurs. However, although STFT has proved useful in numerous application, it is ill suited for the study of signals where the frequency content ranges over several orders of magnitude as the size of the window is the same for all frequencies.

This paper [ref above] study was to adapt wavelet analysis as a tool for discriminating speech sample taken from healthy subjects across two biological states. Speech pressure waveforms were drawn from a study on effects of hormone fluctuations across the menstrual cycle on language functions. Speech samples from the vowel portion of the syllable ‘pa’, taken at the low- and high-hormone phase of the menstrual cycle, were extracted for analysis. Firstly they applied fourier transform to examine the fundamental and format frequencies. Wavelet analysis was used to investigate spectral differences at a more microbehavioural level. The results found are that wavelet coefficients for the fundamental frequency of speech samples taken from the high-hormone phase had larger amplitude than those from the low-hormone phase. This study provided evidence for differences in speech across the menstrual cycle that affected the vowel portion of syllables [105]. This study provided a new tool for examination of behavioural differences in speech linked to hormonal variation.

The continuous wavelet transform (CWT) of a signal is defined as the sum over all time of the signal multiplied by scaled, shifted versions of the mother wavelet. Thus mathematically, the wavelet transform $C_{a,b}(t)$ of a signal f(t), where t represents any independent variable , is defined as

$$C_{a,b}(t) = \frac{1}{|a|^{1/2}} \int_{-\infty}^{\infty} f(t)\psi(\frac{t-b}{a})dt$$

where the function $\psi(t)$ is a mother wavelet and the real number a ($a \neq 0$) and b denote the scaling and translation respectively. $C_{a,b}(t)$ are known as the wavelet coefficients. The constituent wavelets of the original signal can be found by taking the product of each coefficient with appropriately scaled and shifted wavelet. This facilitates the detailed analysis of the signal over a range of frequencies. The scale a is related in a broad sense to frequency by:

$$\mathrm{Fa} = \frac{\Delta\ t x F\ c}{a}$$

Where Fa is the pseudo-frequency in Hz corresponding to the scale a in Hz, Fc is the center frequency in Hz, fc is the center frequency in Hz of the mother wavelet and Δt is the sampling period. For the case of the db 10 mother wavelet, the center frequency is 0.68421 Hz.

3. 3 NONLINEAR TECHNIQUES:

Despite the limited technological success of the linear model in several applications, such as speech coding, synthesis and recognition, there is strong theoretical and experimental evidence [106, 107 - 112] for the existence of important nonlinear aerodynamic phenomena during the speech production that cannot be accounted for by the linear model. Various factors contributing to nonlinearities in speech production are (i) turbulent air flow through vocal tract, (ii) coupling produced between different parts of vocal tract (iii) neuro muscular response to stimuli [106].

The investigation of speech nonlinearities can proceed in at least two directions: (i) numerical simulation of the nonlinear differential equations governing the 3-D dynamics of the speech air flow in the vocal tract [108, 110]. (ii) development of nonlinear signal processing systems suitable to detect such phenomenon and extract related information[112 - 115]. The phonetory system is time varying and consequently the speech signal is nonstationary. This can be clearly understood if one closely observes the amplitude of the speech samples.

Nonlinear time series analysis aims at understanding the dynamics of a system using the time series of a single available variable. Chaos theory states that random inputs are not the only source for irregular output of a system: nonlinear chaotic systems with purely deterministic equations of motion can produce very irregular data. Extensive research conducted on nonlinear dynamical systems over recent years have proved that conventional time domain and frequency domain approaches to real world systems are far from optimal.

Many quantities in nature fluctuate in time. Examples are the stock market, the weather, seismic waves, sunspots, heartbeats, and plant and animal populations. Until recently it was assumed that such fluctuations are a consequence of random and unpredictable events. With the discovery of chaos, it has come to be understood that some of these cases may be a result of deterministic chaos and hence predictable in the short term and amenable to simple modeling. Many tests have been developed to determine whether a time series is random or chaotic, and if the latter, to quantify the chaos. A positive maximal Lyapunov exponent derived from the time series, expresses irregular deterministic behavior, which is termed chaotic [116 - 119], whereas dynamical systems with solely non-positive exponents are usually referred to as regular. If chaos is found, it may be possible to improve the short-term predictability and enhance understanding of the governing process of the dynamical system. We mean by talk of a 'dynamical system': a real-world system which changes over time.

While there is a long history of linear time series analysis, nonlinear methods have only just begun to reach maturity. When analyzing time series data with linear methods, there are certain standard procedures one can follow, moreover, the behavior may be completely described by a relatively small set of parameters. Linear methods interpret all regular structure in the data set, such as dominant frequency, as linear correlations. This means, in brief, that the intrinsic dynamics of the system are governed by the linear paradigm that small causes lead to small effects. Since linear equations can only lead to exponentially growing or periodically oscillating solutions, all irregular behavior of the system has to be attributed to some random external input to the system [121]. A brief comparison between linear and nonlinear methods [120] can be found in Table 3.1. Chaos theory has taught us that random input is not the only possible source of irregularity in a system's output: nonlinear, chaotic systems can produce very irregular data with purely deterministic equations of motion plus Table 3.1.

Now, Nonlinear Time Series Analysis (NTSA) is the study of the time series data with computational techniques sensitive to nonlinearity in the data. This was introduced by the theory of chaos to characterize the source complexity [121]. The NTSA theory offers tools that bridge the gap between experimentally observed irregular behavior and deterministic chaos theory. It enables us to extract characteristic quantities of a particular dynamical system solely by analyzing the time course of one of its variables [121-123]. In theory, it would then be possible to collect temperature measurements in a particular city for a given period of time and employ nonlinear time series analysis to actually confirm the chaotic nature of the weather. Despite the fact that this idea is truly charming, its realization is not feasible quite so easily. In order to justify the calculation of characteristic quantities of the chaotic system, the time series must originate from a (i) stationary, (ii) deterministic system.

A deterministic dynamical system is one for which there is a rule, and , given sufficient knowledge of the current state of the system one can use the rule to predict future states; i.e. the future state x_{n+t} can be determined precisely from the current state x_n at any instance n for some value of $t > 0$, by applying the deterministic rule for the dynamical system. The other important requirement before attempting to do quantitative analysis is identification of stationarity of the dynamical system. Dynamical systems that are not stationary are exceedingly difficult to model from time series. Unless one has a priori knowledge of the structure of the underlying system, the number of parameters will greatly exceed the number of available data [124]. It may be noted here that the definition for stationarity is not the same as for linear systems: a linear system is said to be stationary if all its moments remain unchanged with time. A non-stationary system is defined as a one which is subject to temporal dependence based on some outside influence. If we extend our definition of the system to include all outside influences, the system is stationary.

Table 3.1. Comparison of linear and nonlinear signal processing techniques.

Linear signal processing	Nonlinear signal processing
Finding the signal Separate broadband noise from narrow band signal using spectral characteristics. Method: Matched filter in frequency domain.	**Finding the signal** Separate broadband signal from broad band noise using deterministic nature of the signal Method: Mainfold decomposition or statistics on the attractor .
Finding the space Use Fourier space methods to turn difference equations in to algebraic forms $x(t)$ is observed $X(f) = \sum x(t)\, e^{j2\pi f}$ is used	**Finding the space** Time lagged variables from coordinates for a reconstructed state space in m dimensions . $X(t) = [x(t), x(t+\tau), x(t+2\tau), \ldots\ldots x(t+(m-1)\tau)]$ where τ and m are determined by false nearest neighbours and average mutual information.
Classify the signal • Sharp spectral peaks • Resonant frequencies of the system	**Classify the signal** • Lyapunov Exponents • Fractal Dimension measures • Unstable fixed points • Recurrence quantification • Statistical distribution of the attractor
Making models, predict : $x(t+1) = \sum \alpha_k\, x(t-k)$ Find parameters α_k consistent with invariant classifiers- location of spectral peaks	**Making models, predict :** X(t) $\rightarrow X(t+1)$ $X(t+1) = F[X(t), a_1, a_2 \ldots a_p]$ Find parameters a_j consistent with invariant classifier - Lyapunov Exponents,fractal Dimension

Methods and Implementation

Let us suppose that we have a dynamical system which is both stationary and deterministic. To apply the nonlinear time series methods, the dynamics of the system are to be presented in a phase space. When the equations that govern process dynamics are not known, the phase space is reconstructed from a measured time series by using only one observation. The most basic step in this procedure is to rely on a time delayed embedding of the data, i.e. attractor reconstruction. For this purpose, we have to determine the proper embedding delay and embedding dimension. There exist two methods, developed in the framework of nonlinear time series analysis, that enable us to successfully perform these desired tasks. The average mutual information method [125] yields an estimate for the proper embedding delay, whereas the false nearest neighbor method [126] enables us to determine a proper embedding dimension.

In the following sub sections the methods of phase space reconstruction, average mutual information method and false nearest neighbors method are discussed.

Phase Space Reconstruction – Taken's Embedding Theorem

The basic idea of the phase space reconstruction is that a signal contains information about unobserved state variables which can be used to predict the present state. Therefore, a scalar time series $\{x(i)\}$ may be used to construct a vector time series that is equivalent to the original dynamics from a topological point of view.

The problem of how to connect the phase space or state space vector of dynamical variables of the physical system to the time series measured in experiments was first addressed in 1980 by Packard et. al [127] who showed that it was possible to reconstruct a multidimensional state-space vector $X(i)$ by using time delays (or advances which we write as positive delays) with the measured, scalar time series, $\{x(i)\}$. Takens [128] and later Sauer et. al [129] put this idea on a mathematically sound footing by proving a theorem which forms the basis of the methodology of delays. They showed that the equivalent phase space trajectory preserves the topological structures of the original phase space trajectory. Due to this dynamical and topological equivalence, the characterization and prediction based on the reconstructed state space is as valid as if it was made in the true state space. The attractor so reconstructed can be characterized by a set of static and dynamic characteristics. The static characteristics describe the geometrical properties of the attractor whereas the dynamical characteristics describe the dynamical properties of nearby trajectories in phase space.

Thus, given a time series $\{x(j)\}=x(1), x(2), x(3), ...x(N)$ we define points $X(i)$ in an m-dimensional state space as

$$X(i)=[x(i),\ x(i+\tau),\ x(i+2\tau), ...x(i+(m-1)\tau)] \tag{1}$$

for $i=1,2,3,....,N-(m-1)\tau$ where i are time indices, τ, a time delay or sometime referred to as embedding delay, and m is called the *embedding dimension*. Time evolution of $X(i)$ is called a trajectory of the system, and the space, which this trajectory evolves in, is called the reconstructed phase space or simply, embedding space.

While the implementation of Eq. (1), the mathematical statement of Takens' Embedding theorem, should not pose a problem, the correct choice of proper embedding parameters τ and m is a somewhat different matter. The most direct approach would be to visually inspect phase portraits for various τ and m trying to identify the one that looks best. The word "best", however, might in this case be very subjective. In practice this approach for finding embedding parameters are seldom advised since we usually want to analyze a time series that originates from a rather unknown system. Then we would not know if the underlying dynamics that produced the time series had two or twenty degrees of freedom. It is easy to verify that the time required to check all possibilities that might yield a proper embedding with respect to various τ and m is very long. This being said, let it be a good motivation to discuss the average mutual information method and the false nearest neighbor method, which enable us to efficiently determine proper values of the embedding delay τ and embedding dimension m. Let us start with the mutual information method.

Selecting the Time Delay-Average Mutual Information Method

A suitable embedding delay τ has to fulfill two criteria. First, τ has to be large enough so that the information we get from measuring the value of x variable at time $(i+\tau)$ is relevant and significantly different from the information we already have by knowing the value of the measured variable at time i. Only then it will be possible to gather enough information about all other variables that influence the value of the measured variable to successfully reconstruct the whole phase space with a reasonable choice of m. Note here that generally a shorter embedding delay can always be compensated with a larger embedding dimension. This is also the reason why the original embedding theorem is formulated with respect to m, and says basically nothing about τ. Second, τ should not be larger than the typical time in which the system looses memory of its initial state. If τ would be chosen larger, the reconstructed phase space would look more or less random since it would consist of uncorrelated points. The latter condition is particularly important for chaotic systems which are intrinsically unpredictable and, hence, loose memory of the initial state as time progresses. This second demand has led to suggestions that a proper embedding delay could be estimated from the autocorrelation function where the optimal τ would be determined by the time the autocorrelation function first decreases below zero or decays to $1/\exp$. For nearly regular time series, this is a good thumb rule, whereas for chaotic time series, it might lead to spurious results since it based solely on linear statistic and doesn't take into account nonlinear correlations.

The cure for this deficiency was introduced by Fraser and Swinney [130]. They established that delay corresponds to the first local minimum of the average mutual information function $I(\tau)$ which is defined as follows:

$$I(\tau) = \sum P(x(i), x(i+\tau)) \log_2 \left[\frac{P(x(i), x(i+\tau))}{P(x(i))P(x(i+\tau))} \right] \qquad (2)$$

where $P(x(i))$, is the probability of the measure $x(i)$, $P(x(i+\tau))$ is the probability of the measure $x(i+\tau)$ and $P(x(i), x(i+\tau))$ is the joint probability of the measure of $x(i)$ and $x(i+\tau)$ [130]. The average mutual information is really a kind of generalization to the nonlinear phenomena from the correlation function in the linear phenomena. When the measures $x(i)$ and $x(i+\tau)$ are completely independent, $I(\tau) = 0$. On the other hand when $x(i)$ and $x(i+\tau)$ are equal, $I(\tau)$ is maximum. Therefore plotting $I(\tau)$ *versus* τ makes it possible to identify the best value for the time delay, this is related to the first local minimum.

While it has often been shown that the first minimum of $I(\tau)$ really yields the optimal embedding delay, the proof of this has a more intuitive, or shall we rather say empiric, background. It is often said that at the embedding delay where $I(\tau)$ has the first local minimum, $x(i+\tau)$ adds the largest amount of information to the information we already have from knowing $x(i)$, without completely losing the correlation between them. Perhaps a more convincing evidence of this being true can be found in the very nice article by Shaw [131], who is, according to Fraser and Swinney, the idea holder of the above reasoning. However, a formal mathematical proof is lacking. Kantz and Schreiber [121] also report that in fact there is no theoretical reason why there should even exist a minimum of the mutual information. Nevertheless, this should not undermine ones trustworthiness in this particular method, since it has often proved reliable and well suited for the appointed task.

Once the time delay has been agreed upon, the embedding dimension is the next order of business. Let us now turn to establishing a proper embedding dimension m for the examined time series.

Selecting Embedding Dimension- False Nearest Neighbors Method

In general, the aim of selecting an embedding dimension is to make sufficiently many observations of the system state so that the deterministic state of the system can be resolved unambiguously. It is best to remember that in the presence of observational noise and finite quantization this is not possible. Moreover, it has been shown that even with perfect observations over an arbitrary finite time interval, a correct embedding will still yield a set of states indistinguishable from the true state [132]. Most methods to estimate the embedding dimension aim to achieve unambiguity of the system state. The archetype of many of these methods is the so-called False Nearest Neighbors (FNN) technique [133, 134].

The false nearest neighbor method was introduced by Kennel et al. [126] as an efficient tool for determining the minimal required embedding dimension m in order to fully resolve the complex structure of the attractor, i.e. the minimum dimension at which the reconstructed attractor can be considered completely unfolded. Again note that the embedding theorem by Takens [128] guarantees a proper embedding for all large enough m, i.e. that is also for those that are larger than the minimal required embedding dimension. In this sense, the method can be seen as an optimization procedure yielding just the minimal required m. The method relies on the assumption that an attractor of a deterministic system folds and unfolds smoothly with no sudden irregularities in its structure. By exploiting this assumption, we must come to the conclusion that two points that are close in the reconstructed embedding space have to stay sufficiently close also during forward iteration. If this criterion is met, then under some sufficiently short forward iteration, originally proposed to equal the embedding delay, the distance between two points $X(n)$ and $X(p)$ of the reconstructed attractor, which are initially only a small distance apart, cannot grow further as we fix a threshold value for these distances in computation. However, if an n-th point has a close neighbor that doesn't fulfill this criterion, then this n-th point is marked as having a false nearest neighbor. We have to minimize the fraction of points having a false nearest neighbor by choosing a sufficiently large m. As already elucidated above, if m is chosen too small, it will be impossible to gather enough information about all other variables that influence the value of the measured variable to successfully reconstruct the whole phase space. From the geometrical point of view, this means that two points of the attractor might solely appear to be close, whereas under forward iteration, they are mapped randomly due to projection effects. The random mapping occurs because the whole attractor is projected onto a hyper plane that has a smaller dimensionality than the actual phase space and so the distances between points become distorted [135].

In order to calculate the fraction of false nearest neighbors, the following original algorithm is used.

Consider each vector $X(n)=[x(n),\ x(n+\tau),\ x(n+2\tau),.......x(n+(m-1)\tau)]$ in a delay coordinate embedding of the time series with delay τ and embedding dimension m. Look for its nearest neighbor $X(p)$ and $X(p)=[x(p),\ x(p+\tau),\ x(p+2\tau),.......x(p+(m-1)\tau)]$. The nearest neighbor is determined by finding the vector $X(p)$ in the embedding which minimizes the Euclidean distance $R_n=\|X(n)-X(p)\|$. Now consider each of these vectors under an $m+1$ dimensional embedding,

$$X'(n)=[x(n),\ x(n+\tau),\ x(n+2\tau),.......x(n+(m-1)\tau),x(n+m\tau)]$$

$$X'(p)=[x(p),\ x(p+\tau),\ x(p+2\tau),.......,x(p+(m-1)\tau),x(p+m\tau)]$$

In an $m+1$ dimensional embedding, these vectors are separated by the Euclidean distance $R'_n=\|X'(n)-X'(p)\|$. The first criterion by which Kennel, et. al., identify a false nearest neighbor is if

Criterion 1: $$\left[\frac{R_n'^2-R_n^2}{R_n^2}\right]^{1/2}=\frac{|x(n+m\tau)-x(p+m\tau)|}{R_n}>R_{tol} \qquad (3)$$

R_{tol} is a unit less tolerance level for which Kennel, et. al. [124], suggest a value of approximately 15. This criterion is meant to measure if the relative increase in the distance between two points when going from m to $m+1$ dimensions is large. 15 was suggested based upon empirical studies of several systems, although values between 10 and 40 were consistently acceptable. The other criterion Kennel, et. al., suggest is

Criterion 2: $$\left[\frac{R'_n}{R_A}\right]>A_{tol} \qquad (4)$$

Where A_{tol} is called absolute tolerance. This was introduced to compensate for the fact that portions of the attractor may be quite sparse. In those regions, near neighbors are not actually close to each other. Here, R_A is a measure of the size of the attractor, for which Kennel, et. al., use the standard deviation of the data. If either (3) or (4) hold, then $X(p)$ is considered a false nearest neighbor of $X(n)$. The total number of false nearest neighbors is found, and the percentage of nearest neighbors, out of all nearest neighbors, is measured. An appropriate embedding dimension is one where the percentage of false nearest neighbors identified by either method falls to zero.

The above combined criteria correctly identify a suitable embedding dimension in many cases. By now we have equipped ourselves with the knowledge that is required to successfully reconstruct the embedding space from an observed variable. This is a very important task since basically all methods of nonlinear time series analysis require this step to be accomplished successfully in order to yield meaningful results.

3.4 INVARIANT PARAMETERS :

Lyapanov Exponents

The exponential divergence of nearby trajectories is calculated by the Lyapunov exponent. It is a measure of the rate of attraction or repulsion. If two nearby trajectories on a chaotic attractor start off with a separation d0 at time t = 0, then the trajectories diverge so that their separation at time t, denoted by d(t) satisfies the expression

$d(t) = d_0 e^{\lambda t}$ (1) where λ is called the Lyapunov exponent for the trajectories.

The Lyapunov exponents provide a coordinate-independent measure of the asymptotic local stability of properties of a trajectory. In a geometry representation, we can imagine a small infinitesimal ball of radius $\epsilon(0)$ centered on a point $\Phi(0)$ in state space. Under the action of the dynamics the center of the ball may move, and the ball becomes distorted. Since the ball is infinitesimal, this distortion is governed by the linear part of the flow. The ball thus remains an ellipsoid. Suppose the principal axis of the ellipsoid at time t are of length $\epsilon(t)$. The spectrum of Lyapunov exponents for trajectory $\Phi(t)$ is defined as

$$\lambda_i = \lim_{t\to\infty} \lim_{\varepsilon(0)\to\infty} \frac{1}{t} \log \frac{\varepsilon_i(t)}{\varepsilon(0)} \qquad (2)$$

The Lyapunov exponents depend on the trajectory $\Phi(t)$. Their values are the same for any state on the same trajectory, but may be different for states on different trajectories.

Lyapunov exponents are convenient for categorizing steady state behaviour. The trajectory of an n-dimensional state space have n Lyapunov exponents. This is often called the Lyapunov spectrum and it is conventional to order them according to size. The qualitative features of the asymptotic local stability properties can be summarized by the sign of each Lyapunov exponent; a positive Lyapunov exponent indicating an unstable direction, and a negative exponent indicating a stable direction. If the exponent is positive then the trajectories diverge and the system is chaotic. However, for an attractor, contraction must outweigh expansion and so,

$\sum_{i=1}^{m} \lambda_i < 0$ (3) The geometrical meaning of positive Lyapunov exponents is that there exist directions in which the motion on average is unstable such that nearby trajectories in these directions will diverge from the original orbit. Although the orbits is unstable, its stable directions provide sufficient volume contraction so that the orbit is confined to some bounded region in state space. At least one Lyapunov exponent must be zero for any limit set other than an equilibrium point. To produce a strange attractor the system must be dissipitative and hence must have at least one negative Lyapunov exponent. Furthermore, at least one Lyapunov exponent must be zero for any limit set other than an equilibrium point. Also for a chaotic system, at least one Lyapunov exponent must be positive. It follows that a strange attractor must have at least lyapunov exponents. Hence, chaos can only occur in third-order autonomous, second-order non-autonomous or higher order continuous time systems.

Dimensions

Dimension of an attractor corresponds to the number of equations (variables) necessary to describe a system. A system attractor could be defined to be n-dimensional if in a neighborhood of every point it looks like an open subset of R^n. This is how in differential topology the dimension of a manifold is defined. For instance, fixed point is of 0 dimension, while a limit cycle while locally looks like an interval is one dimensional and a torus is two dimensional. The neighbourhood of any point of a strange attractor however has a fine structure and does not manifolds and do not have integer dimension.

There are several ways to generalize dimension to the fractional case, some of which are listed below.

Capacity dimension: The simplest type of dimension measure, which is also referred as a **Fractal dimension**. To compute this measure, the attractor in phase space is covered by a regular grid of volume elements (cubes, spheres etc.) of diameter ε. If the attractor is a D-dimensional manifold, the number of volume elements needed to cover it for small ϵ is given by, $N(\varepsilon)=k\varepsilon^{-D}$ for some constant k. The definition of capacity dimension is obtained from this by taking the ϵ-limit.

$$D_0 = \lim_{\epsilon \to 0} \frac{\ln(-\sum_{i=1}^{N} pi \ln pi)}{\ln(1/\epsilon)} \quad \lim_{\epsilon \to 0} \frac{\ln N(\epsilon)}{\ln(1/\epsilon)}$$

For manifolds, Do is equal to the dimension of the manifold and is an integer while for objects that are not manifolds it gives a nonintegeral value.

Information dimension: While Do is a metric concept and does not utilize the information on the time behaviour of the system, information dimension is a probabilistic measure defined in terms of the relative frequency of visitation of a trajectory. This dimension is defined as,

$$D_1 = \lim_{\epsilon \to 0} \frac{\ln(-\sum_{i=1}^{N} pi \ln pi)}{\ln(1/\epsilon)}$$

Here, pi is the relative frequency at which a typical tracjectory enters the with volume element of the covering is the amount of information needed to specify the state of the system to accuracy ϵ if the state is known to be on the attractor. Hence the name information dimension.

Correlation dimension: Yet another probabilistic measure is the correlation dimension defined as

$$D2 = \lim_{\epsilon \to \infty} \frac{\ln \sum_{i=1}^{N(\epsilon)} pi^2}{\ln \epsilon}$$

An easy way to estimate this dimension is by determination of the correlation function for N points given by,

$$C(\epsilon) = \lim_{\epsilon \to \infty} \frac{1}{N^2} \{ \text{ the number of pairs xi such that } ||xi - xj|| < \epsilon$$

The Correlation Dimension or attractor dimension of a time series can be estimated from the Correlation Integral C (r) of the phase space vectors.

$$C(r) = \frac{1}{N(N-1)} \sum_{i=1}^{N} \sum_{j \neq 1}^{N} \theta(r - | x_i - x_j |)$$

where θ is Heaviside unit step function defined as 1 if the distance between vectors is less than r and 0 if distance is greater than r, N is the number of data points in the phase space, X_i, X_j are points of trajectory in the phase space, the distance r is a radius around each reference pointXi.

Grassberger and Procassia showed that C (r) obeys the scaling law

$C (r) = r^{D2}$

Therefore $D2 = \lim_{r \to 0} \frac{\log C(r)}{\log (r)}$

Where D is the dimension. C(r) is a measure of the probability that two arbitrary points xi, xj of phase space will be separated by distance r. Plotting log C (r) versus log (r) allows us to calculate D2 from the slope of the curves. Main advantage of dimensional analysis is the investigation of time series with noise like spectra. If the phase space representation of such signal converges to an attractor , its dimensionality is a finite number. Otherwise the time series has stochastic properties and cannot be considered as a signal from deterministic system..

D_2 is calculated for the standard Lorenz system using Grassberger and Procassia algorithm as shown in figure 3.5. We get the logarithm of the correlations sum, the correlation dimension.

$$\begin{aligned}\dot{X} &= -\sigma(X - Y) \\ \dot{Y} &= rX - Y - XZ \\ \dot{Z} &= b(XY - Z)\end{aligned}$$

were σ, r and b are parameters

were σ = 10, b = 8/3 and r = 99

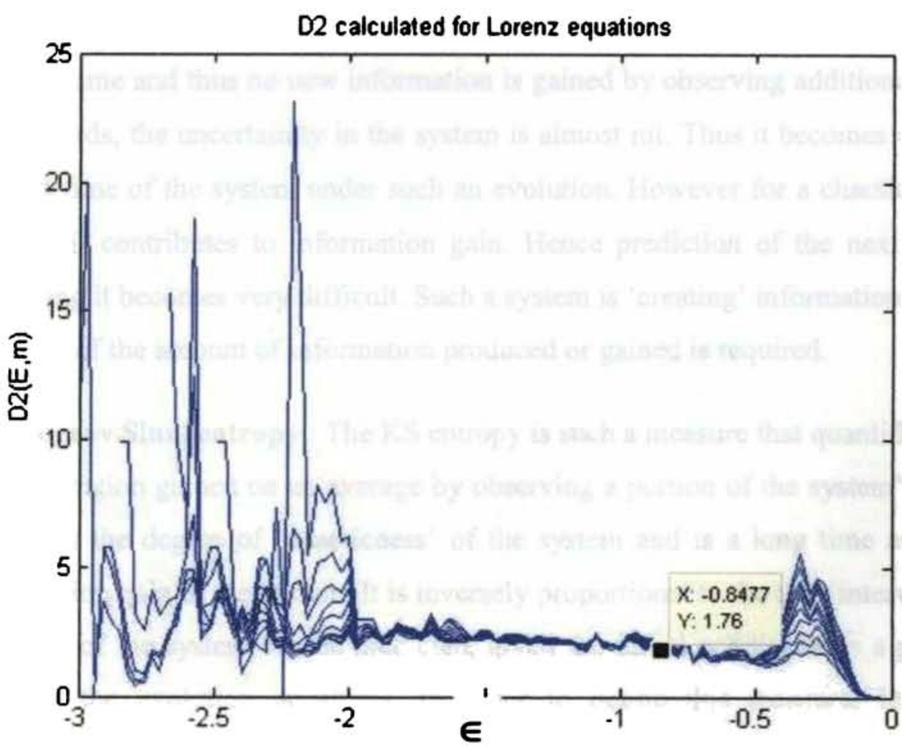


Figure 3.5. Calculation of correlation dimension for Lorenz equation.

With these different measures defined, the natural question is to find out how these are related to each other, Capacity dimension considers the attractor as a static object completely ignoring that it is subject to a dynamical flow. In practical experiments and simulations the attractor is not directly seen, only the trajectories over a finite period are observed. Hence the probabilistic measures are of greater use in practical settings than the capacity dimension. In general, $D_2 \leq D_1 \leq D_0 \leq$ since D_2 and D_1 being probabilistic in nature depend on the relative frequency at which each volume element is visited and hence will be equal to D_0 only when these frequencies are all equal.

Entropies

The dynamical characteristics of the system are studied using the measures of entropy. These measures, which help in understanding how the trajectories evolve in time, are specific of dynamical systems while dimensions can be defined for any fractal measure or point set. Entropy estimates the average information gained by observing a system's state to a precision ϵ. In the case of a fixed point or periodic orbit each orbit remains the same in time and thus no new information is gained by observing additional orbits. Or in other words, the uncertainity in the system is almost nil. Thus it becomes easy to predict the outcome of the system under such an evolution. However for a chaotic system, each new orbit contributes to information gain. Hence prediction of the next orbit without observing it becomes very difficult. Such a system is 'creating' information and a suitable measure of the amount of information produced or gained is required.

Kolmogrov Sinai entropy : The KS entropy is such a measure that quantifies the amount of information gained on an average by observing a portion of the system's evolution. It measures the degree of 'chaoticness' of the system and is a long time average rate of information gain of the system. It is inversely proportional to the time interval over which the state of the system can be predicted, given the initial conditions to a precision ϵ as well as the evolution equations. In order to define this measure, let us consider partitioning the attractor as before into $N(\epsilon)$ boxes $S_1, S_2, \ldots, S_N$ with size ϵ. If m measurements at regular time intervals are made, these will yield a sequence of boxes visited by the tractectory. Let $P(s_1,s_2,\ldots,s_N)$ be the joint probability of finding the trajectory at time Γ in box s1, at time 2Γ in box s2, and so on. The KS entropy is defined as,

$$K= - \lim_{\tau\to 0}\lim_{\varepsilon\to 0}\lim_{m\to\infty}\times[\frac{1}{m\tau}\sum_{s1\ldots sm} P(s1,\ldots sm)\times \ln P(s1,\ldots sm)] \qquad (1)$$

If K approaches 0, or there is no change in information this means the system is fully predictable. In contrast for a stochastic process, K approaches infinity and it attains in between values depending on the irregular nature for chaotic systems.

The entropies can be generalized to a set of order q Renyl entropies, which are dynamical counterparts of Renyi dimensions. These are defined as

$$Kq = - \lim_{\tau\to 0}\lim_{\varepsilon\to 0}\lim_{m\to\infty}\times[\frac{1}{m\tau}\frac{1}{q-1}\ln \sum_{s1,\ldots,sm} p^{q}(s1,\ldots,sm) \qquad (2)$$

The K entropy in eq. (1) for which q=2 in eq. (2) is however the easiest to compute among all the entropies.

Kolmogorov Entropy provides a measure of the rate of information flow in the system. It is thus the dynamical evolution of the system. K2 is computed from time series data using the equation

$$K_2 = \frac{1}{T(d'-d)}\ln\frac{Cd(r)}{Cd'(r)} \qquad (3)$$

Where d and d' are two embedding dimensions.

K_2 is calculated for the standard Lorenz system

$$\begin{aligned} \dot{X} &= -\sigma(X-Y) \\ \dot{Y} &= rX - Y - XZ \\ \dot{Z} &= b(XY - Z) \end{aligned}$$

using Grassberger and Procassia algorithm as shown in figure 3.6.

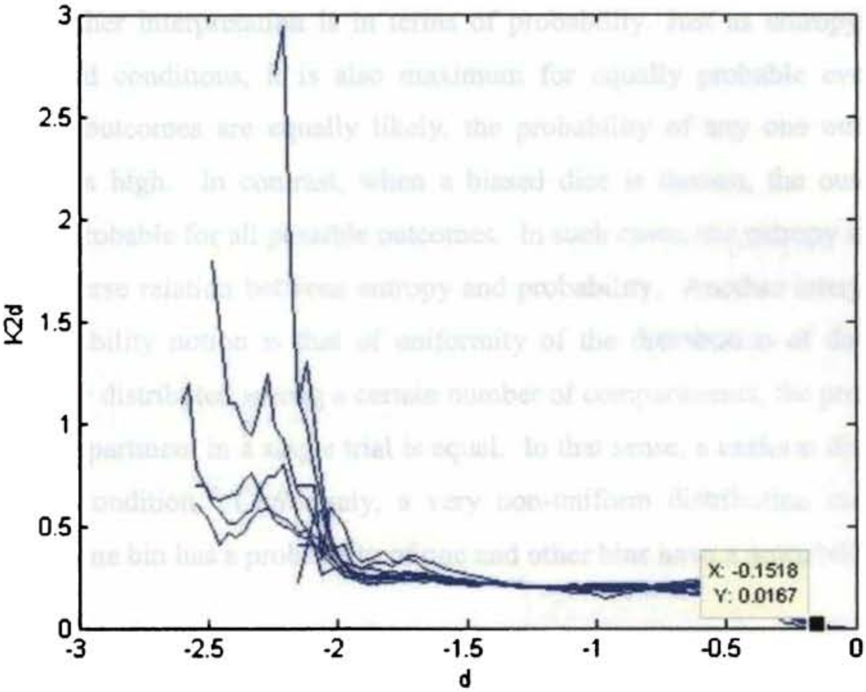


Figure 3.6. Calculation of Kolmogorov entropy (K_2)

Entropy : Various forms and Interpretations :

Information and Entropy: Interpretations

Information is one of the many interpretations of entropy. Chronologically it was relatively a very late interpretation. The concept of entropy was conceived and its application to dynamical interpretations was introduced in Clausius's theory of thermodynamics. In this theory, entropy is defined as the measure of unavailable energy which arises due to heat loss as well as other actions such as chemical reactions, mixing, change from solid to liquid to gas. These processes involve not only an increase in entropy but also an accompanying decrease in the orderly arrangement of constituent atoms implying an increase in disorder. For example, atoms and molecules are ordered in a solid thus having low entropy whereas in the case of liquids the entropy will be higher with less ordering of atoms or molecules. Therefore the measurement of entropy became regarded as the measurement of degree of disorder or disorganization of a system.

Still another interpretation is in terms of probability. Just as entropy is maximum for disordered conditions, it is also maximum for equally probable events i.e. when all possible outcomes are equally likely, the probability of any one outcome is low and entropy is high. In contrast, when a biased dice is thrown, the outcome will not be equally probable for all possible outcomes. In such cases, the entropy is lowest. So there is an inverse relation between entropy and probability. Another interpretation based on the probability notion is that of uniformity of the distribution of data. If the data is uniformly distributed among a certain number of compartments, the probability of getting each compartment in a single trial is equal. In that sense, a uniform distribution is a high entropy condition. Conversely, a very non-uniform distribution means low entropy, because one bin has a probability of one and other bins have a probability of zero.

Yet another notion of entropy is uncertainty. The uncertainty can be pertained to the outcome of an experiment about to be run, or it can pertain to the state of a dynamical system. When the outcome of an event is absolutely certain then uncertainty is zero indicating zero entropy. For eg: absolute certainty means probability P=1, for which case the entropy will be zero.

Another idea is related to randomly distributed observations versus reliable predictability. When there is disorder, and great uncertainty, predictions cannot be based on any known structure or pattern and can only be done probabilistically. In such cases where predictability is low, entropy will be high. In contrast, something well organized or nearly certain is usually very much predictable resulting in low entropy. The idea of many possible outcomes suggests diversity. Another idea of interpreting entropy is related to the information content of an event. In a given probability distribution there is an information value of so many bits. Furthermore, a relatively large number of bits means a relatively large number of information and vice versa. Hence entropy is the average amount of new information gained from a measurement. Table. 3.2 shows the different cases where low and high entropies occur with respect to the above discussed interpretations.

Table 3.2 Different cases where low and high entropies occur

High Entropy	Low Entropy
1. Large proportion of energy unavailable for doing work	Large proportion of energy available for doing work
2. Disorder, disorganization, thorough mix	Order, high degree of organization
3. Equally probable events, low probability of a selected event	Preordained outcomes, high probability of a selected event
4. Uniform distribution	Highly uneven distribution
5. Great uncertainty	Near certainty , high reliability
6. Randomness ,unpredictability	Non randomness, accurate forecasts
7. Much information	Little information

Entropy Measures for Complexity Analysis

Approximate entropy charaterises the regularity of a signal by measuring the presence of similar patterns in a time series. Consider a time series of length N,

From this time series short sequences or patterns $x_m(i)$ of length m are constructed and the quantity C_i^m with tolerance r defined as

$$C_i^m(r) = N^{-1}\{numberof \quad j \le N - m + 1 \quad |d\,[x_m\,(i), x_m(j)] \quad \le r\} \qquad (1)$$

is computed for each $x_m(i)$

This quantity measures the regularity of the patterns by comparing them to a given pattern. Here m is the detail level at which the signal is analysed and r is the threshold which filters out irregularities. The regularity parameter ApEn is defined as

$$ApEn\ (m, r) \quad = \quad \lim_{N \to \alpha} \left[\phi^{\,m}(r) - \phi^{\,m}(r) \right] \qquad (2)$$

where

$$\phi^{m}(r) = (N-m+1)^{-1} \sum_{i=1}^{N-m+1} \ln C_i^m(r) \qquad (3)$$

This gives the relative frequency of finding a vector $x_m(j)$ similar to vector $x_m(i)$ within a tolerance of r and provides a quantitative measure of entropy of the time series. ApEn statistic gives good results and provides general information about the regularity and persistence of a signal. However, in the above evaluation method, vectors or pattern $x_m(i)$ are allowed to self match and therefore results in biased statistics [136, 139, 138]. To overcome this drawback, a modification of ApEn algorithm named Sample entropy (SampEn) is developed [139] which avoid this self matching. SampEn shows a relative consistency compared to ApEn [139]. However, this measure strongly depends on length of the series and also gives biased results in the case of irregular signals.

As in the case of Komogorov-Sinai entropy, both sample and approximate entropy, provide a measure for the information increase over one step of dimension from m to m+1. To be able to resolve complexity on scales larger than this smallest scale, multiscale entropy is introduced [140, 141]. The efficiency of ApEn and SampEn are enhanced by estimating these measures at different time scales. These measures are widely used for charaterising biological signals in clinical application [142, 143].

Though the above methods give reliable results, their applicability to real world signal analysis is limited due to the sensitivity to noise and computation cost. Therefore, a fast and efficient algorithm which is also robust to noise contamination is very essential for online applications. PE is one such measure suitable for analysis of real world data.

Relative entopy:

Relative entropy measures applied to healthy and pathological voice characterization . According to Cover and Thomas [145] , entropy is a quantity defined for any probability distribution with properties that agree with the intuitive notion of information measures . One of the first concept was presented in [146] as the definition of a measure of uncertainity of random variable. Considering a random variable X that assumes values x ε χ where χ is a finite set, the entropy H(X) can be defined by H(X) =

$$- \sum_{x\varepsilon\chi} p(x)\log_2 p(x)$$

The probability of x, Pr{X = x}, is denoted by p(x) = 0, $p(x)\log_2 p(x) = 0$ by convention. This quantity is dependent on the distribution of X instead of the actual values of the random variable. The entropy measures the average amount of bits necessary to store outcomes of the random variable.

It is possible to estimate the relative entropy of two distributions[144], p(x) and q(x), in discete-time form, in the following equation

$$D(p||q) = \sum_{x\varepsilon\chi} p(x)\log_2 \frac{p(x)}{q(x)}$$

The relative entropy is also known as the kullback-Leibler(KL) entropy, which may be understood as a measure of the difficulty of discriminating between two distributions. Relative entropy is well suited due to its sensibility to uncertainties. Recently Paulo Roge´rio Scalassara et. al., [73] applied relative entropy to characterize healthy patients from pathological voice of patients suffereing from vocal disorders like nodule in vocal fold and Reinke's edema. The study showed that nonlinear dynamical methods seem to be suitable technique for voice signal analysis. Due to the chaotic component in the signals the signals are characterized by an increase in the signal complexity and unpredictability . The results showed that pathological groups had higher entropy values in accordance with other vocal acoustic parameters presented viz., jitter, shimmer, Harmonic to noise ratio . Measures of entropy are intimately related to the predictability of signals. These measures can be used to evaluate forecast skill of a system.

Permutation Entropy

Permutation Entropy [147] is a complexity measure which has aspects of both dynamical systems and entropy measures. PE calculation relies on the order relations between neighboring values of a time series. It estimates complexity as the entropy of the distribution of permutations of groups of time samples. PE can efficiently detect the regular and complex nature of any signal and extract useful information about the dynamics. Thus the variation of PE as a function of time can effectively indicate dynamical change in any real world data. As this method does not require direct calculations of embedding dimension and time delay, this gives faster output and makes it suitable for online application of real time processes. It is robust against dynamical as well as observational noise [147].

Calculation of Permutation Entropy

Computation of PE is based on comparison of neighbouring values in the time series of any dynamical variable of a system. It has been shown that any continuous time series representing a dynamical system can be mapped on to a symbolic sequence [147, 148, 149]. According to the embedding theorem, any arbitrary time series $X = \{x_1, x_2, \ldots\ldots x_T\}$ can be mapped on to an 'n' dimensional space with vectors $X_i = \{x_i, x_{i+\tau}, x_{i+2\tau}, \ldots\ldots x_{i+(n-1)\tau}\}$ where n is the embedding dimension and τ is the delay time for embedding calculated using appropriate methods like false nearest neighbour calculation and first minimum of autocorrelation function [150]. For any arbitrary vector X, the components are n number of real values of the time series $\{x_t, x_{t+\tau}, x_{t+2\tau} \ldots\ldots\ldots x_{t+(n-1)\tau}\}$ from time instant 't' to '$t+(n-1)\tau$'. Assuming $\tau = 1$, each point in the n dimensional space represented by its corresponding vector will therefore be equivalent to a short sequence of the time series consisting of n number of real values as $\{x_t, x_{t+1}, x_{t+2,\ldots} \ldots\ldots x_{t+(n-1)}\}$. If the components of each vector are arranged in ascending order, it will represent a pattern of evolution. Thus each of the vectors can be considered as a symbolic sequence which will be one of the $n!$ possible permutations of 'n' distinct symbols. The probability distribution of each pattern π can be represented as

$$p(\pi) = \frac{\#\{t \mid t \le T-n, (x_{t+1,\ldots\ldots,} x_{t+n})\, has\, type\, \pi\}}{T-n+1} \tag{6}$$

where π represents a pattern and # represents the number of occurrences. Permutation entropy of order $n \geq 2$ is defined as the Shannon entropy of the $n!$ patterns or symbolic sequences and can be written as

$$H(n) = -\sum p(\pi) \log p(\pi) \tag{7}$$

where the sum runs over all n! permutations or sequences. $H(n)$ lies between 0 and $\log(n!)$. For increasing or decreasing sequence of values, $H(n) = 0$, whereas for random series where all $n!$ possible permutations appear with same probability, $H(n) = \log(n!)$. For a time series representing some dynamics, $H(n) < \log(n!)$. Therefore, normalised PE per symbol of order 'n' is given by $H(n)/\log(n!)$. Thus PE characterizes the system dynamics, with low values indicating regular behaviour. Any increase in PE value will thus represent an increase in irregularity in the dynamics. For detection of dynamical changes from time series it is first partitioned into non-overlapping windows of suitable length T. PE for each window is calculated using Eq.(6) and Eq.(7). Any change in the dynamics of the system will be reflected in the variation of PE with respect to moving window. For a reliable estimation of PE, the window length T should be greater than $n!$ [147]. The order of PE should not be too small as this will not give enough number of distinct states. Too large values of order 'n' will demand large values of window size which will not effectively detect dynamical changes and also will create memory restrictions. Optimum values of order of PE are reported to be around 3 to 8[147,148]. In our analysis PE of order 5 is used for a window size of 512 samples for the time series of the audio signal.

As the patterns can be calculated in a very fast and easy way, calculation time of PE is negligibly less compared to other classical nonlinear methods. In this, only two pairs of values are compared at a time. PE based method is 100 times faster than Lyapunov exponent based method [147] due to the fact that neighbourhood searching is not needed. Also we deal with order relations between values instead of values themselves, the permutation entropy is robust with respect to noise corrupting the data.

Permutation entropy has a practically invariance property. If $y_t = f(x_t)$ where f is an arbitrary strictly increasing (or decreasing) real function, then $H(n)$ is same for x_t and y_t. Such nonlinear function f occurs, for example, when measuring physiological data with different equipments. Addition of observational noise causes only a small increase in the value of entropy [151] where as there is hardly any effect on the entropy due to dynamical noise. However in the presence of high noise the ability of PE to distinguish the change in dynamics decreases.

Standard Data Test on PE

Effectiveness of PE is verified on regular chaotic and random data sets. For this, normalised PE for a regular sine wave of amplitude (peak to peak) 0. 2 and a random signal of amplitude 0.2 with 5000 data points each are calculated. Fig 3.7 (a) and (b) shows a sine wave and random signal respectively and their corresponding PE are shown in Fig 3.7 (c) and (d). Permutation entropy for regular sine wave is 0.114 and it is 0.9387 for random signal. Hence it is confirmed that PE values corresponding to regular signal is low whereas for random variation it shows high values.

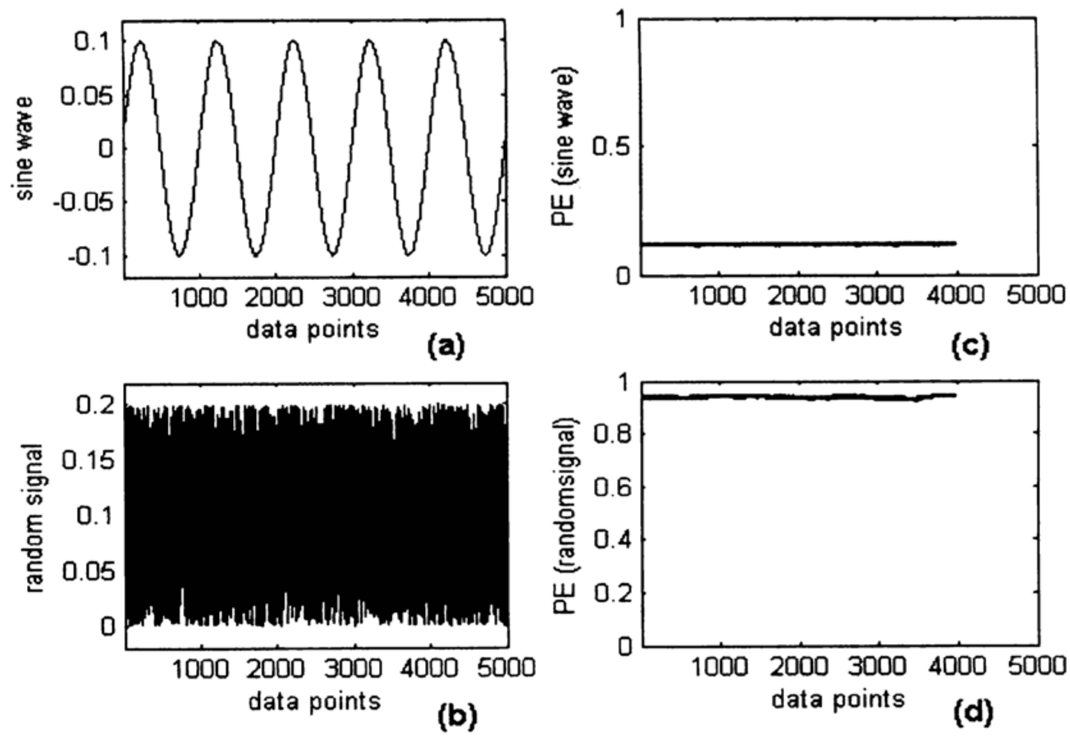


Fig. 3.7 Variation of PE for regular and random signal. (a) sine wave (b) random signal (c) PE for sine wave (d) PE for random signal

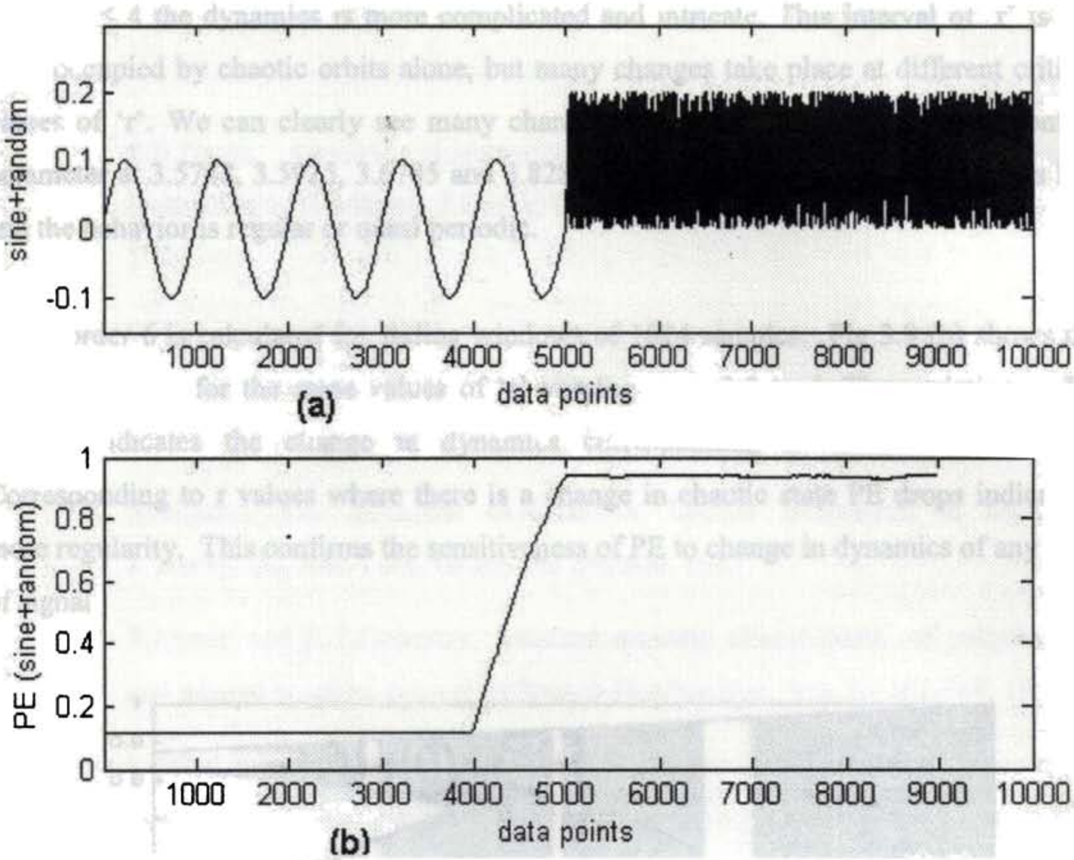


Fig. 3.8 Variation of PE for regular signal connected to random signal
(a) sine wave connected to random signal (b) PE for sine wave connected to random signal

When regular sine wave is connected with random signal as given in Fig 3.8 (a) the PE value suddenly jumps from 0.114 to 0.9387 as indicated in Fig 3.8(b). This clearly shows that the PE is sensitive to change in regularity. The sudden variation from regular to random state is clearly indicated by the abrupt change in PE values.

To get the feeling of the variations of entropy, results are also verified on chaotic signals with change in dynamics for different parameter values. Bifurcation diagram of logistic map $x_{t+1} = rx_t(1 - x_t)$ is used to study the variation in PE in chaotic signal. Fig. 3.9(a) shows the bifurcation diagram of logistic map for 5000 parameter values corresponding to variation of 'r' from 3.5 to 4. For control parameter r less than 3.57 the logistic map exhibits period doubling phenomenon, and a chaotic dynamics is observed at 3.57. For

$3.57 \leq r \leq 4$ the dynamics is more complicated and intricate. This interval of 'r' is not fully occupied by chaotic orbits alone, but many changes take place at different critical values of 'r'. We can clearly see many changes of dynamics as a function of control parameter at 3.5748, 3.5925, 3.6785 and 3.828. At these points the chaotic nature is lost and the behavior is regular or quasi periodic.

PE of order 6 is calculated for sliding windows of 1024 samples. Fig 3.9 (b) shows the variation of PE for the same values of 'r' varying from 3.5 to 4. The variation in PE clearly indicates the change in dynamics corresponding to different 'r values'. Corresponding to r values where there is a change in chaotic state PE drops indicating more regularity. This confirms the sensitiveness of PE to change in dynamics of any type of signal.

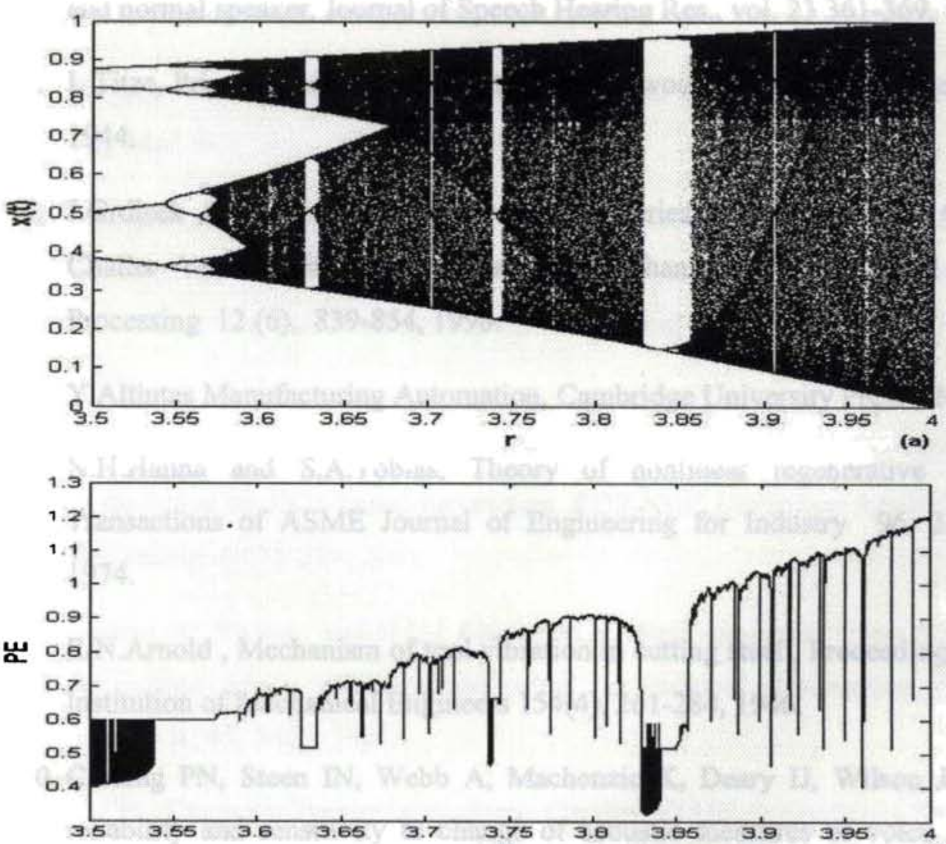


Fig. 3.9 Logistic equation for varying control parameter 'r'

(a) Bifurcation diagram (b) Variation of PE

3.5 REFERENCES

1. S.B.Davis, Acoustic characteristic of laryngeal pathology, In Speech Evaluation in Medicine, J.Darvy, Ed. NewYork, Grune and Stratton, 77-104, 1981.

2. S.Hiki, S.Imaizumi, M.Hirano,H.Matsuhita and Y.Kita, Acousticical analysis for voice disorders, IEEE Proceedings International Conference on Acoustic Speech Signal Processing, 613-616, 1976.

3. M.Hirano, The laryngeal examination, Speech Evaluation in Medicine J. Darvy, Ed. NewYork, Grune and Stratton, 1981.

4. T.Murry and E.T.Dohertry, Selected acoustic characteristic of pathological and normal speaker, Journal of Speech Hearing Res., vol. 23 361-369, 1980.

5. I. Titze, Principle of Voice production, Engle wood cliffs, NJ, Prentice – Hall, 1944.

6. J.Grdisek ,E.Govekar and I.Grabec , Time Series Analysis in Metal Cutting: Chatter versus Chatter-Free Cutting, Mechanical Systems And Signal Processing 12 (6), 839-854, 1998.

7. Y.Altintas Manufacturing Automation, Cambridge University Press , 2000.

8. N.H.Hanna and S.A.Tobias, Theory of nonlinear regenerative chatter. Transactions of ASME Journal of Engineering for Industry 96, 247-255, 1974.

9. R.N.Arnold , Mechanism of tool vibration in cutting steel , Proceedings of the Institution of Mechanical Engineers 154(4), 261-284, 1946.

10. Carding PN, Steen IN, Webb A, Machenzie K, Deary IJ, Wilson JA: The reliability and sensitivity to change of acoustic measures of voice quality, Clinical Otolaryngology, 29(5):538-544, 2004.

11. M. Frohlich, D. Michaelis, H.W. Srube: Acoustic 'Breathiness Measures' in the description of Pathologic Voices, Porc. ICASSP-98, Seattle, WA, vol. 2, 937-940, 1998.

12. Dejonckere PH, Bradley P, Clemente P, Cornut G, Crevier-Buchman L, Friedrich G, Van De Heyning P, Remacle M, Woisard V: A basic protocol for functional assessment of voice pathology, especially for investigating the efficacy of (phonosurgical) treatments and evaluating new assessment techniques. Guideline elaborated by the Committee on Phoniatrics of the European Laryngological Society (ELS). Eur Arch Otorhinolaryngol, 258(2):77-82, 2001.

13. Michaelis D, Frohlich M, Strube HW: Selection and combination of acoustic features for the description of pathologic voices. Journal of the Acoustical Society of America, 103(3):1628-1639, 1998.

14. Boyanov B, Hadjitodorov S: Acoustic analysis of pathological voices. IEEE Eng Med Biol Mag, 16(4):74-82, 1997.

15. Godino-Llorente JI, Gomez-Vilda P: Automatic detection of voice impairments by means of short-term cepstral parameters and neural network based detectors. Ieee Transactions on Biomedical Engineering, 51(2):380-384, 2004.

16. Alonso J, de Leon J, Alonso I, Ferrer M: Automatic detection of pathologies in the voice by HOS based parameters. EURASIP Journal on Applied Signal Processing, 4:275-284, 2001.

17. Kennel, R. Brown , and H.D.I Aberbanel, Determining embedding dimension for phase- space reconstruction using geometrical construction, Physical Review A, 45, 3403, 1992.

18. E. Ott, Chaos in Dynamical Systems, Cambridge University Press, New York 1993.

19. J.D Farmer, E. Ott and J. A Yorke , The dimension of chaotic attractors, Physica D, 7, 153-180, 1983.

20. J.Theiler, Estimating fractal dimension, Journal of Optical Society of America, 7, 1055-1073, 1990.

21. P. Grassberger, I. Procassia, Measuring the strangeness of strange attractors, Physica D, 9, 189-208, 1983b.

22. A. Wolf, J.B. Swift, H.L. Swinney, and J.A. Vastano Determining Lyapunov exponents from a time series, Physica D, 16, 285-317, 1985.

23. P. Grassberger, I. Procassia, Estimation of Kolmogorov entropy from chaotic signal, Physical Review A, 28, 2591-2593, 1983c.

24. Behlau M, Madazio G, Feijó D, Pontes P. Avaliação da Voz. In: Behlau M (org.) Voz – O Livro do Especialista. Vol. I. Rio de Janeiro: Revinter; Cap. 3, 86-180., 2001.

25. J. L. Flanagan, Speech Analysis, Synthesis and perception, 2nd Ed., Springer-verlag, new York, 1972.

26. W. Koenig, H.K. Dunn, and L.Y. Lacy, The Sound Spectrograph, J. Acoust. Soc. Am., Vol. 17, pp. 19-49, July 1946.

27. K. Johnson. Acoustic and auditory phonetics. Blackwell Pub., Malden, Mass., 2nd edition, 2003.

28. P. Ladefoged. A course in phonetics. Harcourt College Publishers, Fort Worth, 4th edition, 2001.

29. Moore P, Leden H von. Dynamic variations of the vibratory pattern in the normal larynx. Folia Phoniatr;10:205-38, 1958.

30. Lieberman P. Perturbations in vocal pitch. JAcoust SocAm ; 33:597-603, 1961.

31. Michel J. Vocal fry and harshness [PhD dissertation]. Gainesville, FL: University of Florida, 1964.

32. Hiller SM. Automatic acoustic estimation of waveform perturbations. [PhD dissertation]. Edinburgh, Scotland: University of Edinburgh, 1985.

33. Anders LC, Hollien H, Hurme P, Sonninen A, Wendler J. Perception of hoarseness by several classes of listeners, Folia Phoniatr; 40:91-100, 1988.

34. Askenfelt A, Hammarberg B., Speech waveform perturbation analysis. Speech Transmission Laboratory—quarterly progress and status report. Stockholm: Royal Institute of Technology, 440-9, 1980.

35. Askenfelt A, Hammarberg B., Speech waveform perturbation analysis revisited. Speech Transmission Laboratory quarterly progress and status report. Stockholm: Royal Institute of Technology, 449-68, 1981.

36. Askenfelt A, Sjolin A., Voice analysis in depressed patients: rate of change of fundamental frequency related to mental state. Speech Transmission Laboratory--quarterly progress and status report. Stockholm: Royal Institute of Technology, 71-84, 1980.

37. Coleman RF, Effect of median frequency levels upon the roughness of jittered stimuli. J Speech Hear Res, 12: 330-6, 1969.

38. Coleman RF, Wendahl RW, Vocal roughness and stimulus duration. Speech Monogr; 34:85-92, 1967.

39. Davis SB., Computer evaluation of laryngeal pathology based on inverse filtering of speech. Speech Communication Research Laboratory, SCRL Monograph 13, Santa Barbara, CA, 1976.

40. Deal R, Emanuel F., Some waveform and spectral features of vowel roughness. J Speech Hear Res ; 21:250-64, 1978.

41. Emanuel FW, Sansone FE Jr. Some spectral features of normal and simulated rough vowels. Folia Phoniatr; 21:401-15, 1969.

42. Hammarberg B, Fritzell G, Gauffin J, Sundberg J, Wedin L. Perceptual and acoustic correlates of abnormal voice qualities. Acta Otolaryngol; 90:441-51, 1980.

43. Hollien H, Michel J, Doherty ET. A method for analyzing vowel jitter in sustained phonation. J Phonet; 1:85-91, 1973.

44. Horii Y., Some statistical characteristics of voice fundamental frequency. J Speech Hear Res; 18:192-201, 1975.

45. Horii Y., Fundamental frequency perturbation observed in sustained phonation. J Speech Hear Res; 22:5-19, 1979.

46. Horii Y. Vocal shimmer in sustained phonation. J Speech Hear Res; 23:202-9, 1980.

47. Horii Y., Jitter and shimmer differences among sustained vowel phonations. J Speech Hear Res; 25:12-4, 1982.

48. Isshiki N, Okamura H, Tanabe M, Morimoto M. Differential diagnosis of hoarseness. Folia Phoniatr; 21:9-19, 1969.

49. Kane M, Wellen CJ., Acoustical measurements and clinical judgments of vocal quality in children with vocal nodules. Folia Phoniatr; 37:53-7, 1985.

50. Kasprzyk PL, Gilbert HR. Vowel perturbation as a function of vowel height. J Acoust Soc Am; 57:1545-6, 1975.

51. Kasuya H, Kobayashi Y, Kobayashi T. Characteristics of pitch period and amplitude perturbations in pathologic voice. 1CASSP-83: 1372-5, 1983.

52. Kempster GB, Kistler DJ., Selected acoustic characteristics of pathological and normal speakers: a re-analysis. J Speech Hear Res;26:159-60, 1983.

53. Kitajima K, Gould WJ., Vocal shimmer in sustained phonations of normal and pathological voice. Ann Otolaryngol, 85:377-81, 1976.

54. Koike Y., Application of some acoustic measures for the evaluation of laryngeal dysfunction. Studia Phonetica; 7:17-23, 1973.

55. Koike Y, Takahashi H, Calcaterra TC., Acoustic measures for detecting laryngeal pathology. Acta Laryngol; 84: 105-17, 1977.

56. Leden H von, Koike Y., Detection of laryngeal disease by computer techniques. Arch Otolaryngol; 91:33-40, 1970.

57. Ludlow CL, Naunton RF, Bassich CJ. Procedures for the selection of spastic dysphonia patients for recurrent laryngeal nerve section. Otolaryngol Head Neck Surg; 92: 24-31, 1984.

58. Murry T, Doherty ET., Selected acoustic characteristics of pathological and normal speakers: a re-analysis: a reply to Kempster and Kistler. J Speech Hear Res;26:160, 1983.

59. Sherman D, Linke E., The influence of certain vowel types on degree of harsh vowel quality. J Hear Speech Disord;17:401-8, 1952.

60. Sorensen D, Horii Y., Frequency and amplitude perturbation in the voices of female speakers. J Commun Disord; 16:57-61, 1983.

61. Sorensen D, Horii Y., Directional perturbation factors for jitter and shimmer. J Commun Disord;17:143-51, 1984.

62. Sorensen D, Horii Y, Leonard R. Effects of laryngeal topical anesthesia on voice fundamental frequency perturbation. J Speech Hear Res;23: 274-83, 1980.

63. Murry T, Doherty ET., Selected acoustic characteristics of pathological and normal speakers. J Speech Hear Res; 23:361-9, 1980.

64. Ludlow CL, Coulter DC, Gentges FH., The differential sensitivity of frequency perturbation to laryngeal neoplasms and neuropathologist. In: Bless DM, Abbs JH, eds. Vocal fold physiology: contemporary research and clinical issues. San Diego, CA: College Hill,:381-92, 1983.

65. Hecker M, Kreul E., Descriptions of the speech of patients with cancer of the vocal folds. Part h Measures of fundamental frequency. J Acoust Soc Am; 49:1275-82, 1971.

66. Baken RJ., Clinical measurements of speech and voice. Boston, MA: Little, Brown, 1987.

67. Iwata S, Leden H von. Pitch perturbations in normal and pathological voices. Folia Phoniatr; 22:413-24, 1970.

68. Iwata S., Periodicities of pitch perturbations in normal and pathological voices. Laryngoscope; 82:87-95, 1972.

69. John Laver, Steven Hiller, and Janet Mackenzie Beck, Special Article Acoustic Waveform Perturbations and Voice Disorders, Journal of Voice vol. 6, No. 2, pp. 115-126, Raven Press,- Ltd., New York, 1992.

70. Hiller SM., Automatic acoustic estimation of waveform perturbations. [PhD dissertation]. Edinburgh, Scotland: University of Edinburgh, 1985.

71. Haydée F. Wertzner, Solange Schreiber, Luciana Amaro, Analysis of fundamental frequency, jitter, shimmer and vocal intensity in children with phonological disorders, Rev Bras Otorrinolaringol.V.71, n.5, 582-88, sep./oct. 2005.

72. Christine D.L. van Gogh, Joost M. Festen, Irma M. Verdonck-de Leeuw, Andrew J. Parker , Louis Traissac, Anthony D. Cheesman, Hans F. Mahieu, Acoustical analysis of tracheoesophageal voice, Speech Communication 47, 160–168,ELSEVIER, 2005.

73. Paulo Roge'rio Scalassara, Marı'a Eugenia Dajer, Carlos Dias Maciel, Rodrigo Capobianco Guido, Jose' Carlos Pereira, Relative entropy measures applied to healthy and pathological voice characterization, Applied Mathematics and Computation 207, 95–108, ELSEVIER, 2009.

74 Yu Zhang and Clancy McGilligan, Liang Zhou, Mark Vig and Jack J. Jiang, 2004 Acoustical Society of America, Nonlinear dynamic analysis of voices before and after surgical excision of vocal polyps, J. Acoust. Soc. Am., Vol. 115, No. 5, Pt. 1, Pages: 2270–2277, May 2004.

75. V. Uloza, Effects on voice by endolaryngeal microsurgery, Eur. Arch. Otorhinolaryngol. 256, 312–315, 1999.

76. I. R. Titze, Workshop on acoustic voice analysis; Summary statement, National Center for Voice and Speech, Denver, CO, p. 36, 1995.

77. M. P. Karnell, A. Chang, A. Smith, and H. Hoffman, Impact of signal type of validity of voice perturbation measures, NCVS Status and Progress Report 11, 91–94, 1997.

78. I. R. Titze and H. Liang, Comparison of F0 Extraction Method for High-Precision Voice Perturbation Measurements, J. Speech Hear. Res. 36, 1120–1133, 1993.

79. S. M. Zeitels, R. E. Hillman, R. Desloge, M. Mauri, and P. B. Doyle, Phonomicrosurgery in singers and performing artists: treatment outcomes, management theories, and future directions, Ann. Otol. Rhinol. Laryngol. 111, 21–40, 2002.

80. J. J. Jiang, Y. Zhang, and C. N. Ford, Nonlinear dynamic of phonations in excised larynx experiments, J. Acoust. Soc. Am. 114, 2198–2205, 2003.

81. PARLEY W. NEWMAN, RICHARD W. HARRIS, and LAURENCE M. HILTON , VOCAL JITTER AND SHIMMER IN STUTTERING, J. Flency disorder 14, 87-95, 1989.

82. Deal, R.E., and Emanuel F.W., Some waveform and spectral features of vowel roughness. Journal of Speech and Hearing Research, 21, 250-264, 1978.

83. Horii, Y., Jitter and shimmer differences among sustained vowel phonations. Journal of Speech and Hearing Research, 25, 12-14, 1982.

84. Horii, Y., Vocal shimmer in sustained phonation. Journal of Speech and Hearing Research, 23, 202-209, 1980.

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

86. Smith, B.E., Weinberg, B., Feth, L.L., and Horii, Y. Vocal roughness and jitter characteristics of vowels produced by esophageal speakers. Journal of Speech and Hearing Research, 21, 240-249, 1978.

87. Stromsta, C., Elements of Stuttering. Oshtemo, MI: Atsmorts Publishing, 1986.

88. Horii, Y., Some statistical characteristics of voice fundamental frequency. Journal of Speech and Hearing Research, 18, 192-201, 1975.

89 Kedem, B., Spectral Analysis and Discrimination by Zero-crossings. Proceedings of IEEE, 74, No. 11, 1477-1492, Nov. 1986 .

90. L.R.Rabiner and R.W.Schafer, Digital Processing of Speech Signals. Englewood Cliffs, NJ, Prentice-Hall, 1978.

91. Arai-T and Yoshida-Y, Study on zero-crossings of speech signals by means of analytic signal, Journal-of-the-Acoustical-Society-of-Japan. Vol. 46, no.3, pp.242-246, March 1990.

92. Varada-S; Sankar-R, Hardware strategies for end-point detection IEEE, New York, NY, USA; vi=417 pp.p.163-7, 1995.

93. Santhoshkumar-U; Anto-BP; Sridhar-CS Digit recognition using energy envelope and zerocrossing rate IEEE, New York, NY, USA; 1098 pp. p.270-3, 1989.

94. Doh-Suk Kim, Jae-Hoon Jeong, Jae-Weon Kim and Soo-Young Lee, Feature Extraction Based on Zerocrossings With Peak Amplitudes for Robust Speech Recognition in Noisy Environments., IEEE Transaction on Audio Electro magnetic accost Vol AU 17 PP 61-64, Mar 1996.

95. Russell J. Niederjohn and Meir Lahat, A Zerocrossing Consistency Method for Formant Tracking of Voiced Speech in High Noise Levels, IEEE Transactions on Acoustics, Speech and Signal Processing, Vol. ASSP-33, No. 2, April 1985.

96. WIESLAWA KUNISZYK – JOZKOWIAK, The statistical analysis of Speech Envelopes in Stutterers and Non-stutterers, J. Fluency Disord. 20, 11-23, 1995.

97. Rees JM, Staszewski WJ, Winkler JR. Case study of a wave event in the stable atmospheric boundary layer overlying an Antarctic ice shelf using the orthogonal wavelet transform. Dyn Atmos Oceans, 34;245-61, 2001.

98. Staszewski WJ, Worden K. Wavelet analysis of time-series :coherent structures, chaos and noise, Int J Bifurcation Chaos; 9(3):455-71, 1999.

99. Torrence C, Compo GP, A practical guide to wavelet analysis . Bull AmMeteor Soc;79(1):61-78, 1998.

100. Lung SY., Wavelet feature selection using fuzzy approach to text independent speaker recognition. IEICE Trans Fundam Electron Commun Comput Science; E88A(3):779-81, 2005.

101. Sreenivasan P. Jaieson LH., High-quality audio compression using an adaptive wavelet packet decomposition and psychoacoustic modeling. IEEE Trans Signal Process; 46(4):1085-93, 1998.

102. Enders J. Geng W, Li P, Frazier MW, Scholl DJ. The shift-invariant discrete wavelet transform and application to speech waveform analysis, J. Acoustic Soc. Am; 117(4):2122-33, 2005.

103. Seok JW, Bae KS., Comparison of glottal closure instants obtained by using wavelet transform of speech signal and EGG signal. IEICE Trans Inform Syst; E82D(11):1486-8, 1999.

104. Wang JF, Chen SH, Shyuu JS., Wavelet transforms for speech signal processing. J Chinese Inst Eng;22(5):549-60, 1999.

105. Adaptation of wavelet transform analysis to the investigation of biological variations in speech signals Julia M. Rees, Gavita Regunath, Sandra P. Whiteside, Meghana B. Wadnerkar, Patrica E. Cowell, ScienceDirect, Medical Engineering & Physics 30, 865-871, 2008.

106. H. M. Teager and S. M. Teager, Evidence of Nonlinear Sound Production Mechanisms in the Vocal Tract in speech Production and Speech Modelling, W . J. Hardcastle and A. Marchal, Eds., NATO Advanced study Institute Series D, vol.55, Bonas, France, July 1989.

107. J.F.Kiser, Some observations on vocal tract operation from a fluid flow point of view, in Vocal Fold Physiology: Biomechanics, Acoustics and Phonatory Control, I.R.Titze and R.C.Scherer,Eds. Denver ,CO, Denver Center for Performing Arts, 358-386, 1983.

108. T.J.Thomas, A finite element model of fluid flow in the vocal tract, Comput. Speech Lang. vol. 1, 131-151, 1986.

109. R.S.McGowan, An aeroacoustics approach to phonation, Journal of Acoustic Society of America 2, vol.83, 696-704, 1988.

110. G.Richard,D.Sinder, H.Duncan,Q.Lin, J.Flanagan, S.Levinson, M.Krane, S.Slimon and D.Davis, Numerical stimulation of fluid flow in the vocal tract, In Proceeding Eurospeech, 1995.

111. A.Barney, C. Shable and P.Davies, Fluid flow in a dynamical mechanical model of the vocal fold and tract: part 1& 2, Journal of Acoustic Society of America 1, vol 105 pp. 444-466, 1999.

112. P. Maragos, A.Dimakiu and I. Kokkainos , Some advances in nonlinear speech modeling using modulations, fractals and chaos, Proceedings of International Conference on DSP, Santorini, Greece, 2002.

113. I.Kokkinos, P.Maragos, Nonlinear Speech Analysis Using Models for Chaotic Systems, IEEE transactions on speech and audio processing 6, vol. 13 1098-1109, 2005.

114. H.L.John, Liana Gavidia-Ceballos and James F.Kaiser , A Nonlinear Operator-Based Speech Feature Analysis Method with Application to Vocal Fold Pathology Assessment, AIEEE Transactions On Biomedical Engineering 3, vol. 45 300-313, 1998.

115. P. Maragos, Modulation and fractal Models for Speech Analysis and Recognitio , Proceedings of COST-249, Meeting, Porto ,Portugal, pp. 1-7, 1998.

116. H. G. Schuster, Deterministic chaos, VCH, Weinheim, 1989.

117. E. Ott, Chaos in dynamical systems, Cambridge University Press, Cambridge 1993.

118. S. H. Strogatz, Nonlinear dynamics and chaos, Addison-Wesley, Massachusetts, 1994.

119. H. Kantz and T. Schreiber, Nonlinear time series analysis, Cambridge University Press, Cambridge (1997).

120. H. Abarbanel, Analysis of observed chaotic data, 1996, Springer.

121. H. Kantz and T. Schreiber, Nonlinear time series analysis, Cambridge University Press, Cambridge, 1997.

122. H. D. I. Abarbanel, Analysis of observed chaotic data, Springer, New York, 1996.

123. J. C. Sprott, Chaos and time-series analysis, Oxford University Press, Oxford, 2003.

124. M. Small, Applied nonlinear time series analysis, World Scientific, Singapore, 2005.

125. A. M. Fraser and H. L. Swinney, Independent coordinates for strange attractors frommutual information, Physical Review A 33 2, 1134–1140, 1986.

126. M. B. Kennel, R. Brown, and H. D. I. Abarbanel, Determining embedding dimension for phase-space reconstruction using a geometrical construction, Physics Review A 45, 3403, 1992.

127. N.H. Packard, J.P. Crutchfield, J.D. Farmer et al., Geometry from a time series, Physical Review Letters 45, 712-716, 1980.

128. F. Takens, Detecting strange attractors in turbulence, Dynamical Systems and Turbulence, Warwick, Lecture Notes in Mathematics 898, Springer-Verlag, 366-381, 1981.

129. T. Sauer, J.A. Yorke, and M. Casdagli, Embedology, Journal of Statistical Physics 64, 579-616, 1991.

130. A. M. Fraser and H. L. Swinney, Independent coordinates for strange attractors from mutual information, Physical Review A 33, 1134-1140, 1986.

131. R. Shaw, Strange attractors, Chaotic behavior and information flow. Z. Naturforsch., 36a, 80-112, 1981.

132. K. Judd and L. Smith, Indistinguishable states-perfect model scenario , Physica D, 151 224-242, 2001.

133. J.D. Farmer, E.Ott and J. A Yorke, The dimension of chaotic attractors, Physica D, 7 153-180, 1983.

134. J. Theiler, Estimating fractal dimension, Journal of Optical Society of America, 7 1055-1073, 1990.

135. M. Perc, Introducing nonlinear time series analysis in undergraduate courses, Fizika A (Zagreb) 15 (2) 91–112, 2006.

136. S.M.Pincus, Approximate entropy (ApEn) as complexity measure , Chaos,1 vol 5 110-117, 1995.

137. S.M.Pincus , I.M.Gladstone, and R.A.Ehrenkranz, A regularity statistic for medical data analysis,, J. Clin. Monit ,1 vol.7 pp. 335-345, 1991.

138. S.M.Pincus, Approximate entropy as a measure of system complexity, in Proceedings of .Nat.Acad.Sci.USA, vol . 88 pp.2297-2301, 1991.

139. J.S Richman and J.R Moorman, Physiological time series analysis using approximate entropy and sample entropy, American journal of Physiology, Heart Circ.Physiol.,vol .278 pp.H2039-2049, 2000.

140. M. Costa, A.L. Goldberger, and C. –K. Peng, Multiscale entropy analysis of complex physiological time series, Physical Review Letters, 6 vol. 88 pp 068-102, 2002.

141. M. Costa, A.L. Goldberger, and C. –K. Peng, Multiscale entropy to distinguish physiologic and synthetic RR time series, Comput. Cardiol vol. 29 pp. 137-140, 2002.

142. D.E Lake, J.S. Richman and J.R Moorman, Sample entropy analysis of neonatal heart rate variability, Am. J. Physiol. Regul. Integr.comp. Physiol, vol.283 pp.R789-R797, 2002.

143. M.Costa, A.L.Goldberger, and C.K Peng, Multiscale entropy analysis of biological signals, Physical .Review E , vol.71 pp.021906, 2005

144. Timothy DelSole, Predictability and information theory. Part I: Measures of predictability, J. Atmos. Sci. 61 (20) 2425–2440, 2004

145. T.M. Cover, J.A. Thomas, Elements of Information Theory, John Wiley and Sons Inc., New York, 1991.

146. C.E. Shannon, A mathematical theory of communication, Bell Syst. Tech. J. 27 379–423, 623–656, 1948.

147. Bandt C, Pomp B, Permutation Entropy: A natural complexity measure for time series, Physical Review Letters, 17 88 174102 1-4, 2002.

148. I. Veisi, N. Pariz, A. Karimpour , Fast and Robust Detection of Epilepsy in Noisy EEg Signals Using Permutation Entropy, IEEE conference, 1 4244-1509, 2007.

149. Y. Cao, W. Tung, J.B. Gao , V.A. Protopopescu, M. Hively, Detecting dynamical changes in time series using the permutation entropy, Physical Review E, 2004.

150. H.Kanz,T.Schreiber, Nonlinear Time Series Analysis, Cambridge University Press , 1997.

151. Y. Cao, W. Tung, J.B. Gao , V.A. Protopopescu, M. Hively, Detecting dynamical changes in time series using the permutation entropy, Physical Review E, 70, 2004.

CHAPTER 4 - RESULTS AND DISCUSSION ON STUTTERING

In this chapter analysis of speech signals of healthy and stuttering subjects are carried out. Vocal sound signals are recorded from normal as well as stuttering subjects for different phonemes for Nonlinear analysis with an aim of characterising normal and stuttering vocal sound signals. Linear acoustic analysis is also carried-out on these signals. In order to study the efficiency of speech therapy for two weeks (post-treatment) rendered on the stuttered subjects the nonlinear and linear acoustic analysis is carried out on the vocal signals viz. the vowels /അ/, /ഇ/ and /ഉ/ and on the sentence "ഇരുവരും വലിയ ചങ്ങാതിമാരാണ്" the meaning in English is "Both are good friends". The results indicate that PE is an efficient measure for characterising the audio signal and establishing quantitatively the improvement in the progress of post treatment rendered by speech therapist and clinicians to Persons With Stammering(PWS) time to time.The results are verified using the evolution of PE, maximum lyapunov exponent(λ_{max}), Total PE, Correlation Dimension(D_2) Kolmogorov entropy (K_2) and interdependence of phonation number.

4.1 STUTTERING SPEECH SIGNAL ANALYSIS:

Generally, classification of speech disorders is considered as a very difficult complex problem, however some typical artifacts associated with stuttering are commonly recognized. The speech rhythm is determined as sequence of durations of vowels and consonants in speech [1]. Stuttering is the subject of interest of researchers in various domain like speech physiology and pathology, psychology, acoustic signal analysis. It is a poorly understood communication disorder with 1% global prevalence [2]. It is a complex multidimensional speech fluency disorder resulting from the disco-ordinaion in articulation, phonation and breathing [3-5]. Van Riper defines stuttering as a temporal disruption of the simultaneous and successive programming of muscular movements required to produce a speech sound or its link to the next sound [5].

Traditional analysis techniques of stuttererred speech are based on estimation of mean duration of syllable repetitions and other related measures of frequency of different speech disfluency or mean number of repeated units per instances of sound or syllable [6]. There are some cases of so called "hidden stuttering" when the speaker practices word avoidance. Also, there often are blockades in the stutterer's speech, which may be judged by the listeners as normal pauses in speaking. Because of this stutterers are currently evaluated using visual methods also [6]. Other methods are based on extracting acoustic information from the intensity time profile [7-20]. These studies concentrated on short fragments of utterances, the approximate duration of one fragment being one or a few phonemes.

Eventhough several such methods have been proposed for evaluation of patients progress in speech therapy [21-25] most of them are not suitable for automatic characterization. Of late several attempts have been made to develop methods that can assist in automatic detection and characterization of speech [21]. Conventionally, to assess and classify dysfluency is to translate and use audiovisual cues on the recorded speech and to locate occurrences of repetitions, syllable/word injections and prolongation of speech [8]. Good results are achieved using motor measures [15 25]. It is indicated that movement of jaw is the best measure and allowing easy detection in a study carried out by [25] of relationships between speech motor deficiencies and stuttering severity. On the other hand, from an acoustic point of view, it is possible to analyse an electric signal representing disordered speech. The results can form the basis for an in-depth diagnosis of the patient. Acoustic analysis of stuttered speech has attracted much attention due to its plausible application as suitable inputs to learning systems and for automatic classification of stuttered and fluent speech. Which may help in developing electronic devices to assist in speech therapy.

Other studies have been concerned with speech signal features [12 23;24 26]. Speech envelopes have been analysed by [26], resulting in some parameters that were later used in automatic search for lexical dysfluency [11]. [27] compared area under the speech envelopes of utterances of fluent and stuttered speakers. Inorder to look into significant changes in disordered speech acoustic analysis were also performed[28-32]. Recently there has been several studies to probe into the dynamical nature of stuttered and normal speech [33-40] such studies has advantages in providing important clues to ideal features to be accessed for external control. Such attempts using Hausdorff Dimension and Information Kolmogrov entropy has revealed strong correlation between Hausdorff Dimension of normal speech with that of chaotic logistic equation and fractal dimensions corresponding to boundary between bifurcation and chaos zones for stuttered speech. Nonlinear quantitative measures are found to be useful in classifying normal and pathological signals and also modeling speech signals [41-44].

Characterising and estimating the severity of stuttered speech is been equally interesting as well as complex problem. Research in this direction has potential applicability in methods for providing support to improve speech quality. For an effective evaluation and comparison of speech quality improvement clinicians and therapist should have quantitative information about the improvement that could be provided by specific methods[45]. Here, we propose to evaluate the PE [46] of the whole signal as a signal quantifying measure of the complexity of the total signal.

Nonlinear dynamical analysis of speech signal has potential application in characterizing any therapeutic correction methods; extensive description of speech before and after therapy are essential as they are based on perceptual or acoustic parameters. Dynamics of speech signal has nonlinear characteristics. Considering the dynamical similarities between normal and stuttered speech with that of different dynamical regimes of logistic equations one can infer that stuttered speech signals are less complex.

Existing nonlinear methods of Correlation Dimension and Kolmogorov entropy used for classifying normal and pathological speech is sensitive to noise contamination and nonstationarity. Also these algorithms are found to be time consuming. For using any quantitative measure for assessing and classifying speech disorders such methods should essentially be robust to noise and allow fast evaluation. This is more so for classifying and evaluating severity of stuttered speech. Here we propose the use of a fast nonlinear analysis technique viz. Permutation Entropy (PE) [46] analysis, to detect the dynamics of stuttering from speech audio signal captured using microphone. PE is a complexity measure which is robust against dynamical as well as observational noise [47][48]. It is a regularity statistic which relies on the order relations between neighboring values of a time series and is applicable to any real world data. It gives quantitative information about the complexity of a time series. Thus the variation of PE as a function of time can effectively indicate dynamical change. Corresponding to r values in the logistic equation, where there is a change in chaotic state PE drops indicating more regularity. This confirms the sensitiveness of PE to change in dynamics of any type of signal.

4.2 METHOD AND RESULTS OF STUDY ON FLUENT AND STUTTERED SUBJECTS.

Nonlinear measures like Correlation dimension(D_2), Maximum Lyapunov exponents(λ_{max}) and Kolmogorov entropy(K_2) are calculated for each of the signals vowel /അ/ ,vowel /ഇ/ and vowel /ഉ/ of Malayalam language. Permutation entropy(PE) analysis is also computed on each of the vocal samples. In addition, conventional acoustic analysis is also performed on the signals using zero crossing, fundamental frequency(F0), perturbation methods like jitter and shimmer.

Tables 4.4a, to 4.6b summarises the results of the nonlinear analysis evaluated for the recorded speech signal in Malayalam language . Vowels /ആ/, /ഈ/ and /ഉ/ from ten normal subjects and ten sutterered subject are analysed. The first, second and third columns shows Maximum Lyapunov exponent(λ_{max}), Correlation dimension(D_2), Kolmogrov entropy(K_2) . Fourth columns shows the PE of order 5 calculated for the complete signal. The data length of the signals varies between 2000 and 4000 samples for all the vowels for normal subjects . It is between 3328 to 6524, 3914 to 6769 and 3453 to 7736 samples for vowels /ആ/, /ഈ/, /ഉ/ respectively for stuttering subjects.

VOWEL "ആ"				
NORMAL	**Lyap_spec (λ_{max})**	**D_2**	**K_2**	**Total PE**
SUBJECT 11	0.069	2.46	0.03	0.72
SUBJECT 12	0.072	2.34	0.02	0.70
SUBJECT 13	0.052	2.55	0.09	0.71
SUBJECT 14	0.059	2.34	0.03	0.66
SUBJECT 15	0.052	2.26	0.02	0.60
SUBJECT 16	0.05	2.19	0.01	0.66
SUBJECT 17	0.072	2.33	0.11	0.72
SUBJECT 18	0.051	2.38	0.04	0.67
SUBJECT 19	0.034	2.05	0.02	0.61
SUBJECT 20	0.108	1.55	0.03	0.65
Average	0.062	2.25	0.04	0.67

Table 4.4a Nonlinear parameters for vowel /ആ/ for normal subjects.

VOWEL "ആ"				
STUTTERING	Lyap_spec (λ_{max})	D_2	K_2	Total PE
SUBJECT 1	0.045	1.51	0.032	0.59
SUBJECT 2	0.062	0.71	0.041	0.59
SUBJECT 3	0.047	1.19	0.080	0.65
SUBJECT 4	0.055	1.91	0.021	0.67
SUBJECT 5	0.041	0.92	0.043	0.65
SUBJECT 6	0.034	1.47	0.019	0.59
SUBJECT 7	0.039	1.24	0.034	0.60
SUBJECT 8	0.034	0.87	0.013	0.68
SUBJECT 9	0.069	1.22	0.007	0.68
SUBJECT 10	0.134	0.88	0.045	0.66
Average	0.056	1.19	0.034	0.64

Table 4.4b Nonlinear parameters for vowel /ആ/ before therapy for stuttering subjects.

From table 4.4a and 4.4b it can be inferred that the speech signals of stuttering subjects are less complex in nature compared to those of normal subjects . In the case of vowel /a/ the mean value of maximum lyapnaov exponent for normal subjects is 0.062 whereas for stuttering group it is 0.056. The mean value of correlation dimension for normal group is 2.25 whereas it is 1.19 in the case of stuttered group. A similar lowering of average mean value is observed for kolmogorov entropy which is 0.04 for normal and 0.03 for stutterered group. The PE value is in accordance with the above measures . In the case of vowel /ആ/ the average mean value of PE for normal subjects and stutterring group are 0.67 and 0.64 respectively. Similar decrease is observed in case of vowels /e/ and /u/ also (Table 4.5a and 4.5b and 4.6a and 4.6b). The mean value of maximum lyapunov exponent of stuttered subjects for vowel /ഉ/ is found to be 0.284 whereas it 0.451 for normal subject groups. Similarly correlation dimension of 1.87 is lowered to 1.25 from normal to stuttered group. A similar decrease is observed in mean value of kolmogorov entropy, which is 0.032 for normal subject group and 0.0152 for stuttered group. The value of PE also evidently shows the lowered complexity of the signal. Also average value of PE of stuttered group is found to be 0.61 which is much lower than 0.76 of normal subject group. Similarily for vowel /ഉ/ all the nonlinear quantitative parameters showed a decrease for the stuttering subject group indicating lowered complexity of speech signal of stutterers. Average value of maximum lyapunov exponent for the vowel /ഉ/ is 0.038 and 0.036 for normal and stuttered groups respectively. Corresponding values of average correlation dimension of the two groups are 2.41 and 1.45. Average kolmogorov entropy for normal and stuttered groups are 0.02 and 0.018 respectively. Corresponding values of PE for these groups are 0.47 and 0.40.

VOWEL "ഊ"				
NORMAL	**Lyap_spec (λ_{max})**	**D2**	**K2**	**Total PE**
SUBJECT 11	0.179	1.43	0.02	0.73
SUBJECT 12	0.026	1.22	0.017	0.76
SUBJECT 13	0.038	1.54	0.094	0.79
SUBJECT 14	0.033	1.57	0.034	0.80
SUBJECT 15	0.017	2.65	0.025	0.68
SUBJECT 16	0.025	2.45	0.094	0.61
SUBJECT 17	0.014	1.93	0.023	0.84
SUBJECT 18	0.003	1.33	0.038	0.77
SUBJECT 19	0.086	2.05	0.027	0.77
SUBJECT 20	0.03	2.54	0.065	0.82
Average	0.451	1.87	0.032	0.76

Table 4.5a Nonlinear parameters for vowel /ഊ/ for normal subjects.

VOWEL "ഈ"				
STUTTERING	Lyap_spec (λ_{max})	D2	K2	Total PE
SUBJECT 1	0.026	0.58	0.004	0.53
SUBJECT 2	0.035	1.41	0.029	0.73
SUBJECT 3	0.048	1.00	0.032	0.75
SUBJECT 4	0.016	1.60	0.013	0.56
SUBJECT 5	0.029	1.23	0.028	0.60
SUBJECT 6	0.041	1.10	0.013	0.57
SUBJECT 7	0.023	1.55	0.006	0.66
SUBJECT 8	0.021	1.64	0.006	0.58
SUBJECT 9	0.029	1.10	0.010	0.57
SUBJECT 10	0.016	1.23	0.011	0.56
Average	0.284	1.25	0.0152	0.61

Table 4.5b Nonlinear parameters for vowel /ഈ/ before therapy for stuttering subjects.

VOWEL "ഉ"				
Normal	Lyap_spec (λ_{max})	D2	K2	Total PE
SUBJECT 11	0.034	2.46	0.032	0.50
SUBJECT 12	0.054	2.48	0.024	0.47
SUBJECT 13	0.023	2.76	0.074	0.46
SUBJECT 14	0.051	2.34	0.045	0.46
SUBJECT 15	0.023	2.24	0.058	0.38
SUBJECT 16	0.037	2.35	0.036	0.44
SUBJECT 17	0.046	2.42	0.030	0.50
SUBJECT 18	0.032	2.42	0.020	0.56
SUBJECT 19	0.044	2.26	0.006	0.44
SUBJECT 20	0.033	2.35	0.033	0.50
Average	0.038	2.41	0.02	0.47

Table 4.6a Nonlinear parameters for vowel /ഉ/ for normal subjects.

VOWEL "ഉ"				
STUTTERING	**Lyap_spec (λ_{max})**	**D2**	**K2**	**Total PE**
SUBJECT 1	0.029	1.78	0.012	0.43
SUBJECT 2	0.038	2.40	0.002	0.45
SUBJECT 3	0.043	0.87	0.005	0.47
SUBJECT 4	0.026	0.81	0.017	0.45
SUBJECT 5	0.043	0.88	0.012	0.35
SUBJECT 6	0.044	0.79	0.007	0.37
SUBJECT 7	0.025	1.66	0.067	0.41
SUBJECT 8	0.032	1.65	0.009	0.36
SUBJECT 9	0.045	2.35	0.010	0.37
SUBJECT 10	0.044	1.33	0.139	0.34
Average	0.036	1.45	0.018	0.40

Table 4.6b Nonlinear parameters for vowel /ഉ/ before therapy for stuttering subjects.

The Lyapunov exponent, Correlation dimension and Kolmogorov entropy, all of which estimates the complexity of the system from various characteristics of the attractor of the system and PE which estimates the complexity of the system based on symbolic characteristics of the time series shows a lowered value for stuttering subjects compared to that of the normal subjects. The mean value of maximum lyapunov exponent of stuttering subjects shows a lower value compared to the normal subjects indicating lowered complexity in the stuttering signals of these subjects for all the vowels. Similar behaviour is observed in the case of Correlation dimension, Kolmogorov entropy and PE. The decrease in these nonlinear quatitative measures indicates a lowered complexity or increased regularity in these signals. Phonetic analysis of stuttering behavior has evidences of predominance of pure repetitions[50]. This can be treated as the existence of repetitive blocks of same syllables which will lead to increased regularity of the signals. These results reinforces the efficiency of nonlinear measures in distinguishing the dynamical differences between normal and stuttered speech signals.

VOWEL "ഔ"					
NORMAL	Time (ms)	ZCross	F0(Hz)	%jit	%shm
SUBJECT 11	300	27	845.6	4.99	32.57
SUBJECT 12	310	25	911.4	4.65	27.49
SUBJECT 13	410	9	735.4	6.17	18.58
SUBJECT 14	340	22	819.1	5.3	44.48
SUBJECT 15	300	20	742	1.78	11.38
SUBJECT 16	280	22	760.1	2.93	22.71
SUBJECT 17	340	37	880.2	3.32	26.1
SUBJECT 18	340	23	770.2	6.98	56.62
SUBJECT 19	300	34	854	2.99	16.33
SUBJECT 20	340	36	812	3.73	45.65
Average		25.5 ± 8.5	813 ± 60.2	4.284 ± 1.61	30.19 ± 14.56

Table 4.1a linear parameters zero crossing rate, Fundamental Frequency (F0), jitter and shimmer for evaluating acoustic properties for signal (vowel /ഔ/).

VOWEL "ഔ"					
STUTTERING	Time (ms)	ZCross	F0(Hz)	%jit	%shm
SUBJECT 1	400	14	734.1	4.12	34.34
SUBJECT 2	420	18	644	3.34	18.82
SUBJECT 3	590	14	700.8	6.77	29.74
SUBJECT 4	480	21	880.7	4.06	24.91
SUBJECT 5	350	17	769.4	4.17	24
SUBJECT 6	520	11	715.7	6.3	27.16
SUBJECT 7	580	10	659.6	32.74	37.59
SUBJECT 8	510	16	720	6.93	31.2
SUBJECT 9	460	15	650.9	21	39.33
SUBJECT 10	460	17	877.4	2.94	26.22
Average		15.3 ± 3.3	735.26 ± 85.3	9.24 ± 9.79	29.33 ± 6.40

Table 4.1b linear parameters zero crossing rate, Fundamental Frequency F(0), jitter and shimmer for evaluating acoustic properties for signal (vowel /ഔ/).

Vowel "ஐ"					
NORMAL	Time (ms)	ZCross	F0(Hz)	%jit	%shm
SUBJECT 11	290	19	383.9	4.49	14.31
SUBJECT 12	270	22	331.8	3.35	14.91
SUBJECT 13	330	4	277.1	3.86	19.42
SUBJECT 14	310	11	361.2	4.45	18.35
SUBJECT 15	300	11	286.1	2.28	11.64
SUBJECT 16	280	9	351	1.93	12.31
SUBJECT 17	410	13	373.3	5.21	16.95
SUBJECT 18	320	21	281.9	11.95	47.17
SUBJECT 19	330	12	353	6.62	37.6
SUBJECT 20	390	17	313.2	5.25	45.45
Average		13.9±5.72	331.25±39.4	4.94±2.83	23.81±13.94

Table 4.2a linear parameters zero crossing rate, Fundamental Frequency (F0), jitter and shimmer for evaluating acoustic properties for signal (vowel /ஐ/).

Vowel "ஐ"					
STUTTERING	Time (ms)	ZCross	F0(Hz)	%jit	%shm
SUBJECT1	350	9	237.5	3.2	16.41
SUBJECT2	460	1	340.3	4.56	33.61
SUBJECT3	540	2	270.3	6.4	22.13
SUBJECT4	540	12	331.5	5.77	41.23
SUBJECT5	390	2	305.8	3.85	41.48
SUBJECT6	410	5	390.5	3.73	19.95
SUBJECT7	580	3	381.1	18.68	33.47
SUBJECT8	530	17	337.2	4.87	24.48
SUBJECT9	410	3	287.5	3.02	19.92
SUBJECT10	370	11	392.2	1.96	12.49
Average		6.5±5.42	327.39±52.43	5.60±4.77	26.52±10.2

Table 4.2b linear parameters zero crossing rate, Fundamental Frequency F(0), jitter and shimmer for evaluating acoustic properties for signal (vowel /ஐ/).

Vowel "ഉ"					
NORMAL	Time (ms)	ZCross	F0(Hz)	%jit	%shm
SUBJECT11	310	12	381.2	6.17	25.59
SUBJECT12	330	9	351.1	11.06	52.26
SUBJECT13	310	4	406.3	5.44	17.5
SUBJECT14	260	12	375.7	3.87	20.24
SUBJECT15	400	6	356	24.63	31.81
SUBJECT16	250	13	399.7	2.44	11.95
SUBJECT17	410	13	450	3.72	18.51
SUBJECT18	260	13	383.8	4.07	23.87
SUBJECT19	280	13	383.8	3.28	20.31
SUBJECT20	350	12	363.1	3.63	18.45
Average		10.7 ± 3.26	385.07 ± 28.80	6.83 ± 6.71	24.05 ± 11.2

Table 4.3a linear parameters zero crossing rate, Fundamental Frequency F(0), jitter and shimmer for evaluating acoustic properties for signal (vowel /ഉ/).

	Vowel "ഉ"				
Stuttering	Time (ms)	ZCross	F0(Hz)	%jit	%shm
SUBJECT 1	310	9	338.6	5.9	44.53
SUBJECT 2	320	7	375.4	1.67	14.88
SUBJECT 3	430	12	397.4	6.36	44.08
SUBJECT 4	500	12	359.6	3.99	19.38
SUBJECT 5	440	10	365.9	5.07	47.17
SUBJECT 6	460	7	410.6	4.99	43.78
SUBJECT 7	700	7	318.9	18.89	36.88
SUBJECT 8	510	9	371.8	3.96	30.35
SUBJECT 9	500	8	342.5	3.97	30.77
SUBJECT 10	310	11	420.2	1.61	15.74
Average		9.2 ± 1.98	370.09 ± 32.3	5.64 ± 4.91	32.75 ± 12.52

Table 4.3b linear parameters zero crossing rate, Fundamental Frequency F(0), jitter and shimmer for evaluating acoustic properties for signal (vowel /ഉ/).

Tables 4.1a to 4.3b summarises the results of the acoustic analysis evaluated for the recorded speech signal /അ/, /ഇ/ and /ഉ/ for ten normal subjects and ten stuttering subjects. Table 4.1a and b shows the acoustic parameters of normal and stuttering subjects respectively for vowel /അ/ . Similarily Table 4.2a, b and 4.3a, b summarizes the corresponding parameters for vowels /ഇ/ and /ഉ/ respectively.

From Table 4.1a and 4.1b it can be inferred that the average value of fundamental frequency F0 of all the vowels is lower for the stuttering group subjects compared to the normal subjects. Average F0 for normal subject group for vowel /അ/ is found to be 813Hz whereas for the stuttering group it is found to 735Hz. In the case of vowel / ഇ / the mean value of F0 as shown in Table 4.2a and 4.2b is lower for the stuttering subject group. The average value of F0 for normal subject group is 331Hz whereas it is 327Hz for stuttering group. From Table 4.3a and 4.3b it can be inferred that average F0 of vowel /ഉ / is 385Hz whereas it is 370 Hz for the stuttering groups.

These results indicate that the Fundamental frequency is affected by stuttering, but the change is unique to each vowels. It was earlier observed[50] the changes in fundamental frequency depends on the type of stuttering. However, these changes showed patterns that differed according to the type and severity of stuttering behavior that followed the chosen moment of stuttering.

Average value of percentage jitter for vowel /അ/ and /ഇ/ is higher for the stuttering group compared to the normal. Average percentage Jitter for vowel /u/ is found to be similar in both groups. Percentage shimmer did not show any significant difference between the two groups for /അ/ and /ഇ/ but shows a slight increase for vowel /ഉ /. Average percentage jitter for the vowel /അ/ is 4.284 and 9.237 for normal and stuttered groups respectively. The corresponding values are 4.939 and 5.604 in the case of vowel /ഇ/. For vowel /ഉ/ percentage jitter corresponding to normal and stuttered groups are 6.83 and 5.64. Average value of percentage shimmer in the case of vowel /അ/ is 30.19 and

29.33 for normal and stuttered groups respectively. For the vowel /ഓ/ the corresponding values are 23.81 and 26.52. For the vowel /ഉ/ average value of percentage of shimmer for normal and stuttered groups are 24.05 and 32.76 respectively. Eventhough the jitter and shimmer doesn't show significant changes in all the phonations the observed changes are towards the higher side which shows that phonations of stutter's were less stable than that of normal subjects in terms of both vocal frequency and intensity.

Results of analysis using Zero Crossing Rate(ZCR) (Table 4.1a to 4.3a) on the above signals shows decrease in zero crossing rate for group of stutterers with respect to those of normal subjects in the case of all the vowels. The ZCR is simply the count of crossing the zero throw fixed window size. It is said to occur if successive samples have different algebraic signs. The ZCR is a time-domain algorithm and it deeply depends upon the frequency of the input signal. ZCR is proportional to the dominant frequency. If a signal is a pure sinusoidal waveform, the dominant frequency is the only one in the spectrum. This frequency equals the number of zero crossings of the signal in one second. In other words, it equals the value of the ZCR if the rate is taken every second. For non-sinusoidal periodic waveform, the dominant frequency has the largest amplitude. For unvoiced sounds like pure fricatives or noisy breaks zero crossing rate will assume higher values whereas for voiced sounds it will be close to zero. In the case of stuttered speech of vowels the lowered value of zero crossing rate can be attributed to the repeated utterance of the same syllable which increases the number of voiced segments [51].

Inorder to estimate the efficiency of these nonlinear parameters in accessing the improvement obtained by a particular treatment method the vocal signals of the same subject of the stuttering group are recorded and analysed after two weeks of therapy.

Vowel "അ"				
SUTTERING	**Lyap_spec (λ max)**	**D2**	**K2**	**Total PE**
SUBJECT 1	0.26	1.28	0.007	0.69
SUBJECT 2	0.21	2.27	0.026	0.60
SUBJECT 3	0.08	0.42	0.015	0.74
SUBJECT 4	0.49	1.72	0.019	0.74
SUBJECT 5	0.13	0.47	0.014	0.67
SUBJECT 6	0.19	2.46	0.14	0.68
SUBJECT 7	0.11	0.44	0.021	0.63
SUBJECT 8	0.07	1.22	0.006	0.63
SUBJECT 9	0.09	0.21	0.0026	0.73
SUBJECT 10	0.99	0.63	0.047	0.79
Average	**0.26**	**1.11**	**0.030**	**0.69**

TABLE 4.7 Nonlinear parameters for vowel "അ" after two weeks speech therapy.

Vowel " ഉ "				
SUTTERING	**Lyap_spec (λ_{max})**	**D_2**	**K_2**	**Total PE**
SUBJECT 1	0.029	0.79	0.02	0.79
SUBJECT 2	0.02	1.51	0.021	0.78
SUBJECT 3	0.031	0.466	0.004	0.84
SUBJECT 4	0.044	2.96	0.015	0.69
SUBJECT 5	0.037	0.461	0.0084	0.86
SUBJECT 6	0.321	0.749	0.011	0.85
SUBJECT 7	0.218	0.578	0.015	0.68
SUBJECT 8	0.119	0.912	0.019	0.82
SUBJECT 9	0.079	0.276	0.005	0.71
SUBJECT 10	0.208	0.672	0.024	0.77
Average	0.55	**0.937**	**0.0142**	**0.78**

TABLE 4.8 Nonlinear parameters for vowel "ഉ" after two weeks speech therapy.

Vowel " ഉ "				
SUTTERING	Lyap_spec (λ_{max})	D_2	K_2	Total PE
SUBJECT 1	0.053	1.46	0.18	0.54
SUBJECT 2	0.082	2.37	0.012	0.51
SUBJECT 3	0.038	0.373	0.015	0.41
SUBJECT 4	0.227	2.71	0.007	0.54
SUBJECT 5	0.054	0.659	0.004	0.53
SUBJECT 6	0.075	1.58	0.015	0.49
SUBJECT 7	0.041	0.151	0.004	0.39
SUBJECT 8	0.037	0.539	0.0048	0.43
SUBJECT 9	0.035	0.369	0.009	0.45
SUBJECT 10	0.288	1.38	0.035	0.65
Average	**0.093**	**1.159**	**0.0286**	**0.49**

TABLE 4.9 Nonlinear parameters for vowel "ഉ" after two weeks speech therapy.

Tables 4.4a to 4.6b, 4.7, 4.8, 4.9 summarises the results of the nonlinear analysis evaluated for the recorded speech signal /അ/, /ഇ/ and /ഉ/ for ten normal subjects and ten stuttering subjects after two weeks therapy. For all the three vowels the Maximum lyapunov exponent(λ_{max}) and PE clearly indicates an increase in the post-treatment after two weeks, compared to the pre-treatment case, however correlation dimension and kolmogorov entropy don't show the same behaviour. For a comparison of the results conventional acoustic analysis are conducted on these signals.

In the case of vowel /അ/ , the mean value of maximum lyapunov exponent for normal, stuttered subjects before treatment and stuttered subject after two weeks of therapy (Table 4.4a, 4.4b and 4.7) are 0.062 , 0.056 and 0.26 respectively . For the vowel /ഇ/ the corresponding values are 0.451, 0.284 and 0.55 respectively (Table 4.5a, 4.5b and 4.8). In the case of letter /ഉ/ these values are 0.038, 0.036 and 0.093 respectively. (Table 4.6a, 4.6b and 4.9).

Average Correlation dimension for vowel /അ/ for normal, stuttered subjects before and after therapy are 2.25, 1.19 and 0.26 respectively. Corresponding values of vowel /ഈ/ are 1.87, 1.25, and 0.94 respectively. For vowel /ഉ/ these values are 2.41, 1.45 and 1.159 respectively. Average value of Kolmogorov entropy of vowel /അ/ for normal subjects, stuttering before and after treatment are 0.04, 0.034 and 0.03 respectively. In the case of vowel /ഈ/ these values are 0.032, 0.0152 and 0.0142 respectively. Corresponding values for the vowel /ഉ/ are 0.02, 0.018 and 0.029 respectively.

Average value of PE of the vowel /അ/ for normal subjects, stuttering subjects before and after treatment are 0.67, 0.64 and 0.69 respectively. Correspondingly in the case of vowel /ഈ/ are 0.76, 0.61 and 0.78 respectively. In case of vowel /ഉ/ these values are 0.47, 0.40 and 0.49 respectively.

From the above results it can be observed that all the nonlinear measures are efficient in differentiating normal subjects from sutterered subjects. All these measures shows a decrease in their values for the stutterering subject group indicating increased regularity which can be attributed to the repeated blocks of the same syallable.

However, in accessing the dynamic changes in speech signals of the stuttering subjects after speech therapy, some of the conventional measures do not give conclusive results. Among these measures the lyapunov exponent and PE shows indication of improvement in speech signal after therapy with the change in its value towards normal whereas, the measures like correlation dimension and kolmogorov entropy are not so efficient. These measures shows much higher values compared to even normal subjects. Correlation dimension and kolmogorov entropy are measures which are very sensitive to noise contamination of the signal. During the therapy the subjects are trained to control the stuttering by implementing slow rate and light articulation of utterances. This incorporates more silent blocks in their speech which would have earlier been repeated blocks of syllables. These silent blocks will present as background noise to the assessment system, thereby increasing their values of quantifiers. Thus the higher values of the conventional nonlinear measures in post treatment cases than the normal subjects can be attributed to the presence of silent blocks after therapy.

VOWEL "ഒ"					
STUTTERING	Time (ms)	ZCross	F0(Hz)	%jit	%shm
SUBJECT 1	360	31	845	24.48	38.8
SUBJECT 2	410	32	714	19.15	34.6
SUBJECT 3	310	24	724	1.04	12.83
SUBJECT 4	370	40	934	1.42	5.48
SUBJECT 5	370	26	708	2.2	12.06
SUBJECT 6	310	32	736	0.61	6.1
SUBJECT 7	310	20	745	3.52	12.17
SUBJECT 8	380	29	776	2.75	25.94
SUBJECT 9	480	22	715	17.2	33.25
SUBJECT 10	390	37	895	1.65	9.92
Average		29.3 ± 6.4	779.2 ± 82.49	7.4 ± 9.1	19.12 ± 12.6

Table 4.10 linear parameters zero crossing rate, Fundamental Frequency (F0), jitter and shimmer for evaluating acoustic properties for signal (vowel /ഒ/) after 2 weeks therapy.

VOWEL "ഉ"					
STUTTERING	Time (ms)	ZCross	F0(Hz)	%jit	%shm
SUBJECT1	370	12	360	4.32	15.08
SUBJECT2	420	13	331	1.46	11.94
SUBJECT3	400	12	311	2.84	15.85
SUBJECT4	360	14	353	1.31	5.69
SUBJECT5	450	12	320	3.68	31.29
SUBJECT6	350	23	435	1.25	7.34
SUBJECT7	350	13	322	2.16	19.06
SUBJECT8	410	16	348	4.2	19.46
SUBJECT9	490	9	317	3.21	12.46
SUBJECT10	490	12	367	2.16	9.46
Average		13.6 ± 3.74	346.4 ± 36.7	2.66 ± 1.16	14.76 ± 7.40

Table 4.11 linear parameters zero crossing rate, Fundamental Frequency F(0), jitter and shimmer for evaluating acoustic properties for signal (vowel /ഉ/) after 2 weeks therapy.

Vowel "ഉ"					
STUTTERING	Time (ms)	ZCross	F0(Hz)	%jit	%shm
SUBJECT 1	350	12	352	3.3	1.36
SUBJECT 2	440	14	422	1.2	9.98
SUBJECT 3	340	13	398	1.39	9.76
SUBJECT 4	350	13	62	7.5	14.46
SUBJECT 5	370	12	309	2.52	14.54
SUBJECT 6	280	15	438	0.34	2.21
SUBJECT 7	250	12	402	1.6	13.84
SUBJECT 8	390	15	391	1.42	15.74
SUBJECT 9	500	13	368	1.97	11.52
SUBJECT 10	390	16	446	2.52	8.71
Average		13.5 ± 1.43	358.8 ± 112.05	2.38 ± 1.98	10.21 ± 5.02

Table 4.12 linear parameters zero crossing rate, Fundamental Frequency F(0), jitter and shimmer for evaluating acoustic properties for signal (vowel /ഉ/) after 2 weeks therapy.

Table 4.10, 4.11 and 4.12 shows the results of acoustic analysis of speech signals of the same group of stuttering subjects recorded after two weeks of therapy. In the case of post-treatment of vowel /അ/ the mean value of fundamental frequency F0 of stuttering group increased to 779Hz from 735hz but is still lower than the mean F0 of the normal group which is 813Hz (Table 4.1a and Table 4.10). Similarly in the case of vowel /ഇ/ the mean value of fundamental frequency F0 of stuttering group increased to 346Hz from 327Hz but is higher than the mean F0 of the normal group which is 331Hz (Table 4.2a and Table 4.11). But in the case of vowel /ഉ/ the mean value of fundamental frequency F0 of stuttering group decreased from 370Hz to 358Hz and is lower than the mean F0 of the normal group which is 385Hz (Table 4.3a and Table 4.12).

The zero crossing rate is higher for all the vowels after two weeks therapy. The average value of zero crossing rate for vowel /അ/ is found to increase from 15 to 29 (Table 4.1a and 4.10).Similar is the case for vowel /ഇ/ and vowel /ഉ/ which is found to increase from 6 to 13 and 9 to 13 respectively(Table 4.2a and b) (Table 4.3a and 4.11). The percentage of Jitter value is lower in all cases for post treatment case. Shimmer is also found to decrease for all the vowels in the post treatment.

From a perturbation point of view, short utterances are preferred to longer once which require a minimum duration of speech to stabilize long term measurements of perturbation. Therefore we propose to analyze the effectiveness of nonlinear measures in characterizing stuttered speech. Considering the non-stationary nature of continuous speech data in addition to convention measures like maximum Lyapunov exponent, Correlation dimension, Kolmogorov entropy, a new measure robust to noise and non-stationarity namely Permutation Entropy(PE) is also evaluated.

Eventhough, single phonetic analysis is widely used for characterizing vocal pathologies analysis of longer utterances are found to a better approach for characterizing stuttered speech. Continuous speech produces interactive segmental affects of articulator movement and multiple voicing onsets and offsets. Various studies based on length of utterance has shown that stuttering is less likely to occur on shorter utterance compared to longer utterance [52-58]also stuttering speakers exhibited more stuttering on the first word of the phrase compared to the single word [59].

Nonlinear measures like Correlation dimension, Lyapunov exponents and Kolmogrov entropy are calculated for a recorded sentence "ഇരുവരും വലിയ ചങ്ങാതിമാരാണ്" in Malayalam language, the meaning in English is "Both are good friend" for the same 10 normal and 10 stuttering subjects . Permutation entropy is also calculated on the vocal sample sentence signal before speech therapy and after two weeks therapy. In addition conventional acoustic analysis is performed on the sentence (signal) using zero crossing rate, fundamental frequency (F0), perturbation methods like jitter, shimmer .

Sentence “ഇരുവരും വലിയ ചങ്ങാതിമാരാണ്”				
Normal	**Lyap_spec (λ_{max})**	**D2**	**K2**	**Total PE**
SUBJECT 11	0.098	0.58	0.003	0.71
SUBJECT 12	0.063	0.3	0.009	0.73
SUBJECT 13	0.062	0.96	0.01	0.75
SUBJECT 14	0.067	0.34	0.009	0.69
SUBJECT 15	0.072	0.24	0.008	0.76
SUBJECT 16	0.041	0.49	0.009	0.73
SUBJECT17	0.719	0.53	0.037	0.75
SUBJECT 18	0.054	0.17	0.0134	0.82
SUBJECT 19	0.074	0.27	0.004	0.68
SUBJECT 20	0.063	0.87	0.007	0.75
Average	0.13	0.48	0.01	0.74

Table 4.16 Average value of nonlinear parameters calculated for a sentence of Normal subjects.

Sentence “ഇരുവരും വലിയ ചങ്ങാതിമാരാണ്”				
Stuttering	**Lyap_spec (λmax)**	**D_2**	**K_2**	**Total PE**
SUBJECT 1	0.074	0.39	0.0111	0.61
SUBJECT 2	0.091	0.25	0.008	0.65
SUBJECT 3	0.031	0.41	0.009	0.66
SUBJECT 4	0.072	0.08	0.002	0.62
SUBJECT 5	0.4187	0.71	0.002	0.6
SUBJECT 6	0.074	0.22	0.012	0.59
SUBJECT 7	0.054	0.19	0.001	0.65
SUBJECT 8	0.062	0.44	0.01	0.7
SUBJECT 9	0.069	0.35	0.004	0.65
SUBJECT 10	0.078	0.15	0.002	0.7
Average	0.102	0.32	0.006	0.64

Table 4.17 Nonlinear parameters for stuttering subjects calculated before therapy for sentence

Stuttering	Sentence "ഇരുവരും വലിയ ചങ്ങാതിമാരാണ്"			
	Lyap_spec (λ max)	**D2**	**K2**	**Total PE**
SUBJECT 1	0.088	0.754	0.006	0.65
SUBJECT 2	0.093	1.06	0.005	0.65
SUBJECT 3	0.123	1.91	0.005	0.62
SUBJECT 4	0.104	0.179	0.049	0.63
SUBJECT 5	0.124	2.62	0.00892	0.72
SUBJECT 6	0.098	0.39	0.00954	0.66
SUBJECT 7	0.082	1.82	0.0304	0.65
SUBJECT 8	0.157	0.372	0.0094	0.7
SUBJECT 9	0.11	0.253	0.0172	0.65
SUBJECT 10	0.119	0.64	0.0113	0.7
Average	0.11	1.00	0.015	0.66

Table 4.18 Nonlinear parameters for stuttering subjects calculated after therapy for sentence

Tables 4.16 , 4.17 and 4.18 summarises the results of the Nonlinear analysis evaluated for the recorded speech signal "ഇരുവരും വലിയ ചങ്ങാതിമാരാണ്" for ten normal subjects and ten stuttered subjects before therapy and after therapy. The first, second and third coloumns shows Maximum lyapnov exponent(λ_{max}) , Correlation dimension(D_2) and Kolmogrov entropy(K_2) . Fourth coloumn shows the PE of order 7 calculated for the complete signal. The data length of the signals of the normal subjects varies between 15277 and 22656 and that of stuttered subjects varies between 17030 and 22444 samples. The corresponding data length of the signals varies between 22006 and 51021 samples for stuttered subjects after two weeks therapy.

The mean value of maximum lyapunov exponent of the stuttered subjects shows a lower value compared to the normal subjects indicating increased regularity in the vocal signals of these subjects. For normal subjects mean value of maximum lyapunov exponent is 0.13 whereas it is 0.10 for stuttered subjects as given in Table 4.16 and 4.17. Similar decrease is found in the case of correlation dimension and kolmogrov entropy . Average correlation dimension for normal and stuttered subjects are 0.48 and 0.32 respectively as shown in Table 4.16 and 4.17. The mean Kolmogorov entropy for normal subjects is 0.01 and stuttered subjects is 0.006 as given in Table 4.16 and 4.17 . The new regularity measure PE shows a lowering in its value in stuttering cases indicating an increase in regularity. Average value of PE for normal and stuttered subjects are 0.74 and 0.64 respectively as given in Table 4.16 and 4.17 . The decrease in these nonlinear quatitative measures points towards the increased regularity induced by the stuttering. These methods give clear indications of the increased regularity induced by stuttering .

Sentence "ഇരുവരും വലിയ ചങ്ങാതിമാരാണ്"					
Normal	Time (ms)	ZCross	F0(Hz)	%jit	%shm
SUBJECT 11	2180	16	519.5	3.2	28.63
SUBJECT 12	1540	18	501	3.53	31.12
SUBJECT 13	2010	19	529.7	3.36	20.42
SUBJECT 14	2210	14	452.1	4.53	28.95
SUBJECT 15	2830	12	431.6	2.98	27.07
SUBJECT 16	2210	10	428.2	3.20	28.63
SUBJECT 17	1560	15	527.3	2.66	34.32
SUBJECT 18	1580	20	406.1	4.86	36.92
SUBJECT 19	2120	17	471.8	3.34	27.05
SUBJECT 20	2120	15	501.6	16.53	46.62
Average		15.6	476.89	4.82	30.97

Table. 4.13 Average value of linear parameters calculated for a sentence for normal subjects

Stuttering	Sentence "ഇരുവരും വലിയ ചങ്ങാതിമാരാണ്"				
	Time (ms)	ZCross	F0(Hz)	%jit	%shm
SUBJECT1	2180	19	469.7	3.52	30.22
SUBJECT2	1540	12	422.5	3.05	18.75
SUBJECT3	2010	18	378	3.28	41.97
SUBJECT4	2210	20	558.6	3.66	30.44
SUBJECT5	2830	20	388.7	21.32	31.37
SUBJECT6	2210	16	512.6	3.34	39.92
SUBJECT7	1560	13	538.9	3.09	24.35
SUBJECT8	1580	20	538.6	3.19	18.90
SUBJECT9	2120	18	459.5	3.21	28.16
SUBJECT10	2120	16	555.3	3.41	28.88
Average		18	482.24	5.11	29.3

Table 4.14. Average value of linear parameters calculated for a sentence before speech therapy .

Stuttering	Sentence "ഇരുവരും വലിയ ചങ്ങാതിമാരാണ്"				
	Time (ms)	ZCross	F1(Hz)	%jit	%shm
SUBJECT1	4320	16	550.4	2.59	22.23
SUBJECT2	3030	16	491.1	1.64	13.05
SUBJECT3	3890	9	489.7	4.09	26.85
SUBJECT4	4620	25	573.4	3.79	19.66
SUBJECT5	4280	14	489.6	1.76	29.37
SUBJECT6	3510	18	481.8	3.10	31.79
SUBJECT7	1990	15	531.6	1.94	14.26
SUBJECT8	3410	18	591.7	3.09	18.68
SUBJECT9	3590	15	498.1	3.42	24.38
SUBJECT10	2920	24	587.4	3.97	15.03
Average		17	528.48	2.94	21.53

Table 4.15 Average value of linear parameters calculated for a sentence after two weeks speech therapy

For assessing, Tables 4.13, 4.14 and 4.15 summarises the results of the acoustic analysis evaluated for the recorded speech signal "ഇരുവരും വലിയ ചങ്ങാതിമാരാണ്" for ten normal subjects and ten abnormal subjects before therapy, after therapy for two weeks and normal subjects.

The fundamental frequency F0 from Table 4.13 and 4.14 show that the average value of F0 of the sentence is higher for stuttered subjects compared to the normal subjects. Average F0 for normal subject group is found to be 476Hz whereas for the stuttered group it is found to 482Hz. After two weeks speech therapy to the 10 stuttered subjects the average value of F0 of the sentence is found to increase , it is found to be 528Hz from 482Hz.

The average percentage jitter recorded while reading a sentence is higher for the stuttered group compared to the normal. Average percentage jitter for stuttered group is found to be 5.11Hz and for the normal group it is found to be 4.82Hz . After two weeks therapy the percentage of jitter is found to decrease from 5.11 to 2.94.

Percentage shimmer did not show any significant difference between the two groups. After 2 weeks therapy the average percentage of shimmer for the stuttered group is found to have a lower value compared to the normal control group. It is found to reduce from 29 to 21.

Results of analysis using Zero crossing rate done on the above sentence where found to have a lower value for stuttering group compared to the normal group. It is found to have 12 for stuttered group whereas for normal it is found to have 15 before therapy. During post treatment the average zero crossing rate is found to increase in the case of stuttered group compared to the normal group. It is found to increase from 12 to 17.

Now, for assessing the efficiency of nonlinear measures in evaluating the changes in dynamical features of speech signal of stuttered subjects after two weeks of therapy . The speech signals for the same sentence is recorded and analysed from the same subject group after two weeks of therapy. Table 4.18 shows the nonlinear parameters like lyapunov exponent, correlation dimension, kolmogorov entropy and permutation entropy. The results summarized in the table 4.18 indicate the difference in diagnostic efficiency of various parameters in characterization of stuttered signals after therapy. Comparing the results of nonlinear analysis of normal and stuttered subjects before and after therapy(Table 4.16 , Table 4.17 and Table 4.18) it can be inferred that among these measures the maximum lyapunov exponent(λ_{max}) and PE are better indicators of improvement in speech quality. The maximum lyapunov exponent of normal subjects is 0.13 whereas for stuttering subjects it is 0.10 . In the post therapy case the mean value of maximum lyapunov exponent increase towards the normal side which is indicated by its value 0.11 . The PE values of normal , pre and post treatment case also shows the same trend. Mean PE value of normal subject group is 0.74 and it is 0.64 for stuttering group. In the post treatment case it increases to 0.66 showing a tendency of suppression of the increased regularity. The other parameters like correlation dimension, kolmogorov entropy shows values of post treatment case to be much higher than the normal once.
The results of acoustic as well as nonlinear analysis on individual vowels and continuous speech data indicate the efficiency of nonlinear measures in characterizing the change in

dynamics of stuttered speech after a specific period of speech therapy. Conventional nonlinear measures like correlation dimension and kolmogorov entropy do not perform very well on raw data indicating dynamical changes. In the case of post-treatment of vowels and continuous speech data, the conventional nonlinear measures like correlation dimension and kolmogorov entropy do not perform very well. These measures being sensitive to non-stationarity and noise contamination produce ambiguous results mainly due to the presence of silent blocks in trained speech. Among the conventional nonlinear measures maximum lyapunov exponent performs slightly better. The newly introduced entropy measure mainly permutation entropy shows efficiency in indicating the improvement in speech performance after the therapy. The robustness of this measure to noise contamination and non-stationarity helps this measure in effectively characterizing the level of improvement in speech process after therapy. PE being a bounded measure is more efficient in comparing different methods of therapy.

In general the results can be summarized as follows :

- All nonlinear measures are efficient in differentiating normal subjects from stuttered subjects.
- All nonlinear measures show decrease in their value for stuttering subject groups indicating increased regularity.
- In characterizing the level of improvement after therapy from individual phonemes,
    - Conventional measures of correlation dimension(D_2), maximum lyapunov exponent and kolmogorov entropy(K_2) don't give uniform results for the vowels analysed.
- PE shows uniform results for the vowels analysed.
    - In case of post treatment the value of PE is more than the normal subjects which can be attributed to the existence of silent blocks during the trained speech.
- Linear acoustic measures of individual phonemes viz.,

    - Fundamental frequency and zero crossing rate show clear difference in its values between vowel and stuttered subjects.
    - Percentage jitter doesn't give uniform results in all cases.
    - None of the above measures are effective in characterizing the level of improvement after therapy.
- In the case of continuous speech data, all the nonlinear measures shows characteristic difference in its values between normal and stuttered subjects.
- PE in characterizing the improvement after therapy from continuous speech data
    - Among the conventional nonlinear measures only the maximum lyapunov exponent characterisies the improvement obtained other measures of correlation dimension and kolmogorov entropy don't show effectiveness in this case.
    - The new measure of entropy, Permutation Entropy (PE) clearly characterizes the improvement in speech quality by the increase in its values towards normal.

In general it can be concluded that in an online setup the use of PE in continuous speech data is recommendable for characterizing the effectiveness of various methods of speech therapy for stuttering.

PE can serve as an efficient bounded quantitative measure which can be used for comparing the treatment efficacy of various methods of therapy. Also the robustness of this measure towards non-stationarity and noise contamination of the signals enhances its applicability as an online tool. This method can be used for direct application in stuttering assessment programs of any clinical setup. It can also be used for characterizing the speech signals on real time basis for simultaneous assessment in the recording stage itself.

For investigating the variation of PE with change in dynamics of the speech signal , PE of order 7 for sliding windows of 7000 samples with one sample shift is evaluated for the normal and stuttering subjects. Figure 4.1a shows the audio wave form of the speech signal of a normal subject uttering the sentence "ഇരുവരും വലിയ ചങ്ങാതിമാരാണ്" Figure 4.1b shows the corresponding values of PE . Similarily, Figure 4.2a and 4.2b shows the audio wave form and PE values of a stuttering subject respectively for uttering the same sentence. Comparing 4.1b and 4.2b, it can be observed that the variation of PE in the case of stuttering subjects occurs within a wider range compared to that of normal subjects. For the normal subject PE values varies between 0.7 and 0.74 whereas for the stuttering subjects it 0.5 to 0.65. From Figure 4.2b it is clear that PE values corresponding to stuttering events drops indicating lowered complexity. As the speech acquires fluency and voiced sounds become clear PE gradually increases. During the silence section between the two segments where only the background noise is present, PE attains values much higher than those of voiced region. PE values of the succeeding voiced segments comes down. It can be observed that PE values corresponding to stuttering segment is much lower than that of voiced sound and those of the silence section much higher. These results indicate that for stuttering subject dynamic variations are predominant.

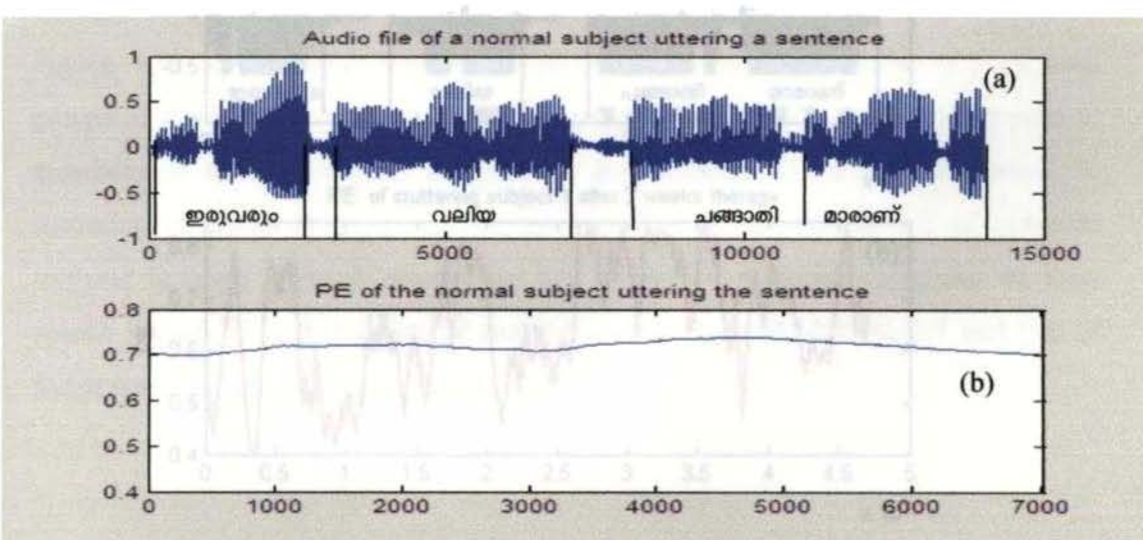


Figure 4.1.a represents the audio file of a normal subject uttering a sentence

Figure 4.1b represents the PE of a same normal subject uttering a sentence

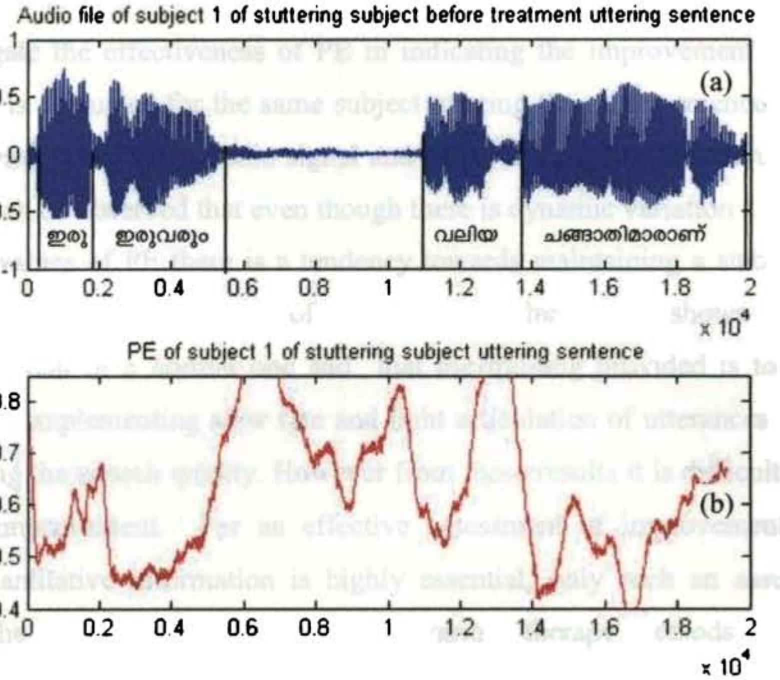


Figure 4.2a represents the audio file of a stuttering subject uttering a sentence
Figure 4.2b represents the PE of a same stuttering subject uttering a sentence

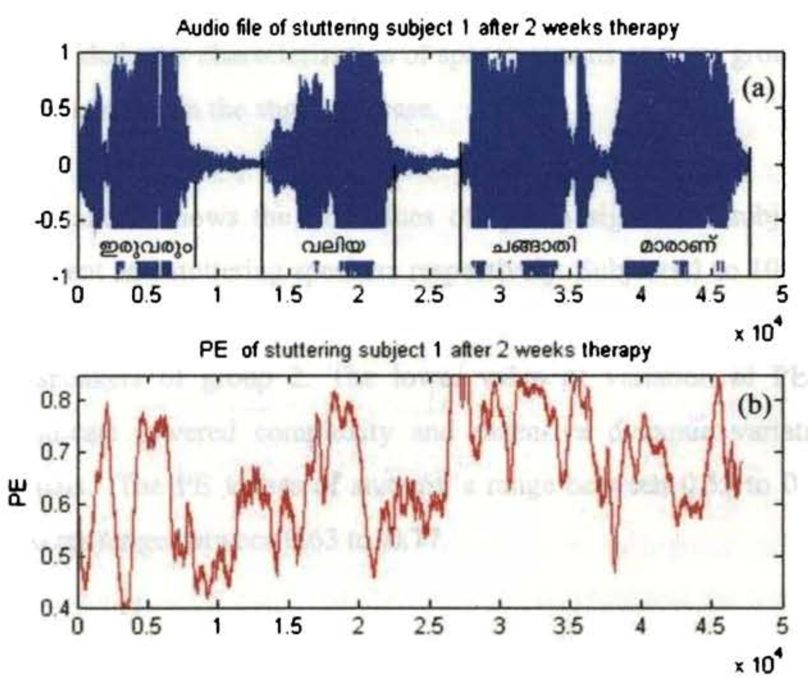
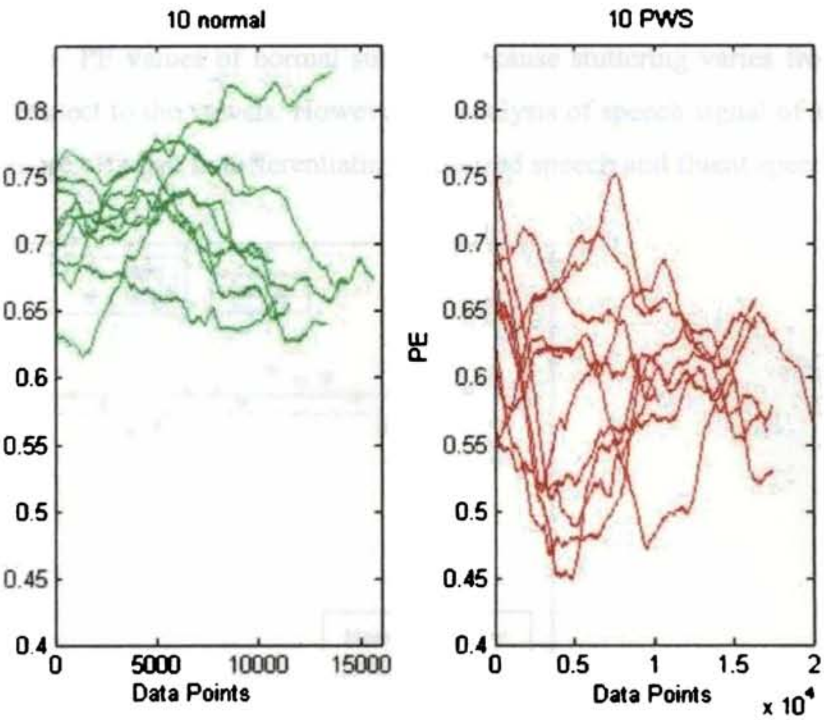


Figure 4.3a shows the audio signal for the stuttering subjects after therapy
Figure 4.3b shows the PE values for the same stuttering subject after therapy.

To investigate the effectiveness of PE in indicating the improvement in quality of speech, PE is evaluated for the same subject uttering the same sentence. Figure 4.3a show the wave form of the audio signal and figure 4.3b shows the corresponding PE values. It can be observed that even though there is dynamic variation as indicated by the varied values of PE there is a tendency towards maintaining a stable region. It can be inferred that the dynamics of speech signal after therapy shows a tendency of approaching that of a normal one and that the training provided is to control the stuttering by implementing slow rate and light articulation of utterances is effective in improving the speech quality. However from these results it is difficult to infer the extent of improvement. For an effective assessment of improvement of speech quality, quantitative information is highly essential, only such an assessment can bring out the difference in effectiveness of various therapy methods and also bring out the effectiveness of a method in different subjects. With this in mind and from the inference that the general range of PE values in the case of stuttering subjects is lower than those of normal subjects, PE of a complete speech signal is evaluated.

This can provide better characterization of speech signals of these groups and also the level of improvement in the stuttering case.

Figure. 4.4a and b shows the PE values of speech signals of subjects from both groups of fluent and stuttering speakers respectively. Subjects 1 to 10 corresponds to members of normal subjects of group 1 and subjects 11 to 20 corresponds to stuttering speakers of group 2. The lower value of variation of PE of stuttered subjects indicate lowered complexity and extensive dynamic variations of their speech signals. The PE values of stutterer's range between 0.55 to 0.7 and that of fluent speakers range between 0.63 to 0.77.



(a) (b)

Figure 4.4 Evolution of PE of speech signals of 10 Normal and 10 stuttering subjects.

For evaluating the efficiency of PE in characterizing speech signals of the above subjects, while uttering Malayalam vowel sounds /അ/ , /ഇ/ and /ഉ/ are subjected to PE analysis. PE of the complete signal corresponds to utterances of /അ/,/ഇ/ and /ഉ/ is evaluated. The total length of these signals varied between 2000 to 4000 data points for normal subjects and 2000 to 7000 data points for stutterering subjects, for vowels /അ/,/ഇ/ and /ഉ/ . For any practical purpose appropriate order should be chosen for evaluation of PE. For vowels PE of order 5 is evaluated for the total signal.

Figure. 4.5, 4.6, 4.7 shows PE of each of the 10 subjects in both group of stutterers and fluent speakers reading vowels /അ/ , /ഇ/, and /ഉ/ . PE is calculated for three vowels for each of the 20 subjects ten fluent and ten stutterers by taking the whole time series as window size and order 5. Overall PE values are higher for normal subjects compared to the stammered subjects this again confirms that the PE of stutterered subjects are lower than the fluent subjects in the case of vowels also. A few PE values of stuttering subjects are close to the PE values of normal subjects because stuttering varies from person to person with respect to the vowels. However PE analysis of speech signal of sentences are found to be more efficient in differentiating stutterred speech and fluent speech.

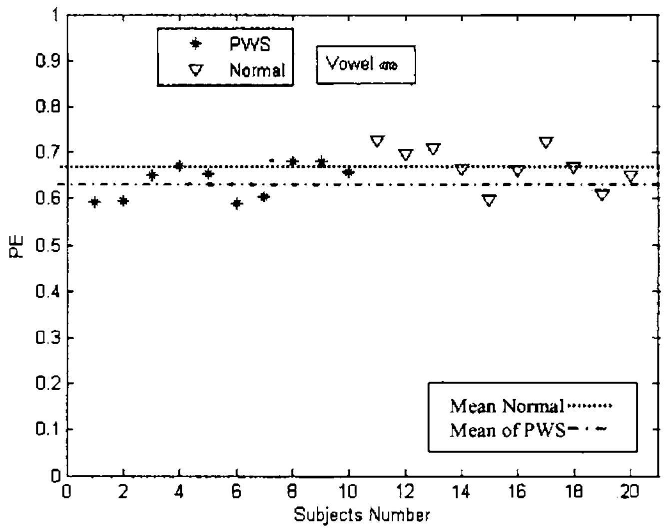


Fig.4.5. PE calculated for first 10 PWS subjects(1-10) and normal subjects(11-20) for Vowel /ഐ/.

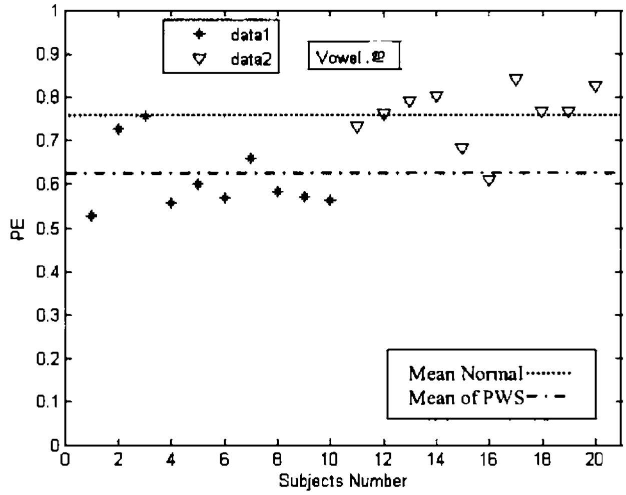


Fig. 4.6. PE calculated for first 10 PWS subjects(1-10) and normal subjects(11-20) for Vowel /ഊ/.

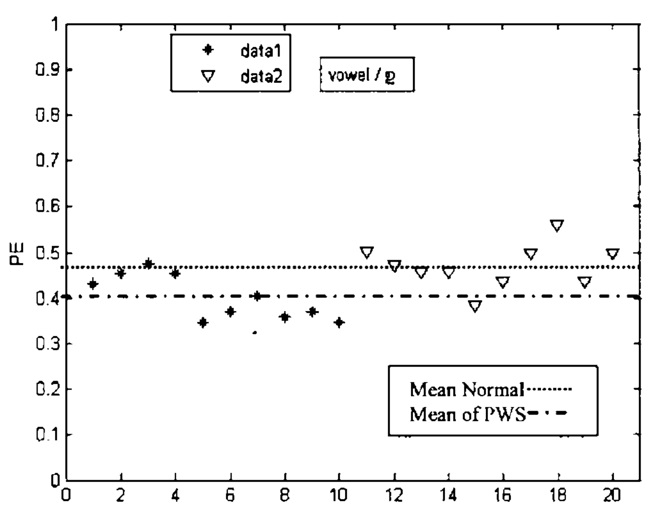


Fig. 4.7. PE calculated for first 10 PWS subjects(1-10) and normal subjects(11-20) for Vowel /ഉ/.

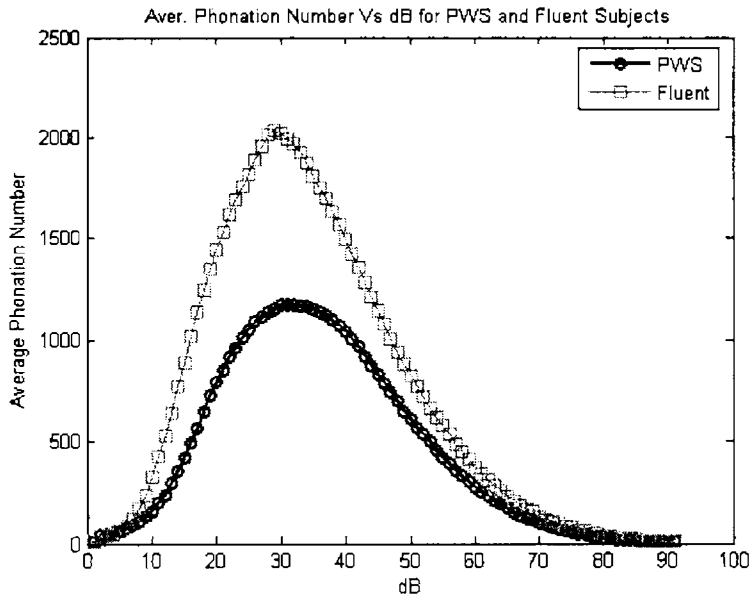


Fig.4.8 variation of average phonation number with respect to sound intensity for PWS and Fluent Subjects

In order to confirm the results obtained by PE in differentiating the group of stutterers and fluent speakers, the average of phonation number corresponding to different sound intensity level for both groups are evaluated. Figure.4.8 shows the average phonation numbers corresponding to different sound intensity levels of speech while reading the sentence " ഇരുവരും വലിയ ചങ്ങാതിമാരാണ്". A definite maximum phonation number can be observed at the level of approximately 27 dB in the case of non stutterers and 31dB in case of stutterers. The Plot in the case of stutterers appears to be flat from 27 to 35 dB where as the plot for fluent subjects shows the average phonation number increasing from 5dB to 27 dB and then decreasing .

Figure. 4.9, shows the PE of each of the 10 stuttering subjects before therapy and after two weeks therapy for uttering the sentence "ഇരുവരും വലിയ ചങ്ങാതിമാരാണ്". Results shows that there is an increase in the Total PE values calculated for the complete signal in the post treatment. Figure. 4.10 shows PE of each of the 10 subjects in both group of stutterers and fluent speakers reading the sentence "ഇരുവരും വലിയ ചങ്ങാതിമാരാണ്" . PE is calculated for reading the sentence for each of the 20 subjects ten fluent and ten stutterers by taking the whole time series as window size and order 7. Overall PE values are higher for normal subjects compared to the stammered subjects this again confirms that the PE of stutterered subjects are lower than the fluent subjects in the case of vowels also. A few PE values of stuttering subjects are close to the PE values of normal subjects because stuttering varies from person to person with respect to the vowels. However PE analysis of speech signal of sentences are found to be more efficient in differentiating stutterred speech and fluent speech.

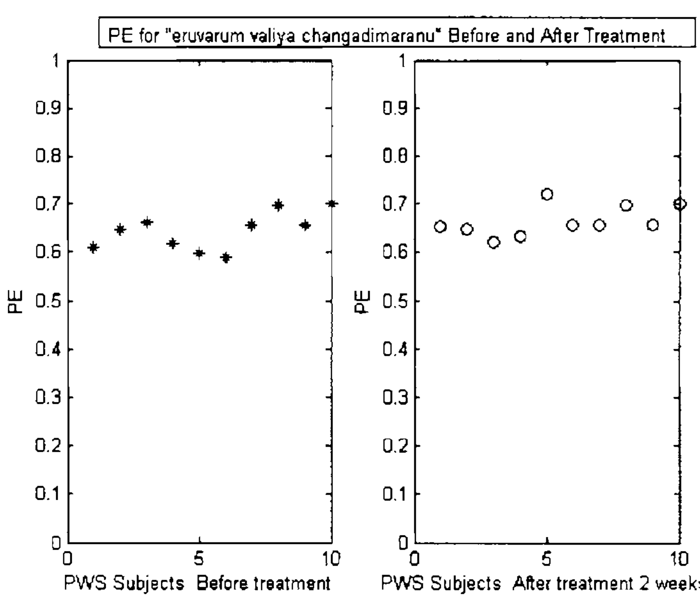


Fig. 4.9. PE calculated for stuttering subjects before treatment and after 2 weeks speech therapy

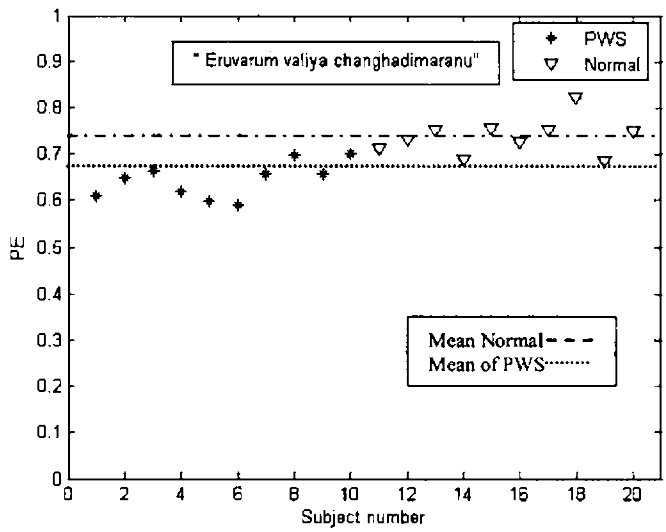


Fig. 4.10 PE calculated for first 10 PWS subjects(1-10) and normal subjects(11-20) for a sentence

4.3 REFERENCES

1. Chaos and Speech Rhythm , Proceeding of International Joint conference on Neural Network, Montreal, Canada, July 31- August 4, 2005.
2. Application of time-frequency analysis methods to speaker verification Proceedings SPIE, Vol. 6313, 63130p; 001:10.17/12.681019, 2006.
3. S R Savithiri Some Acoustic correlates of stuttering A pre-post therapy comparison.
4. Perkins W, Rudas J, Johnson L, Bell J. stuttering :dis-co-ordination of phonation with articulation and respiration, Journal of Speech and hearing disorder; 19: 509-522, 1976.
5. Van Riper C. [The nature of stuttering (2nd Eds.) Prentice-Hall Inc, NJ, 1982.
6. Zebrowski, P.M. "Duration of the speech disfluencies of beginning stutterers", J. Speech-Hear-Res., 34, 3, 1991.
7. Peter Howell and Louise Vause, "Acoustic analysis and perception of stuttered speech", J. Acoustic Soc. Am. 79(5), May 1966.
8. Howell, P., Hamilton, A. & Kyriacopoulos, A., Automatic detection of repetitions and prolongations in stuttered speech. Speech Input/Output: Techniques and Applications, 252-256. IEE Publications, 1986.
9. Howell, P. & Vause, L., Acoustic analysis and perception of vowels in stuttered speech. Journal of the Acoustical Society of America, 79, 1571-1579, 1986.
10. Howell, P., El-Yaniv, N. & Powell, D.J., Factors affecting fluency in stutterers when speaking under altered auditory feedback. In H. Peters and W. Hulstijn (Eds.), Speech Motor Dynamics in Stuttering, 361-369. Springer Press: New York, 1987.
11. Howell, P. & Williams, M., The contribution of the excitatory source to the perception of neutral vowels in stuttered speech. Journal of the Acoustical Society of America, 84, 80-89, 1988.

12. Howell, P. and Wingfield, T., "Perceptual and acoustic evidence for reduced fluency speech in the vicinity of stuttering episodes", Lang. Speech 33, 1, 31-46, 1990.

13. Howell, P. & Williams, M., Acoustic analysis and perception of vowels in children's and teenagers' stuttered speech. Journal of the Acoustical Society of America, 91, 1697-1706, 1992.

14. Howell, P., The acoustic properties of stuttered speech. In C. W. Starkweather and H. F. M. Peters (Eds). Proceedings of the First World Congress on Fluency Disorders, Vol. II (pp. 48-50), Nijmegen, The Netherlands: University Press Nijmegen, 1995.

15. Howell, P., Staveley, A., Sackin, S. and Rustin, L., "Comparison of speech motor development in stutterers and fluent speakers between 7 and 12 years old", J. Fluency Disorders, 20, 1, 243-255, 1995.

16. Howell, P., Au-Yeung, J., Sackin, S., Glenn K., & Rustin, L., Detection of supralexical dysfluencies in a text read by child stutterers. Journal of Fluency Disorders, 22, 299-307, 1997.

17. Howell, P., Sackin, S. & Glenn, K., Development of A Two-Stage Procedure for the Automatic Recognition of Dysfluencies in the Speech of Children Who Stutter: I. Psychometric Procedures Appropriate for Selection of Training Material for Lexical Dysfluency Classifiers. Journal of Speech, Language and Hearing Research, 40, 1073-1084, 1997a.

18. Howell, P., Sackin, S. & Glenn, K. Development of A Two-Stage Procedure for the Automatic Recognition of Dysfluencies in the Speech of Children Who Stutter: II. ANN Recognition of Repetitions and Prolongations with Supplied Word Segment Markers Journal of Speech, Language and Hearing Research, 40, 1085-1096, 1997b.

19. Howell, P., Staveley, A., Sackin, S. and Rustin, L. "Methods of interval selection, presence of noise and their effects on detectability of repetitions and prolongations", J. Acoust. Soc. Amer., 104, 6, 3558-3567, 1998.

20. Howell, P., Au-Yeung, J. and Pilgrim, L., "Utterance rate and linguistic properties as determinants of lexical dysfluencies in children who stutter", J. Acoust. Soc. Amer., 105, 1, 481-490, 1999.

21. Kaczmarek, A. and Skorka, P. "Investigation of the Pitch of the Vocal Tone for Stuttered Speech", 102nd Audio Engineering Society Convention, preprint No. 4486 (L7), Munich, Germany, 1997.

22. Roland-Mieszkowski, M., Czyzewski, A. and Kostek, B. "DSA (Digital Speech Aid) - a new device to decrease or eliminate stuttering. First World Congress of Fluency Disorders", Chapter 11, Munich, 535-539, Munich, Germany, 1995.

23. Kalinowski, J., Armson J. and Stuart A., "Effect of Normal and Fast Articulatory Rates on Stuttering", J. Fluency Disorders, 20, 3, 293-302, 1995.

24. Kostek, B. "Soft Computing in Acoustics, Applications of Neural Networks, Fuzzy Logic and Rough Sets to Musical Acoustics, Studies in Fuzziness and Soft Computing", Physica Verlag, Heilderberg, New York, 1999.

25. Andrzej Czyzewski, Andrzej Kaczmarek Borena Kostek Intelligent Processing of Stuttered Speech, Journal of Intelligent Information Systems, 21:2, 143 –171, 2003.

26. Kuniszyk-Jozkowiak, A. "The statistical analysis of speech envelopes in stutterers and non-stutterers", J. Fluency Disorders, 20, 1, 11-23, 1995

27. Kuniszyk-Jozkowiak, W. "A comparison of speech envelopes of stutterers and nonstutterers", J. Acoust. Soc. Amer., 100, 2, 1105-1110, 1996.

28. Michaelis, D., Frohlich, M. and Strube, H. W., "Selection and combination of acoustic features for the description of pathologic voices", J. Acoust. Soc. Amer., 103, 3, 1628-1639, 1998.

29. Lee B., "Effects of delayed speech feedback", J. Acoust. Soc. Amer., 22, 824-826.
28. Robb, M. and Blomgren, M. (1997). "Analysis of F2 Transitions in the Speech of Stutterers and Non-stutterers", J. Fluency Disorders, 22, 1, 1-16, 1950.

30. Pawlak, Z., "Rough sets", J. Computer and Information Science, 11, 5,341-356, 1982.

31. Polkowski, L. and Skowron A. (Eds.) "Rough Sets in Knowledge Discovery 1: Methodology and Applications", Physica-Verlag, Heidelberg and New York, 1998.

32. Slowinski, R. (ed.) "Intelligent Decision Support. Handbook on Applications and Advances of the Rough Sets Theory". Kluwer Academic Publisher, Dordrecht/ Boston/ London, 1992.

33. Skljarov O.P., "Information model on categories "voice/unvoice" of segments in speech".

34. Skljarov O.P. Neurodynamical Route to chaos and Normal Speech vs. Stuttering//In:2000 Int. Conf. "Control of Oscillations and Chaos, St. Petersburg,, 449-452, 2000.

35. O.P. SKLJAROV, Nonlinear Neurodynamics in Representation of a Rhythm of Speech , Journal of Biological Physics 25: 223–234, 1999.

36. Oleg P. Skljarov, The Control of Dynamic Regimes of the Rhythm within Feigenbaum Scenario of Route to Chaos"

37. Chaos and Speech Rhythm Proceedings of International Joint Conference on Neural Networks, Montreal, Canada, July 31 - August 4, 2005.

38. Oleg Skljarov, Interdisciplinary Research and Speech Rhythm

39. Oleg P. Skljarov', Logistic Mapping as the Means of the Description of the Speech Rhythm in the Vicinity of the Critical Point.

40. Oleg P. Skljarov' Tatyana N. Skljarova, Robust Rhythm as a Conlequence of the Universal Description of Complex System Dyndmics in the Vicinity of Critical Point.

41. Yu Zhang, Jack J. Jiang, and Stephanie M Wallace , Comparison of nonlinear dynamic methods and perturbation methods for voice analysis", J. Acoustic Society America, Vol. 118 (4), October 2005.

42. Lorenzo Matassini, Rainer Hegger, Holger Kantz, Claudia Manfredi, "Analysis of Vocal disorders in a feature space", Medical Engineering & Physics 22, 413-418, 2000.

43. Yu Zhang and J.J. Jiang, "Nonlinear dynamic analysis in signal typing of pathological human voices", Electronic Letters, Vol. 39 No. 13, 26th June 2003.

44. KUMAR. A. and MULLICK, S.K.: 'Attractor dimension, entropy and modelling of speech time series'. Electron. Lett., 26,pp. 1790-1792, 1990.

45. V. Protopopessu and L. M. Hively, Phase-space dissimilarity measures of nonlinear dynamics: Industrial and biomedical applications, Transworld Reaserch Network Recent Res. Devel. Physics, 6: 649-688 ISBN: 81-7895-171-1, 2005.

46. Bandit C, Pomp B, Permutation Entropy: A natural complexity measure for time series Phys. Rev. Lett. 88, 174102, 2002.

47. Yinhe Cao, Wen-wen, Tung, J B Gao, V A Protopopescu, L M Hively (2004), Detecting dynamical changes in time series using the permutation entropy, American Physical Society, 70,046217.

48. Iman Veisi, Naser Pariz, Ali Karimpour, Fast and Robust Detection of Epilepsy in Noisy EEG Signals Using Permutation Entropy, IEEE 200-203, 2007.

49. Joseph G. Sheehan, "Stuttering Behavior: A phonetic analysis", Journal of communication disorders 7, 193-212, 1974

50. Frank J. Falck and patricia sartin lawler, A. Yonovitz, Effect of stuttering on fundamental frequency, J. FLUENCY DISORD. 10, 123-135, 1985.

51. Abdullah I. AL-Shoshan, Speech and music classification and separation: A REVIEW, J. king saud univ., VOL. 19, ENG. SCI. (1), PP. 95-133, RIYADH 1427H./2006.

52. Brown, S., & Moren, A., The frequency of stuttering in relation to word length during oral reading. Journal of Speech Disorders, 7, 153–159, 1942.

53. Brundage & Ratner, 1989; Brundage, S., & Ratner, N. B., Measurement of stuttering frequency in children's speech. Journal of Fluency Disorders, 14, 351–358, 1989.

54. Logan & Conture, 1995; Logan, K., & Conture, E., Length, grammatical complexity, and rate differences in stuttered and fluent conversational utterances of children who stutter. Journal of Fluency Disorders, 20, 36–61, 1995.

55. Silverman, 1972; Silverman, F., Disfluency and word length. Journal of Speech and Hearing Research, 15, 788–791, 1972.

56. Soderberg, 1966; Soderberg, G., The relations of stuttering to word length and word frequency. Journal of Speech and Hearing Research, 9, 584–589, 1966.

57. Wells, 1979; Wells, G., Effect of sentence structure on stuttering. Journal of Fluency Disorders, 4, 123–129, 1979.

58. Wingate, 1967 Wingate, M., Stuttering and word length. Journal of Speech and Hearing Research, 10, 146–152, 1967.

59. Michael Blomgren, Alexander M., Revisting speech rate and utterance length manipulation in stuttering speakers, Journal of Communication Disorders 41, 159-178, 2008.

CHAPTER 5 - RESULTS AND DISCUSSION ON VOCAL DISORDERS

Two groups of samples are used, one from healthy individuals subjects and other with vocal pathologies. Vocal sound signals, അ /ഇ/, and /ഉ/ corresponding to Malayalam alphabets from normal as well as abnormal subjects with vocal disorders are analysed. The recorded audio are converted to digital signals and are subjected to analysis. In this chapter acoustic perturbation methods like, fundamental frequency (F0), jitter, shimmer, zero crossing rate(ZCR) were carried out and non linear measures like maximum lyapunov exponent(λ_{max}), correlation dimension (D_2), Kolmogorov exponent(K_2), and a new measure of entropy viz., Permutation entropy(PE) are evaluated for all two groups of the subjects. Permutation Entropy is a nonlinear complexity measure which can efficiently distinguish regular and complex nature of any signal and extract information about the change in dynamics of the process by indicating sudden change in its value.
The results shows that nonlinear dynamical methods seem to be a suitable technique for voice signal analysis, due to the chaotic component of the human voice. Permutation entropy is well suited due to its sensitivity to uncertainties, since the pathologies are characterized by an increase in the signal complexity and unpredictability. Pathological groups have higher entropy values compared to the normal group. The results clearly distinguish the difference in dynamics between the two cases.

5.1 PATHOLOGICAL SPEECH SIGNAL ANALYSIS

Traditionally spectrographic method which determines the fundamental frequency [1], [2] and formants[3] in the acoustic signal are the most widely used analysis tool by physicians and speech therapist for measurement diagnosis and voice treatment. Eventhough this is the most essential and familiar tool for the clinical community even today. There are other acoustic perturbation measures which are also widely used. This includes Jitter[4],[5] and shimmer[4][5][6] which evaluates variation in fundamental frequency and amplitude and other parameters like Pitch Amplitude(PA) [4], which is the normalized measure of the amplitude of the pitch period peak of the residue signal autocorrelation function, Harmonic to Noise Ratio(HNR)[4] which is used to determine the voice signal noise quantity and coefficient of excess (EX)[4] which is a statistical parameter for noise quantification.

The above mentioned conventional methods are devised based on the linear model assumption which considers voice production as a linear phenomena [4]. Eventhough these methods served the purpose of voice quality assessment to an extend there are also situations[4] where some of these methods could not meet the expected performance. In addition it is now well established that voice production is a complex mechanism involving a number of variables associated to vocal fold biomechanics and aerodynamics of the larynx[4]. With the development of nonlinear dynamical analysis several research teams have investigated the dynamical nature of voice production mechanism. These investigations has proved the low dimensional nonlinear nature of this system[7-11]. These findings have enhanced the importance of nonlinear dynamical methods in the study of normal and pathologically affected voice signals.

A variety of nonlinear dynamic methods has been employed for detecting and assessing both qualitatively and quantitatively various vocal pathologies. Such methods are also used in identifying and assessing other human pathologies [12-14]. Such conventional nonlinear dynamical methods[15-25] are also employed for detecting vocal pathologies as well as assessing the effect of clinical treatments.

Considering the nonlinear dynamical nature of speech production mechanism , voice signal acquired as a time series data can be used to reconstruct the phase space for voice characterisation. Important measures of phase space characteristic are invariant quantities like various dimensions and entropies of attractor formed in the phase space by the voice signal reconstruction. The attractor geometry is capable of providing a deeper understanding of the underlying process. Eventhough these conventional measure of dimension and entropy have been successfully employed for characterizing vocal signals they are found to go wrong in the presence of noise and in the case of short non-stationary data series. Lately, more general measures of entropy which are found to be robust even in the presence of noise are being proposed . The use of entropy is becoming more widely accepted as the concept has been extended from deterministic continuous to stationary random process and discrete dynamical systems. Measures like approximate entropy[26,27] and sample entropy[28-30] are developed for over coming the drawbacks of conventional measure of correlation dimension and kolmogorov entropy. Other measure like discrete entropy[31], relative entropy[4], permutation entropy[32-34] etc., are also proposed with similar aims. Permutation Entropy is a complexity measure suitable for regular, chaotic, noisy or real time based signals. Permutation entropy is efficiently even in the presence of dynamical and/or observational noise. Unlike other nonlinear techniques PE is easier and faster to calculate as the reconstruction of the state space from time series is not required.. In this work we aim at evaluating the diagnostic efficiencies of some of the conventional nonlinear measures like lyapunov exponent, correlation dimension, kolmogrov entropy as well as the newly proposed PE in characterizing vocal pathologies in a real time online setup. For a comparison of the results with traditional acoustic measures some of these conventional linear measures like fundamental frequency, jitter , shimmer and zero crossing rate are also evaluated.

5.2 METHODS

A. Data collection

Ten healthy individuals and ten dysphonic patients served as subjects. All dysphonic subjects were diagnosed with laryngitis, Pharyngitis, vocal cord thickening, laryngitis, tonsillitis with pharangitis, nodule in the vocal fold, paralysis, polyps, cancer and papilloma. Data obtained on each subjects consisted of the microphone signals of recorded vowels /അ/, /ഇ/ and /ഉ/ of Malayalam language at self-selected comfortable pitch and loudness. The audio signals are converted to digital time series by sampling it at a rate of 11Khz.

B. Data Analysis

Nonlinear measures like Correlation dimension, Maximum Lyapanov exponents and Kolmogrov entropy are calculated for each of the signals. Permutation entropy is also evaluated for each of the vocal samples. In addition conventional acoustic analysis is performed on the signals using zero crossing, fundamental frequency, perturbation methods like jitter, shimmer .

5.3 RESULTS AND DISCUSSION:

Tables 5.1a & 5.1b, 5.2a & 5.2b, 5.3a & 5.3b summarises the results of the acoustic analysis evaluated for the recorded speech signal /ആ/, /ഇ/ and /ഉ / for ten normal subjects and ten abnormal subjects. Table 5.1a and 5.1b shows the acoustic parameters of normal and pathological subjects respectively for vowel /ആ/ . Similarily Table 5.2a,b and 5.3a,b summarises the acoustic parameters for vowels /e/ and /u/ .

Normal	**Vowel "ആ"**				
	Time (ms)	ZCross	F0(Hz)	%jit	%shm
SUBJECT 11	300	27	845.6	4.99	32.57
SUBJECT 12	310	25	911.4	4.65	27.49
SUBJECT 13	410	9	735.4	6.17	18.58
SUBJECT 14	340	22	819.1	5.3	44.48
SUBJECT 15	300	20	742	1.78	11.38
SUBJECT 16	280	22	760.1	2.93	22.71
SUBJECT 17	340	37	880.2	3.32	26.1
SUBJECT 18	340	23	770.2	6.98	56.62
SUBJECT 19	300	34	854	2.99	16.33
SUBJECT 20	340	36	812	3.73	45.65
Average		25.5 ± 8.5	813	4.28 ± 1.61	30.19 ± 14.56

Table 5.1a Linear acoustic parameters evaluated for vowel /ആ / for normal subjects.

Vocal Disorder	**Vowel "ആ"**				
	Time (ms)	ZCross	F0(Hz)	%jit	%shm
SUBJECT21	290	10	738.4	3.25	22.27
SUBJECT22	270	2	635.3	11.61	21.48
SUBJECT23	310	6	954.7	3.88	25.06
SUBJECT24	310	24	822.3	10.96	51.22
SUBJECT25	300	1	595.1	5.41	33.03
SUBJECT26	280	5	557	11.9	28.11
SUBJECT27	410	6	822.2	13.92	36.63
SUBJECT28	320	16	724.5	20.68	33.12
SUBJECT29	330	15	872.9	6.34	31.88
SUBJECT30	390	7	697.9	39.75	40.99
Average		9.2± 7.16	743	12.77 ± 10.86	32.3 ± 9.07

Table 5.1b Linear acoustic parameters evaluated for vowel /ആ/ for abnormal subjects.

Normal	Vowel "ஐ"				
	Time (ms)	ZCross	F0(Hz)	%jit	%shm
SUBJECT11	290	19	383.9	4.49	14.31
SUBJECT12	270	22	331.8	3.35	14.91
SUBJECT13	330	4	277.1	3.86	19.42
SUBJECT14	310	11	361.2	4.45	18.35
SUBJECT15	300	11	286.1	2.28	11.64
SUBJECT16	280	9	351	1.93	12.31
SUBJECT17	410	13	373.3	5.21	16.95
SUBJECT18	320	21	281.9	11.95	47.17
SUBJECT19	330	12	353	6.62	37.6
SUBJECT20	390	17	313.2	5.25	45.45
Average		13.9 ± 5.72	331.25	4.94 ± 2.83	23.81 ± 13.94

Table 5.2a Linear acoustic parameters evaluated for vowel /ஐ/ for normal subjects.

Vocal Disorder	Vowel "ஐ"				
	Time (ms)	ZCross	F0(Hz)	%jit	%shm
SUBJECT21	290	19	383.9	8.34	21.02
SUBJECT22	270	22	331.8	4.98	41.4
SUBJECT23	310	4	277.1	4.29	28.18
SUBJECT24	310	11	361.2	4.35	27.24
SUBJECT25	300	11	286.1	6.64	21.32
SUBJECT26	280	9	351	6.42	29.46
SUBJECT27	410	13	373.3	3.65	23.8
SUBJECT28	320	19	337.2	8.13	28.46
SUBJECT29	330	4	291.6	21.41	42.8
SUBJECT30	390	9	392.2	9.42	33.55
Average		12.1 ± 6.2	338.54	7.76 ± 5.17	29.72 ± 7.55

Table 5.2b Linear acoustic parameters evaluated for vowel /ஐ/ for abnormal subjects.

Normal	Vowel "ഉ"				
	Time (ms)	ZCross	F0(Hz)	%jit	%shm
SUBJECT11	310	12	381.2	6.17	25.59
SUBJECT12	330	9	351.1	11.06	52.26
SUBJECT13	310	4	406.3	5.44	17.5
SUBJECT14	260	12	375.7	3.87	20.24
SUBJECT15	400	6	356	24.63	31.81
SUBJECT16	250	13	399.7	2.44	11.95
SUBJECT17	410	13	450	3.72	18.51
SUBJECT18	260	13	383.8	4.07	23.87
SUBJECT19	280	13	383.8	3.28	20.31
SUBJECT20	350	12	363.1	3.63	18.45
Average		10.7± 3.26	385	6.83 ± 6.71	24.05± 11.24

Table 5.3a Linear acoustic parameters evaluated for vowel /ഉ/ for normal subjects..

Vocal Disorder	Vowel "ഉ"				
	Time (ms)	ZCross	F0(Hz)	%jit	%shm
SUBJECT21	460	5	450.2	14.61	21.08
SUBJECT22	370	1	299.1	36.38	31.76
SUBJECT23	540	2	360.4	22.46	26.14
SUBJECT24	500	9	540	6.84	19.61
SUBJECT25	510	3	443.8	3.53	18.47
SUBJECT26	360	4	295.1	8.35	36.99
SUBJECT27	750	10	364.4	12.15	36.04
SUBJECT28	390	6	434	7.89	21.41
SUBJECT29	380	2	286.7	11.5	30.34
SUBJECT30	350	6	515	8.76	51.87
Average		4.8± 3.01	398	13.25± 9.62	29.37 ± 10.39

Table 5.3b Linear acoustic parameters evaluated for vowel /ഉ/ for abnormal subjects.

The fundamental frequency F0 from Table 5.1a and b shows that the average value of F0 of vowel /അ/ is lower for the pathological subjects compared to the normal subjects. Average F0 for normal subject group is found to be 813Hz whereas for the pathological group it is found to be 743Hz. From Table 5.2a and b it can be inferred that average F0 of vowel /ഈ/ is similar in both groups. In the case of vowel /ഉ/ the mean value of F0 is higher for the pathological subject group. Table 5.3a and b shows value of F0 for normal subject group is 385Hz whereas it is 398Hz for pathological group.

The higher value of F0 for all the vowels for pathological groups compared to the normal groups can be attributed as the healthy group have a more defined period[22]. It can also be due to the variability of the parameters for the dysphonic groups due physiological difference between the pathologies occurring and hence the alterations in system dynamics can lead to much different voice signals [22]

Percentage jitter for vowel /അ/ and /ഉ/ is higher for the pathological group compared to the normal group. Table 5.1a, 5.1b and Table 5.3a, 5.3b shows that the value of percentage of jitter for pathological group is 12 and 13 whereas for the normal group it is 4 and 6. Table 5.2a and 5.2b shows that the percentage Jitter for vowel /ഈ/ is also higher for the pathological group but, the difference is not very prominent . Even in the case of percentage shimmer, there is increase for the pathological group however, the difference between the two groups is not very wide.

Results of analysis using Zero crossing rate done on the above signals were found to be in accordance with the percentage jitter. Zero crossing rate of vowel /അ/ and /ഉ/ showed significant lowering in the pathological groups compared to that of normal group. In the case of vowel /ഈ/ also there is a lowering eventhough it is not as observed in the case of other two vowels. The overall value of zero crossing rate calculated for the whole signal is higher for normal subjects compared to the abnormal subjects for all the vowels /അ/ , /ഈ/ and /ഉ/. The number of times the signal crossing the base line is more in the case of normal as the signal appears more periodic whereas the abnormal signals are aperiodic and hence the lower value of zero crossing. The overall value of the jitter i.e., cycle-to-cycle change of frequencies values is lower for the normal subjects compared to the abnormal subjects but there is not much change for the shimmer value (cycle-to-cycle change in amplitudes) of normal and abnormal subjects. These results do not clearly differentiate the normal and pathological subjects. Hence we go in for nonlinear analysis for characterizing the healthy pathological subjects.

Tables 5.4a, 5.4b, 5.5a, 5.5b, 5.6a, 5.6b summarises the results of the nonlinear analysis of recorded speech signal /ആ/, /ഈ/ and /ഉ/ for ten normal subjects and ten pathological subjects. The first, second and third coloumns shows maximum lyapnov exponent(λ_{max}), correlation dimension(D_2) , Kolmogrov entropy(K_2) . The data length of the signals varies between 7614 and 31701 for vowel /ആ/, 7050 and 28049 for vowel /ഈ/ and between 7896 and 33442 for vowel /ഉ/ .

Normal	**Vowel "ആ"**			
	Lyap_spec (λ_{max})	D_2	K_2	Total PE
SUBJECT 11	0.069	2.46	0.03	0.72
SUBJECT 12	0.072	2.34	0.02	0.70
SUBJECT 13	0.052	2.55	0.09	0.71
SUBJECT 14	0.059	2.34	0.03	0.66
SUBJECT 15	0.052	2.26	0.02	0.60
SUBJECT 16	0.05	2.19	0.01	0.66
SUBJECT 17	0.072	2.33	0.11	0.72
SUBJECT 18	0.051	2.38	0.02	0.67
SUBJECT 19	0.034	2.05	0.02	0.61
SUBJECT 20	0.108	1.55	0.03	0.65
Average	0.0619	2.25	0.04	0.67

Table 5.4a Nonlinear parameters evaluated for vowel /ആ/ for normal subjects

Vocal Disorder	Vowel "ഒ"			
	Lyap_spec (λ_{max})	D_2	K_2	Total PE
SUBJECT 21	0.223	2.12	0.029	0.69
SUBJECT 22	0.124	2.72	0.018	0.72
SUBJECT 23	0.057	2.54	0.045	0.59
SUBJECT 24	0.061	2.04	0.022	0.90
SUBJECT 25	1.251	5.91	0.023	0.73
SUBJECT 26	0.048	1.51	0.032	0.78
SUBJECT 27	0.06	3.03	0.028	0.78
SUBJECT 28	0.078	4.12	0.026	0.88
SUBJECT 29	0.069	3.24	0.023	0.78
SUBJECT 30	0.255	4.11	0.009	0.77
Average	0.223	3.13	0.025	0.76

Table 5.4b Nonlinear parameters evaluated for vowel /ഒ/ for abnormal subjects

Normal	Vowel "ഉ"			
	Lyap_spec (λ_{max})	D_2	K_2	Total PE
SUBJECT 11	0.179	1.43	0.02	0.73
SUBJECT 12	0.026	1.22	0.017	0.76
SUBJECT 13	0.038	1.54	0.094	0.79
SUBJECT 14	0.033	1.57	0.034	0.80
SUBJECT 15	0.017	2.65	0.025	0.68
SUBJECT 16	0.025	2.45	0.094	0.61
SUBJECT 17	0.014	1.93	0.023	0.84
SUBJECT 18	0.003	1.33	0.038	0.77
SUBJECT 19	0.086	2.05	0.027	0.77
SUBJECT 20	0.03	2.54	0.065	0.82
Average	0.451	1.87	0.032	0.76

Table 5.5a Nonlinear parameters evaluated for vowel /ഉ/ for normal subjects

Vocal Disorder	Vowel "ഈ"			
	Lyap_spec (λ_{max})	D_2	K_2	Total PE
SUBJECT 21	0.074	2.84	0.03	0.83
SUBJECT 22	0.987	4.19	0.012	0.77
SUBJECT 23	0.962	3.05	0.026	0.72
SUBJECT 24	0.946	1.31	0.017	0.94
SUBJECT 25	0.084	2.71	0.015	0.63
SUBJECT 26	0.049	4.89	0.015	0.88
SUBJECT 27	0.368	1.83	0.028	0.72
SUBJECT 28	0.782	2.66	0.002	0.94
SUBJECT 29	0.057	2.37	0.007	0.90
SUBJECT 30	0.015	2.89	0.005	0.77
Average	0.432	2.87	0.016	0.81

Table 5.5b Nonlinear parameters evaluated for vowel /ഈ/ for abnormal subjects

Normal	**Vowel "ഉ"**			
	Lyap_spec (λ)	D_2	K_2	Total PE
Subject 11	0.034	2.46	0.032	0.50
SUBJECT 12	0.054	2.48	0.024	0.47
SUBJECT 13	0.023	2.76	0.074	0.46
SUBJECT 14	0.051	2.34	0.045	0.46
SUBJECT 15	0.023	2.24	0.058	0.38
SUBJECT 16	0.027	2.35	0.036	0.44
SUBJECT 17	0.016	2.42	0.030	0.50
SUBJECT 18	0.032	2.42	0.020	0.56
SUBJECT 19	0.044	2.26	0.006	0.44
SUBJECT 20	0.033	2.35	0.033	0.50
Average	0.0337	2.41	0.02	0.47

Table 5.6a Nonlinear parameters evaluated for vowel /ഉ/ for normal subjects

Vocal Disorder	Vowel "ഉ"			
	Lyap_spec (λ)	D_2	K_2	Total PE
SUBJECT 21	0.098	3.05	0.026	0.73
SUBJECT 22	0.052	2.59	0.029	0.73
SUBJECT 23	0.081	2.48	0.015	0.67
SUBJECT 24	0.014	0.82	0.022	0.90
SUBJECT 25	1.074	3.12	0.016	0.59
SUBJECT 26	0.252	4.08	0.015	0.86
SUBJECT 27	0.074	1.42	0.022	0.67
SUBJECT 28	0.252	3.01	0.002	0.91
SUBJECT 29	0.075	2.78	0.007	0.80
SUBJECT 30	0.024	3.13	0.005	0.77
Average	0.20	2.65	0.016	0.76

Table 5.6b Nonlinear parameters evaluated for vowel /ഉ/
for abnormal subjects

For vowel /അ/ the mean value of maximum lyapunov exponent of the pathological subjects shows a higher value compared to the normal subjects indicating increased irregularity in the vocal signals of these subjects. Similar increase is found in the case of correlation dimension , kolmogrov entropy , mean PE , single PE. For vowel /ഈ / and /ഉ/ also the nonlinear parameters are found to be higher for the pathological group. The increase in these nonlinear quatitative measures points towards the increased nonlinearity induced by the vocal pathology [24].

In the case of vowel /ഉ/ the results of the nonlinear measures gives contradictory implication. Even though the mean value of lyapanov exponent and correlation dimension shows slight increase in their values for the pathological subject group, the kolmogrov entropy shows a reduction in value for the pathological subject group compared the normal group. The Total PE also indicates increased irregularity in the case of pathological group compared to that of the normal. The contradictory results can be attributed to the effect of noise contamination and short data length of the signals. The sensitiveness of conventional parameters like correlation dimension and kolmogorov entropy are well established and proved in the case of various biological signals. PE is a robust measure suitable for short noisy real time signals [32] its effectiveness is established by various groups [12-14].The results of our analysis also shows that PE is robust to noisy and nonstationarity and is efficient in characterizing pathological signals from raw data. These results reinforces the effectiveness of PE in characterizing the dynamical difference between normal and pathological subject groups even in the presence of noise and with small data sets.

For visual representation of the difference of dynamics between the vocal signals of the of normal and pathological groups the evolution of PE of moving windows of 512 samples is shown in 5.1a, 5.1b , 5.2a, 5.2b and 5.3a, 5.3b . Fig.5.1a and b, 5.2a and b and 5.3a and b represents the variation of PE for letters “അ”,“ഇ”, “ഉ” respectively. Normal subjects are represented by green, abnormal subjects are represented by red. Results of the analysis clearly indicate that PE values of normal subjects are lower than that of abnormal cases.

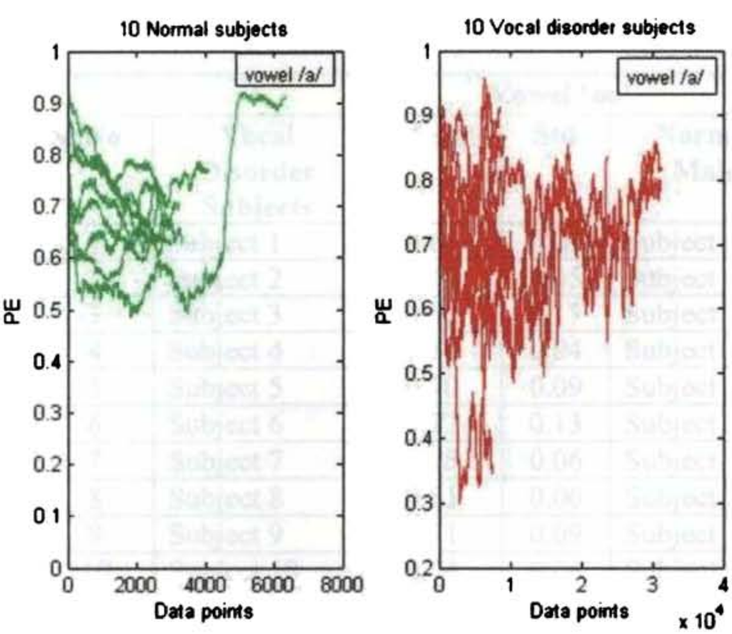


Fig. 5.1 a Fig. 5.1 b

Figure 5.1: Variation of PE for vowel /അ/ for normal and pathological subjects

Figure. 5.1a shows variation of PE of speech signals of ten Normal subjects and Figure. 5.1b show that of Vocal disorder subjects for the audio signal recorded while uttering the vowels /ആ/. PE of order 5 is calculated for sliding window of length 512 samples for these signals. The variation in PE values of normal speakers for reading the vowel /ആ/ lie in range of 0.6 – 0.9 on an average whereas overall range of PE value for that of vocal disorder subjects range between 0.7 – 0.9. There is general increase of complexity in the signals of pathological subjects compared to that of the fluent speakers. This clearly brings out the increased irregularity of pathological speech. The result confirms that the pathological subject dynamics exhibits increased complexity when compared to that of fluent speech signals [45]. Table 5.7 shows the mean and standard deviation(SD) of PE shown in figure. 5.1a and 5.1b. and mean of subjects(1 to 10) ranging from 0.50 to 0.86 and that of normal subjects(11 to 20) ranging from 0.02 to 0.78. The average PE of the pathological subjects is 0.71 and that of the normal subjects is 0.53.. The mean value of PE is higher for pathological subjects compared to the fluent subjects. The larger standard deviation of subjects in the pathological group indicates the wider dynamical variation in this group compared to that of normals.

Vowel 'ആ'						
Sl.No	**Vocal Disorder Subjects**	**Mean**	**Std**	**Normal Male**	**Mean**	**Std**
1	Subject 1	0.65	0.09	Subject 11	0.62	0.04
2	Subject 2	0.69	0.05	Subject 12	0.68	0.03
3	Subject 3	0.50	0.17	Subject 13	0.02	0.02
4	Subject 4	0.86	0.04	Subject 14	0.73	0.05
5	Subject 5	0.70	0.09	Subject 15	0.78	0.04
6	Subject 6	0.72	0.13	Subject 16	0.74	0.05
7	Subject 7	0.75	0.06	Subject 17	0.66	0.03
8	Subject 8	0.81	0.06	Subject 18	0.15	0.06
9	Subject 9	0.71	0.09	Subject 19	0.15	0.03
10	Subject 10	0.74	0.06	Subject 20	0.73	0.04
	Average	0.71	0.08	**Average**	0.53	0.04

Table 5.7 shows the mean and standard deviation of PE for vowel /ആ/

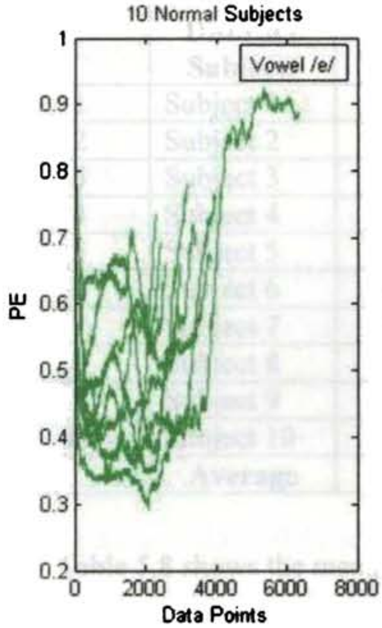


Fig. 5.2 a

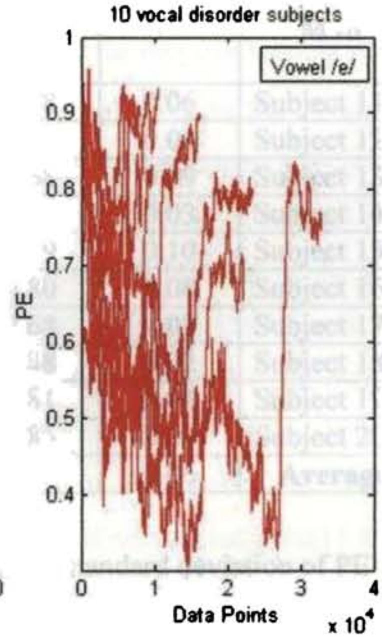


Fig. 5.2 b

Figure 5.2: Variation of PE for vowel /go/ for normal and pathological subjects

Figure. 5.2a shows variation of PE of speech signals of 10 Normal subjects and Figure. 5.2b show that of Vocal disorder subjects for the audio signal recorded while uttering the vowels /go/. PE of order 5 is calculated for non overlapping window of length 512 samples for these signals. The variation in PE values of normal speakers for the vowel /go/ lie in range of 0.4 – 0.8 on an average whereas overall range of PE value for that of vocal disorder subjects range between 0.6 – 0.9. There is general increase of complexity in the signals of pathological subjects compared to that of the fluent speakers. The result confirms that the Pathological subject dynamics exhibits increased complexity when compared to that of fluent speech signals [45]. Table 5.8 shows the mean and standard deviation(SD) of PE shown in figure. 6.5a and 6.5b. and mean of subjects(1 to 10) ranging from 0.59 to 0.88 and that of normal subjects(11 to 20) ranging from 0.58 to 0.79. The average mean PE of the pathological subjects is 0.77 and that of the normal subjects is 0.67.

Vowel 'ഉ'						
Sl.No	Vocal Disorder Subjects	Mean	Std	Normal Male	Mean	Std
1	Subject 1	0.78	0.06	Subject 11	0.85	0.05
2	Subject 2	0.73	0.06	Subject 12	0.83	0.04
3	Subject 3	0.68	0.09	Subject 13	0.78	0.03
4	Subject 4	0.88	0.03	Subject 14	0.71	0.11
5	Subject 5	0.59	0.10	Subject 15	0.89	0.04
6	Subject 6	0.80	0.08	Subject 16	0.86	0.03
7	Subject 7	0.68	0.09	Subject 17	0.82	0.03
8	Subject 8	0.88	0.02	Subject 18	0.70	0.02
9	Subject 9	0.81	0.04	Subject 19	0.60	0.03
10	Subject 10	0.85	0.03	Subject 20	0.50	0.03
	Average	0.77	0.06	**Average**	0.75	0.04

Table 5.8 shows the mean and standard deviation of PE for vowel /ഉ/

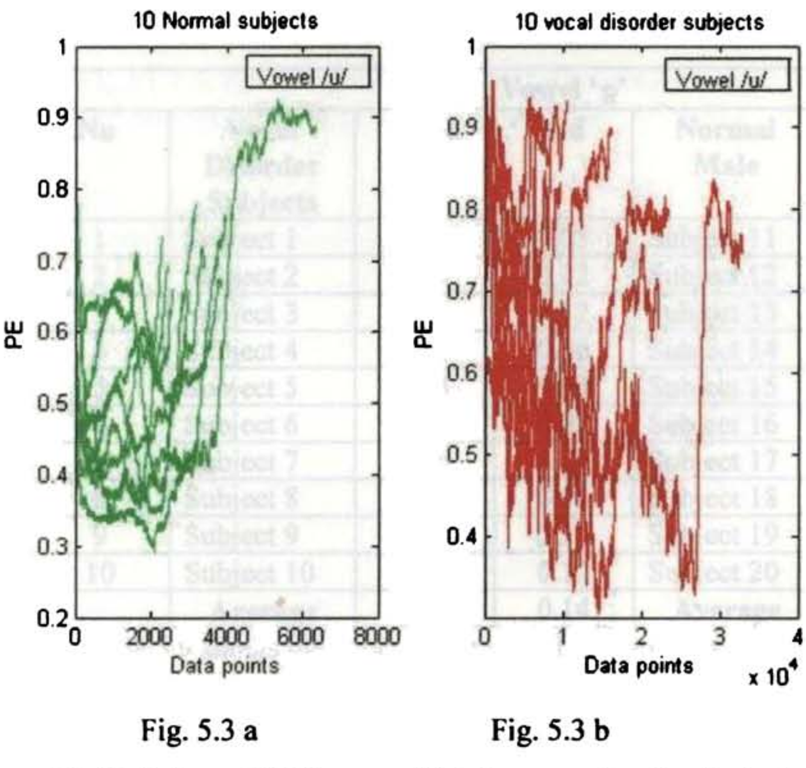


Fig. 5.3 a Fig. 5.3 b

Figure 5.3: Variation of PE for vowel /ഉ/ for normal and pathological subjects

Figure. 5.3a shows variation of PE of speech signals of 10 Normal subjects and Figure. 5.3b show that of Vocal disorder subjects for the audio signal recorded while uttering the vowels /ഉ/. PE of order 5 is calculated for non overlapping window of length 512 samples for these signals. The variation in PE values of normal speakers for reading the vowel /ഇ/ lie in range of 0.4 – 0.8 on an average whereas overall range of PE value for that of vocal disorder subjects range between 0.6 – 0.9. There is general increase of complexity in the signals of pathological subjects compared to that of the fluent speakers. This clearly brings out the increased irregularity of pathological speech. The result confirms that the Pathological subject dynamics exhibits increased complexity when compared to that of fluent speech signals [45]. Table 5.8a and 5.8b shows the mean and standard deviation(SD) of PE shown in figure. 5.6a and 5.6b. and mean of subjects(1 to 10) ranging from 0.55 to 0.86 and that of normal subjects(11 to 20) ranging from 0.41 to 0.63. The average mean PE of the pathological subjects is 0.69 and that of the normal subjects is 0.52.

Vowel 'ഉ'						
Sl.No	**Vocal Disorder Subjects**	**Mean**	**Std**	**Normal Male**	**Mean**	**Std**
1	Subject 1	0.46	0.05	Subject 11	0.46	0.12
2	Subject 2	0.47	0.12	Subject 12	0.40	0.12
3	Subject 3	0.41	0.12	Subject 13	0.47	0.06
4	Subject 4	0.48	0.06	Subject 14	-.63	0.03
5	Subject 5	0.54	0.04	Subject 15	0.55	0.09
6	Subject 6	0.63	0.03	Subject 16	0.47	0.08
7	Subject 7	0.08	0.55	Subject 17	0.68	0.09
8	Subject 8	0.55	0.09	Subject 18	0.61	0.14
9	Subject 9	0.59	0.21	Subject 19	0.81	0.16
10	Subject 10	0.53	0.12	Subject 20	0.55	0.05
	Average	0.47	0.14	**Average**	0.44	0.09

Table 5.9 shows the mean and standard deviation of PE for vowel /ഉ/

Figure 5.4, 5.5 and 5.6 shows PE of each of the ten subjects in both group of abnormal and normal subjects reading vowel /അ/, /ഇ/ and /ഉ/. PE is calculated for three vowels for each of the 20 subjects 10 abnormal and 10 normal by taking the whole time series as window size and order 5. Overall PE values are higher for abnormal subjects compared to the normal subjects which again confirms that the PE of normal subjects is lower than the abnormal subjects.

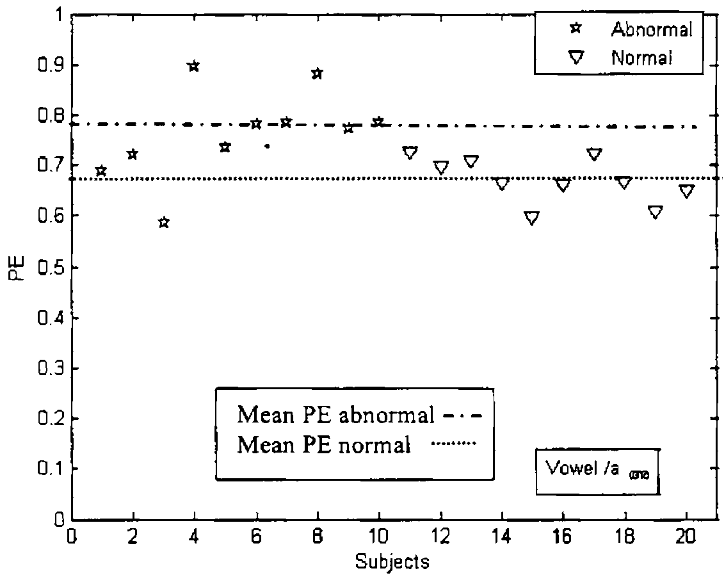


Figure 5.4. PE calculated for first 10 Abnormal subjects(1-10) and normal subjects(11-20) for Vowel /ആ/

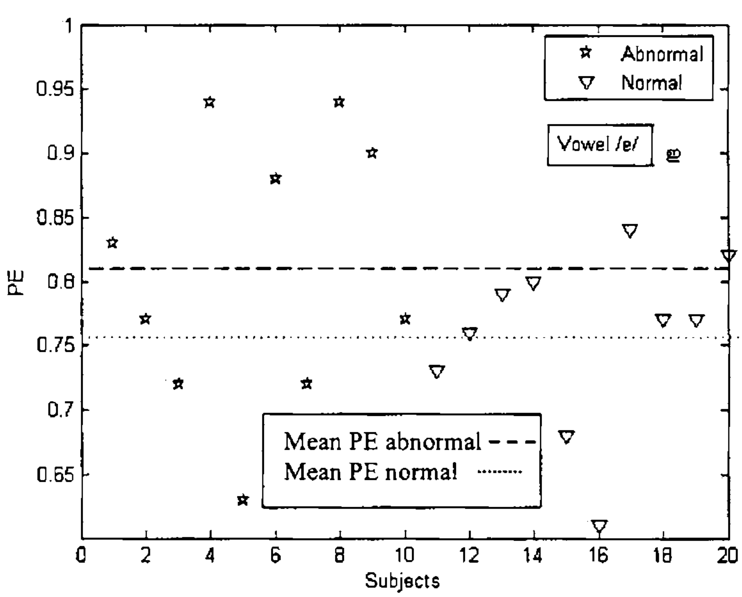


Figure 5.5. PE calculated for first 10 Abnormal subjects(1-10) and normal subjects(11-20) for Vowel /ഇ/

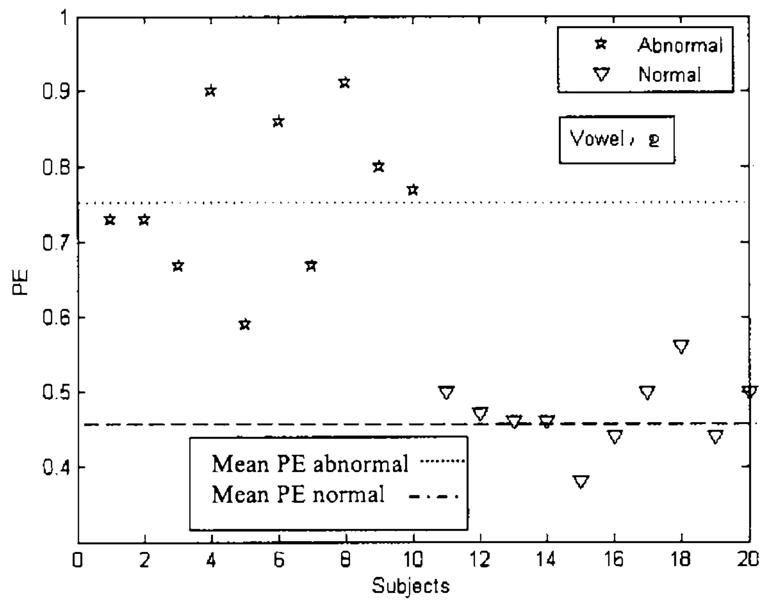


Figure 5.6. PE calculated for first 10 Abnormal subjects(1-10) and normal subjects(11-20) for Vowel /ഉ/

The results of the analysis indicate that irrespective of the gender as well as pathological condition PE values are higher for abnormal cases. There is increased irregularity in pathological group compared to the normal group. The higher values of PE of pathological subjects indicate that with abnormalities in voice signals, irregularities in speech signals increases. This reinforces the concept of presence of bifurcations leading to chaos in signals of vocal disorders.

From these results it is clear that PE is effective in characterizing the amount of disorder in the vocal pathologies. This characteristic of PE can be made useful in the preliminary investigation of vocal disorders in clinical applications. This can be used as a tool by the clinicians during follow-ups for identifying and evaluating the effect of any treatment given to the patients. At every level of treatment the PE data can be stored in the data bank of the patients and can be compared to the PE data before starting the treatment. This may also give first hand information in diagnosis about the level of disorder in pre and post treatment conditions. Hence it can be concluded that PE can be used as a real time indicator in deciding the final strategy of treatment in vocal pathological cases.

5.4 REFERENCES:

[1] Frank J. falck and Patricia Sartin Lawler A. Yonovitz, Effect of stuttering on fundamental frequency, J. FLUENCY DISORD. 10, 123-135, 1985.

[2] Paavo Alku, Juha Vintturi, Erkki Vilkman, Measuring the effect of fundamental frequency raising as a strategy for increasing vocal intensity in soft, normal and loud phonation, ELSEVIER, Speech Communication 38, 321–334, 2002.

[3] Matti Niemi, Juba-Pertti Laaksonen, Timo Peltomaki, Jukka Kurimo, Olli Aaltonen, and Risto-Pekka Happonent, Acoustic Comparison of vowel sounds produced before and after Orthognathic surgery for Mandibular advancement, American Asociation of oral and Maxillofacial surgeons , J Oral Maxillofac surg 64:910-916, 2006.

[4] Paulo Roge´rio Scalassara , Marı´a Eugenia Dajer, Carlos Dias Maciel, Rodrigo Capobianco Guido, Jose´ Carlos Pereira, Relative entropy measures applied to healthy and pathological voice characterization, Applied Mathematics and Computation 207, 95–108, ELSEVIER, 2009.

[5] PARLEY W. NEWMAN, RICHARD W. HARRIS, AND LAURENCE M. HILTON, Vocal jitter and shimmer in stuttering, J. FLUENCY DISORD. 14, 87-95, 1989.

[6] Haydee F. Wertzner, solange Schreiber, Luciana Amaro, Analysis of fundamental frequency. Jitter, shimmer and vocal intensity in children with phonological disorders, Rev Bras Otorrinolaringol, V.71, n.5, 582-88, Brazilian Journal of Otorhinolaryngology 71(5) Part 1, Sep-Oct., 2005.

[7] S. McLaughlin and A. Lowry, Nonlinear dynamical systems concepts in speech analysis, in EUROSPEECH '93, pp. 377–380, 1993.

[8] H. F. V. Boshoff and M. Grotepass, The fractal dimension of fricative speech sounds, in COSMIG '91, pp. 12–16, IEEE, 1991.

[9] S. S. Narayanan and A. A. Alwan, A nonlinear dynamical system analysis of fricative consonants, The Journal of the Acoustical Society of America, vol. 97,pp. 2511–2524, April 1995.

[10] M. Casdagli, Chaos and deterministic versus stochastic non-linear modelling, Journal of the Royal Statistical Society B, vol. 54, no. 2, pp. 303–328, 1991.

[11] H. M. Teager and S. M. Teager, Evidence for nonlinear sound production mechanisms in the vocal tract, in Proc NATO ASI on Speech Production and Speech Modelling, pp. 241–261, 1990.

[12] C J Stam , Nonlinear dynamical analysis of Eeg and MEG: Review of an emerging field , Clinical Neuro phy 116, page 2266-2301, 2005.

[13] Jian Li l and Xingbao Ning Dynamics complexity detection in short-term physiological series using base-scale entropy, Phy Rev E., 73, 052902, 2006.

[14] R A Thuraisingham and G A Gottwald On multi scale entropy analysis for physiological data" Phy. A. 366, page 323-332, 2006.

[15] Yu Zhang and Clancy McGilligan, Liang Zhou, Mark Vig and Jack J. Jiang, 2004 Acoustical Society of America. @DOI: 10.1121/1.1699392, Pages: 2270–2277, Nonlinear dynamic analysis of voices before and after surgical excision of vocal polyps, J. Acoust. Soc. Am., Vol. 115, No. 5, Pt. 1, May 2004.

[16] I. R. Titze, R. Baken, and H. Herzel, Evidence of chaos in vocal fold vibration, in Vocal Fold Physiology: New Frontier in Basic Science, edited by I. R. Titze (Singular), San Diego, pp. 143–188, 1993.

[17] H. Herzel, D. Berry, I. R. Titze, and M. Saleh, ''Analysis of vocal disorders with methods from nonlinear dynamics,'' J. Speech Hear. Res. 37,1008–1019, 1994.

[18] Kakita and H. Okamoto, Visualizing the characteristics of vocal fluctuation from the viewpoint of chaos: an attempt toward qualitative quantification, in Vocal Fold Physiology—Voice Quality Control, edited by O. Fujimura and M. Hirano (Singular), San Diego, pp. 235–248, 1995.

[19] A. Behrman and R. J. Baken, Correlation dimension of electroglottographic data from healthy and pathologic subjects, J. Acoust. Soc. Am. 102, 2371–2379, 1997.

[20] I. Hertrich, W. Lutzenberger, S. Spieker, and H. Ackermann, Fractal dimension of sustained vowel productions in neurological dysphonias: an acoustic and electroglottographic analysis, J. Acoust. Soc. Am. 102, 652–654, 1997.

[21] A. Giovanni, M. Ouaknine, and J-M Triglia, Determination of largest Lyapunov exponents of vocal signal: application to unilateral laryngeal paralysis'' J. Voice 13, 341, 1998.

[22] J. J. Jiang and Y. Zhang, Nonlinear dynamic analysis of speech from pathologic subjects,'' Electron. Lett. 38, 294–295, 2002.

[23] Y. Zhang and J. J. Jiang, Nonlinear dynamic analysis of signal typing of pathological human voices,'' Electron. Lett. 39, 1021–1023, 2003.

[24] Yu Zhang,a_ Jack J. Jiang,b_ and Stephanie M. Wallace, Comparison of nonlinear dynamic methods and perturbation methods for voice analysis, J. Acoust. Soc. Am. 118 (4), Pages: 2551–2560, October 2005.

[25] L. Mattassini, R. Hegger, H. Kanz, C. Manfredi Analysis of vocal disorders in feature space, Medical Engineering & Physics, 22, 423-418, 2000.

[26]. S.M.Pincus, Approximate entropy as a measure of system complexity, in Proceedings of .Nat.Acad.Sci.USA, vol . 88 pp.2297-2301, (1991).

[27] S.M.Pincus, Approximate entropy (ApEn) as complexity measure , Chaos,1, vol 5 110-117, 1995.

[28] S.M.Pincus , I.M.Gladstone, and R.A.Ehrenkranz, A regularity statistic for medical data analysis,, J. Clin. Monit ,1, vol.7 pp. 335-345, 1991.

[29] J.S Richman and J.R Moorman, Physiological time series analysis using approximate entropy and sample entropy, American journal of Physiology, Heart Circ.Physiol.,vol .278, pp.H2039-2049, 2000.

[30] M. Costa, A.L. Goldberger, and C. –K. Peng, Multiscale entropy analysis of complex physiological time series, Physical Review Letters, 6 (2002) vol. 88 pp 068- 102, M. Costa, A.L. Goldberger, C.K. Peng, Phys. Rev. E 71, 021906, 2005.

[31] J.M. Amigo, L. Kocarev, I. Tomovski, Discrete entropy, Phys. D: Nonlinear Phenom., 228 (1), 77–85, 2008.

[32] Bandt C, Pomp B, Permutation Entropy: A natural complexity measure for time series, Physical Review Letters, 17, 88 174102 1-4, 2002.

[33] Y. Cao, W. Tung, J.B. Gao , V.A. Protopopescu, M. Hively, Detecting dynamical changes in time series using the permutation entropy, Physical Review E, 2004.

[34] I. Veisi, N. Pariz, A. Karimpour , Fast and Robust Detection of Epilepsy in Noisy EEG Signals Using Permutation Entropy, IEEE conference, 1, 4244-1509, 2007.

CHAPTER 6 - SUMMARY, CONCLUSIONS AND FUTURE DIRECTIONS

Medical fields require fast, simple and noninvasive methods of diagnostic techniques. Several methods are available because of the growth of technology. To understand the complexity of voice signals, new methods of analysis have been developed, such as nonlinear dynamics aiming at the exploration of voice signals dynamic nature. The work presented in this thesis focuses on characterization of healthy and pathological voice signals with the aid of modern nonlinear methods like maximum Lyapunov exponent, correlation dimension, kolmogorov entropy and a fast and robust entropy measure Permutation Entropy (PE). Nonlinear dynamical methods seem to be a suitable technique for voice signal analysis. Among the nonlinear parameters investigated lyapunov exponent and PE show relatively better performance in characterizing the difference in dynamics between normal and abnormal speech signal.

The results of acoustic as well as nonlinear analysis on individual vowels and continuous speech data indicate the efficiency of nonlinear measures in characterizing the change in dynamics of stuttered speech after a specific period of speech therapy. Conventional nonlinear measures like correlation dimension and kolmogorov entropy do not perform very well on raw data indicating dynamical changes. In the case of post-treatment of vowels and continuous speech data, the conventional nonlinear measures like Correlation dimension (D_2) and Kolmogorov entropy (K_2) do not perform very well. These measures being sensitive to non-stationarity and noise contamination produce ambiguous results mainly due to the presence of silent blocks in trained speech of stuttering subjects. Among the conventional nonlinear measures maximum lyapunov exponent performs slightly better. The newly introduced entropy measure mainly permutation entropy shows efficiency in indicating the improvement in speech performance after the therapy. The robustness of this measure to noise contamination and non-stationarity helps this measure in effectively characterizing the level of improvement in speech process after therapy. PE being a bounded measure is more efficient in comparing different methods of therapy.

The specific ability of nonlinear parameters in characterizing the difference in dynamics is evident in the results of analysis on vocal pathologies also. Among the nonlinear parameters max lyapunov exp and PE are better indicators of the increased dynamical complexity of pathological voice signals. The simplicity and speed of computation of PE enhances its suitability for application in online evaluation setup. Robustness of this measure towards noise and non-stationarity add to its efficiency in characterizing the dynamical difference between normal and pathological, fluent and disfluent speech signals.

The versatile and invariant properties of Permutation Entropy (**PE**) helped us to meet our aims listed below

1. To compare and analyse voice signals using traditional perturbation and nonlinear dynamics time series methods for stutterers and vocal disorder subjects .
2. To Characterise stutterred signals from fluent signals.
3. To estimate the efficacy of speech therapy rendered to stuttered subjects.
4. To Characterise Pathological signals from healthy signals.

6.1 SUMMARY AND CONCLUSION

This thesis presents an analysis of speech signals of different groups of subjects viz., subjects with stuttering, normal subjects and subjects with vocal fold pathologies. Voice signal analysis is carried out using traditional acoustic analysis methods as well as nonlinear time series analysis methods. The results of our analysis helped in distinguishing fluent speech from stuttered speech and the abnormal speech subjects with vocal pathologies from that of normal subjects. The results brings out the effectiveness of nonlinear methods compared to the perturbation methods in characterizing different voice signals . These results suggest that nonlinear analysis method can be beneficial in classifying pathological and disfluent speech signal. The appropriateness of PE for characterizing vocal signals on line clinical setup is also clearly revealed.

To summarise the achievements

1. Application of PE For detection of dynamical change in speech process is carried out.
2. Lower values of PE indicates the presence of repeated blocks of phonemes in speech signal of stuttering subjects.
3. Results are used for judgement of patient performance before and after speech therapy sessions in case of stuttering subjects. This helps in assessing the amount of improvement before and after treatment
4. PE effectively characterize vocal pathologies also .
5. PE values corresponding to abnormal voice is higher indicating increase in irregularity. Increase in PE indicates an increase in irregularity.
6. PE gives reliable results in the presence of noise.
7. Calculation time for PE for window size 512 samples of order 5 is less than nanoseconds.
8. Results obtained using PE analysis is compared with the existing acoustic measures of Jitter, Shimmer , Zero crossing, and fundamental frequency F0.

6.2 FUTURE SCOPE

In this thesis we have presented the result of our analysis of different speech signals of stutterers as well as subjects with vocal disorders. With an aim of characterizing these signals. The results of speech signal analysis using conventional perturbation analysis and nonlinear time series analysis methods are compared for assessing their applicability in an online evaluation setup. However, this work is still an initial study of the subject . Further work along this line can help in modifying and perfecting the approach of PE for online evaluation application. Also further work in investigating the dynamical changes introduced by pathologies in comparision with that of normal voice production will help in better diagnosis thereby increasing the efficiency of treatment modalities. Such methods if combined with the existing analysis tools will help in better classification and early diagnosis of different types of vocal disorders. Such approach will certainly improve the assessment efficiency of existing methods in determining the changes in improvement in speech quality rendered by different therapy methods.

Also the effectiveness of PE for various different pathological conditions can be investigated. The efficiency of other predictability measures can be investigated for a better understanding of the relation of pathologies with the complexity of the voice signals. Furthermore the interconnection between brain dynamics and stuttering is a potential field to be investigated. This can shed light into dynamics/dynamical changes of stuttered speech production.

Research publications:

International journals

- P. M. Radhakrishnan, Bindu M. Krishna, Usha Nair, V. P. N. Nampoori "Nonlinear Dynamic Analysis Of Stutterered Speech:A Preliminary Study Using Permutation Entropy" communicated to International Journal of Information Processing.
- P. M. Radhakrishnan, Bindu M. Krishna, Usha Nair, V. P. N. Nampoori "Permutation entropy measures applied to healthy and pathological voice characterisation" communicated to Journal of Acoustics Society of America.

International Conferences

- P.M Radhakrishnan, Usha Nair, V. N. N. Namboothiri "Characterization of Speech Using Time Series Analysis" ,2007 Jan, Joint Statistical Meeting and International Conference on Statistics, Probability and Related Areas ,Cochin University of Science and Technology Cochin, Kerala, India.

- P.M Radhakrishnan, Usha Nair, V. P. N. Nampoori, "Qualitative Analysis of Speech Signal in Time Series", 2004, Nov , COCOSDA, International conference on speech and Language Technology, New Delhi, India.
- P.M Radhakrishnan, Usha Nair, V. P. N. Nampoori, "Nonlinear Mathematical Technique to understand the Dynamics of Speech Process", 2004, Feb , International Conference on Mathematical Biology, IIT, kanpur, India.
- Usha Nair, P.M Radhakrishnan V. N. N. Namboothiri "Fractal Extraction of Sound Signal in Machining", 2007 Jan, Joint Statistical Meeting and International Conference on Statistics, Probability and Related Areas ,Cochin University of Science and Technology Cochin, Kerala, India.

- T67 -