

A HYBRID MODEL OF VOCAL FOLD VIBRATION
WITH APPLICATION TO SOME PATHOLOGICAL CASES

by

DARRELL WONG

B.E., The University of Auckland, 1982

A THESIS SUBMITTED IN PARTIAL FULFILMENT OF
THE REQUIREMENTS FOR THE DEGREE OF
MASTER OF APPLIED SCIENCE

in

FACULTY OF GRADUATE STUDIES

Department of Electrical Engineering

We accept this thesis as conforming
to the required standard

THE UNIVERSITY OF BRITISH COLUMBIA

August, 1985

• Darrell Wong, 1985

In presenting this thesis in partial fulfilment of the requirements for an advanced degree at the The University of British Columbia, I agree that the Library shall make it freely available for reference and study. I further agree that permission for extensive copying of this thesis for scholarly purposes may be granted by the Head of my Department or by his or her representatives. It is understood that copying or publication of this thesis for financial gain shall not be allowed without my written permission.

Department of Electrical Engineering

The University of British Columbia
2075 Wesbrook Place
Vancouver, Canada
V6T 1W5

Date: August 1985

ABSTRACT

It has been hypothesised (Moore 1976) that vocal fold pathology will manifest itself in voiced sounds when vibratory characteristics are disrupted. This thesis examines the effects that pathologies have on the vocal folds through the use of a computer simulation model of the human phonatory system.

A damped, nonlinear, multiple-mass spring model combined with a transmission line vocal tract model, was developed and mathematically simulated on a computer. Configurational parameters were then varied asymmetrically in order to examine the vibratory characteristics of the system. In particular, the glottal flow and speech signals from the glottal and vocal tract subsystems were observed for perturbations. Next, jitter, shimmer, and harmonics to noise ratio analyses were made and the results compared to a database of analysed speech recordings from Vancouver General Hospital. Finally, an approximate mathematical analysis was made examining the underlying nonlinear oscillatory phenomena.

The study showed that the model, a hybrid between the simple two mass Ishizaka and Flanagan model (1972) and the more complex Titze (1973, 1974) model, was able to simulate the desired asymmetrical conditions. Perturbation phenomena were successfully simulated and the results found to be in good agreement with both real data and data obtained from previously published models. The mathematical analysis revealed the observed perturbations to be characteristic of second and third order subharmonics found in nonlinear oscillatory systems. It was also shown that the driving forces discussed by Titze (1980) (ie the Bernoulli effect, vertical phasing and vocal

tract loading) all appear directly in the proposed dynamical equation.

Table of Contents

ABSTRACT	ii
List of Tables	viii
List of Figures	ix
Glossary of Terms	xv
Acknowledgements	xviii
1.0 INTRODUCTION	1
1.1. Motivation for the Study	1
1.2. Possible Approaches	1
1.3. Models - Past and Present	2
1.3.1 The Linear Model	2
1.3.2 Alternatives to the Linear Model	3
1.3.3 The Single Mass Model	3
1.3.4 Two Mass Models	4
1.3.5 Longitudinally Distributed Masses	4
1.3.6 Continuum Models	7
1.3.7 The Proposed Model	7
1.4. Aims of the Study	8
1.5. Thesis Organisation	8
2.0 VOCAL FOLD PHYSIOLOGY AND ANATOMY	10
2.1. Brief Overview	10
2.2. Control	10
2.3. Descriptors of Vibration	11
2.4. Macro Anatomy of the Laryngeal Musculature	12
2.4.1 Intrinsic Muscle Functions	13

2.5. Tissue Structure	15
2.6. Pathologies	17
2.6.1 Mass Lesions	18
2.6.2 Neurologic Diseases	19
2.7. Summary	21
3.0 MODELLING THE VOCAL FOLD AND ITS VIBRATIONS	22
3.1. A Description of Vibration	22
3.2. Structure of the Model	23
3.3. Nature of the Tensile Forces	24
3.3.1 Lateral springs	25
3.3.2 Longitudinal Forces	25
3.4. Damping Forces	29
3.5. Summary of Tissue Forces	31
3.6. Derivation of Flow Induced Pressures	32
3.6.1 Summary of Flow Induced Forces	36
3.7. Vocal Tract and Subglottis Models	38
3.8. Network Representation	39
3.9. Implementation Notes	39
3.10. Asymmetric Modelling	43
3.11. The Underlying Nature of the Aerodynamic Forces	43
3.11.1 Subglottal Pressure During Closure	43
3.11.2 Negative Bernoulli Forces	44
3.11.3 Vertical Phasing	44
3.11.4 Inertial Loading Forces	46
3.12. Summary	46
4.0 SIMULATION OF THE VOCAL FOLD	48
4.1. Selection of Parameter Values	48

4.1.1 Values for the Normal Vocal Fold	49
4.2. Simulation of the Normal Vocal Fold	53
4.2.1 Speech, Flow, and Area Waves	53
4.2.2 Velocity Wave	56
4.2.3 Phase and Energy Plots	58
4.2.4 Force and Displacement	61
4.2.5 Perturbation Analysis of the Normal Wave	63
4.2.6 Comparisons With Other Data	64
4.3. The Pathological Vocal Fold	68
4.3.1 Unilateral Stiffness Change	68
4.3.2 Localised Stiffness Changes	75
4.3.3 Discussion of Results from Stiffness Changes	78
4.3.4 Unilateral Mass Changes	82
4.3.5 Localised Mass Changes	87
4.3.6 Discussion of Results from Mass Changes	95
4.4. Comparisons with Other Pathological Data	99
4.4.1 Patients from Vancouver General Hospital	100
4.5. Summary	106
5.0 ANALYSIS	108
5.1. The One Mass Model	108
5.2. Fourth Order Systems	122
5.2.1 Expansion to a Fourth Order System	122
5.3. Perturbations	126
5.3.1 Further Examination of Subharmonics	126
5.3.2 Irregular Perturbation	128
5.4. Summary	129
6.0 SUMMARY, CONCLUSIONS AND RECOMMENDATIONS	130

6.1. Summary and Conclusions	130
6.2. Recommendations	132
REFERENCES	134
APPENDICES	137
APPENDIX A: Vocal Tract Measurement Data	137
APPENDIX B: Harmonics to Noise Ratio	138

List of Tables

Table I HNR values for all simulated cases	107
--	-----

List of Figures

Figure	Page
Fig1.1 Ishizaka-Flanagan two mass model. After Ishizaka and Flanagan, (1972).	5
Fig1.2 Multiple Mass Model Used by Titze. After Titze, (1973)	6
Fig2.1 A schematic of laryngeal muscle cartilage function. After Hirano, (1981); 1. thyroid cartilage; 2. cricoid cartilage; 3. arytenoid cartilage; 4. vocal ligament; 5. posterior cricoarytenoid ligament.	14
Fig2.2 Vocal fold tissue structure	16
Fig3.1 Elevated view	24
Fig3.2 Free body diagram of ith mucosal mass	26
Fig3.3 components of the longitudinal forces	27
Fig3.4 Stress-strain characteristic of Tv. After Titze, (1973)	30
Fig3.5 Closure conditions	37

Fig3.6, Network model of phonation	40
Fig4.1 Speech, area and flow waves for normal cords.	54
Fig4.2 (i) displacements and (ii) flow velocity and area, for normal fold	57
Fig4.3 Velocity phase plot showing stable limit cycle	59
Fig4.4 Equivalent restoring force vs displacement, showing energy exchange	60
Fig4.5 Equivalent restoring force. After Titze, (1983)	61
Fig 4.6 Areas and equivalent restoring force as functions of time	62
Fig4.7 Perturbation plots, normal speech, from model	63
Fig4.8 Results from the Ishizaka-Flanagan model using /a/. After Ishizaka and Flanagan, (1972)	65
Fig4.9 Speech and flow, normal. After Davis, (1976)	66
Fig4.10 Speech, ppd and lpa plots for normal case. From Vancouver General Hospital database.	67

Fig4.11 Speech, Flow Waves for Unilateral stiffness change, $T_{act} = 150$	69
Fig4.12 Single Trajectory Limit Cycle Oscillation, Unilateral Stiffness Change, $T_{act} = 150$	70
Fig4.13 ppd and lpa Plots for Unilateral Stiffness Change, $T_{act} = 25$	71
Fig4.14 speech, flow and area waves for unilateral stiffness change, $T_{act} = 25$	72
Fig4.15 Left and right displacements, unilateral stiffness change, $T_{act} = 25$	73
Fig4.16 Energy Exchange Plot, Unilateral Stiffness Change, $T_{act} = 25$	74
Fig4.17 lpa vs period #, Unilateral Stiffness Change, $T_{act} = 25$	75
Fig4.18 ppd, lpa plots for local stiffness change, $T_{act} = 150$	76
Fig4.19 Phase plot Localised Stiffness Change, $T_{act} = 25$	77
Fig4.20 ppd, lpa plots; Localised Stiffness Change, $T_{act} = 25$	78
Fig4.21 Flow and Displacements, normal, $T_{act} = 25$	79

Fig4.22 Phase plot, normal, $T_{act} = 25$	80
Fig4.23 Speech and Flow for Unilateral Mass Changes, $T_{act} = 150$	83
Fig4.24 Displacements for Unilateral Mass Changes, $T_{act} = 150$	84
Fig4.25 Phase plot for Unilateral Mass Changes, $T_{act} = 150$	85
Fig4.26 Speech, flow and area waves for Unilateral Mass Changes, $T_{act} = 25$	86
Fig4.27 Phase plot for Unilateral Mass Changes, $T_{act} = 25$	87
Fig4.28 Speech, flow and area waves for Localised Mass Changes, $T_{act} = 150$	88
Fig4.29 Phase plot for Localised Mass Changes, $T_{act} = 150$	89
Fig4.30 Speech and flow for localised mass changes, $T_{act} = 25$	90
Fig4.31 Displacements for localised mass changes, $T_{act} = 25$	91
Fig4.32 Phase plot for localised mass changes, $T_{act} = 25$	92

Fig4.33 Energy exchange plot for left fold, localised mass changes, $T_{act} = 25$	93
Fig4.34 Energy exchange plot for right fold, localised mass changes, $T_{act} = 25$	94
Fig4.35, ppd and lpa plots, localised mass changes, $T_{act} = 25$	95
Fig4.36, Speech and flow for bilateral local mass changes, $T_{act} = 25$	97
Fig4.37, Energy diagram for bilateral local mass changes, $T_{act} = 25$	98
Fig4.38, ppd and lpa plots for bilateral local mass changes, $T_{act} = 25$	99
Fig4.39 Speech, ppd and lpa plots, a 63 year old male with chronic laryngitis	101
Fig4.40 Speech, ppd and lpa plots, a 65 year old male with T1 Glottic cancer	102
Fig4.41 Speech, ppd and lpa plots, a 62 year old male with T1 Glottic cancer	103
Fig4.42, Frequency vs period number for dry hoarse and harsh hoarse voice.	
After Monsen, (1979).	105

Fig5.1 Four term approximation to $v(t)$	110
Fig5.2a Bernoulli forces	113
Fig5.2b Velocity dependent forces opposing motion	114
Fig5.2c Forces opposing the Bernoulli forces	115
Fig5.2d Velocity dependent forces aiding motion	116
Fig5.3 Frequency response for Duffings equation. After Meirovitch (1975)	120
Fig5.4 P_1 vs f_0 for the Ishizaka-Flanagan model. After Ishizaka and Flanagan, (1972)	121
Fig5.5 Frequency Fluctuation Plot. After Isshiki and Ishizaka (1976)	128

Glossary of Terms

abduction - when the opposing vocal folds separate and the gap widens

adduction - when the opposing vocal folds come together

autonomous - an oscillatory system which does not have a forcing function.

Bernoulli effect - when the flow in a pipe is abruptly constricted, the flow rate is maintained by increasing the flow velocity. As a consequence, the pressure falls, sometimes to a negative value, resulting in a 'sucking' effect.

compliance - acoustic capacitance, the ability to absorb and store more and more air in the same volume by increasing the density of the gas in the chamber.

convergent glottis - the shape of the vocal fold along the z axis determines whether the orifice is narrowest at the exit (convergent) or at the opening (divergent) of the glottal constriction.

dysphonia - synonymous with pathology, but usually referring to pathologies in which function is impaired.

epithelial - referring to the outer skin layer of the vocal fold.

epithelium - the outer layer of the fold which contains the mucosal fluid and covers the vocalis muscle.

flaccid - soft or limp reaction to the aerodynamic forces and the muscle contractions controlling the vocal fold's movement.

glottal - pertaining to the chamber, orifice or gap created between the vocal folds when they are apart.

glottis - the actual opening between the folds.

glottic chink - a triangular opening at the posterior end of the glottis which does not close during adduction. It is usually found in patients with Muscle Tension Dysphonia.

HNR - the Harmonics to Noise Ratio as defined by Yumoto, Gould and Baer (1980). It is a measure of the hoarseness of the voice.

inertance - acoustic inductance, the inertia of the air in the chamber causes a reluctance or delay in movement of the flow.

jitter - the cycle to cycle changes in the fundamental pitch period of the speech signal.

larynx - synonymous with vocal folds and vocal cords. The organic tissue that undergoes vibration.

laryngeal carcinoma - cancer of the larynx.

lpa - largest peak amplitude plot. Plots the largest positive amplitude peak in each period of the speech wave. It is a simple measure of shimmer.

mechanoreceptor - nerves sensitive to mechanical movement.

medial edge - the inner edge of the vocal fold that collides with the opposing fold.

mucosa - a term describing the mucous fluid layer and other fibrous layers that cover the vocalis muscle. In the simulation model, however, it refers to the upper mass that is supposed to represent the movement of the superficial layer of the lamina propria.

mucosal wave - refers to the wave-like motion of the mucosal fluid (the superficial layer of the lamina propria) in the vertical direction. It occurs as a result of the pressure below the folds. The term is synonymous with vertical phasing, which describes the phase delay as the wave travels from the bottom to the top surface of the fold.

ppd - pitch period duration plot. Plots the length (in milliseconds) of each period in the speech signal. The period is determined by a manual period marking procedure.

shimmer - cycle to cycle variations in the amplitude of periodic speech waves.

squamous cell layer - thin, flat, planar surface oriented cells.

subglottal - refers to the chamber below the larynx, which includes the lungs and the trachea.

supraglottal - refers to the vocal tract, the chamber above the larynx from the superior surface of the larynx to the lips.

transglottal pressure - refers to the pressure differential in the vertical axis between the entrance and exit of the glottal chamber. velum - a flap of tissue that can seal the nasal chamber from the vocal tract.

vena contracta - describes the rapid contraction of an orifice in a pipe, which results in a loss in pressure greater than would be found for a pipe with a gradual constriction.

vocal tract loading - the acoustical load of the vocal tract chamber, including the compliance, inertance and resistance.

vocalis - the muscular component of the vocal fold tissue.

Acknowledgements

Recognition of those who participated in this study is imperative, as without their help this thesis would not have been accomplished.

Thanks to my supervisors Dr. Mabo R. Ito and Dr. Murray Morrison for suggesting the topic and providing their knowledge and guidance.

The author acknowledges the financial support provided by the British Columbia Health Foundation.

Special thanks to Neil Cox who willingly contributed to many discussions and also wrote the HNR program, and to Dr. A. Soudack for the time spent discussing aspects of the analysis. To Dr I. R. Titze of the University of Iowa I owe a great deal, for without his participation, interpreting the results would have been far more difficult than they were.

Extra special thanks to my friend Harinath Garudadri for the motivation, to my friend Eleanor Chu-Chong for typing part of the thesis, and lastly to my girlfriend Carmelita Chan who proof read the manuscript and provided understanding and encouragement.

CHAPTER 1

INTRODUCTION

Research on the biophysics of speech has burgeoned in the last two decades with the advent of signal processing techniques and high-speed computers for analysis, synthesis, and simulation. However, the speech process is still not fully understood, and thus far, little work has been done in the area of synthesising pathological speech.

1.1. Motivation for the Study

The purpose of this study is to determine whether abnormal speech patterns are a direct manifestation of vocal fold pathology. This is of interest to medical researchers and speech scientists because it may aid in characterising the changes in the speech and relating them to the biomechanical changes caused by pathology. Hopefully, as the link between pathologies and their subsequent speech is better understood, improved features in the speech indicative of abnormality will be found. This could then lead to the earlier diagnosis of pathologies.

A further benefit is an improved understanding of the interaction between the air flow through the vocal fold opening and the tissue of the fold, and as a result the general understanding of vocal fold oscillation will be improved.

In this study, a computer model capable of simulating conditions found in pathological cases will be developed which, it is hoped, will allow abnormal speech to be synthesised, analysed, and ideally, classified according to specific pathologies.

1.2. Possible Approaches

The problem of quantifying the presence and effect of a pathology may be approached in two ways; (1) by analysing the speech wave through the use of signal processing and pattern recognition techniques and (2) through the synthesis of pathologies by computer modelling (ie. analysis by synthesis).

Researchers approaching the problem through the analysis of speech waves have experienced limited success – due in large part to the fact that this is essentially a "Black Box" approach. Features indicative of specific pathologies have been difficult to obtain. Consequently, it has been difficult to determine direct cause-effect relationships between the pathology type and specific perturbations in the resulting speech. Success has mainly occurred in cases where the discrimination is between pathological and normal classes only, without classification of pathology type (the features used with the greatest success have been those related to jitter (cycle to cycle period fluctuations) and shimmer (cycle to cycle amplitude fluctuations)). Nonetheless, as an aid to clinicians and as an objective method for quantifying perceptual tests, pattern analysis techniques are a useful tool with good possibilities in the future.

The second approach, analysis by synthesis, is the approach taken by this study.

1.3. Models - Past and Present

1.3.1 The Linear Model

As a first approximation, speech is considered to be a linear process (Fant 1960). The speech process is modelled as two linearly separable filters, representing the glottal and supraglottal chambers, excited by an impulse train. The glottal volume velocity, or flow pulse train, is the output of the two-pole low pass filter representing the glottal chamber. It is used as the input for the supraglottal, or vocal tract filter, which is either a cascade of two-pole filters or an acoustic tube of varying diameters and lengths, modelled as a transmission line network.

Because the model presented by Fant is linear, no interaction between the subglottal, glottal, and vocal tract systems is assumed. Work by Ishizaka and Flanagan (1972), Ananthapadmanabha and Fant (1982), and Rothenberg (1983) has shown this assumption to be clearly incorrect. Interaction does indeed occur between the glottis and the vocal tract and models based on this (Flanagan and Landgraf, 1968), (Ishizaka and Flanagan, 1972), (Titze, 1973) (Ananthapadmanabha and Fant, 1982) have been found to give synthesized speech a more natural sound as compared to the strictly linear synthesis systems (eg.LPC).

Another limitation of the linear model is that mathematical abstraction of the filtering effect of the vocal folds, (ie. its representation as a two-pole low pass filter), precludes modelling using physical or biomechanical parameters.

1.3.2 Alternatives to the Linear Model

To date, efforts by Flanagan and Ishizaka (1972) and Titze (1973, 1974, 1976, 1979) in this area have led to the development of a number of more complex models of the vocal cords - ones which involve nonlinear aerodynamics, nonlinear mass-spring systems, the simulation of multiple discrete mass systems, and which allow the use of biomechanical parameters. In these models the vocal cords are interactively coupled to the subglottal and vocal tract systems.

1.3.3 The Single Mass Model

The first quantitative self oscillating model was created by Flanagan and Landgraf (1968). The cords were modelled as single opposing masses fixed to the lateral laryngeal wall by nonlinear (cubic) springs and nonlinear dampers. Only lateral displacement was allowed. The external flow-induced force applied to the masses was determined by the pressure in the cavity between the opposing folds. The pressure was a function of the nonlinear resistance to the flow, which was in turn determined by

the cross-sectional area through which the flow occurred. Since the cross-sectional area was dependent on the position of the masses, a feedback relationship was created between the flow induced force and the tissue displacements.

A transmission line model of the vocal tract was concatenated to the Flanagan-Landgraf model allowing interaction between the vocal folds and the vocal tract. The model was self contained with excitation being provided by the constant lung pressure source. The one-mass model was successful in demonstrating sustained oscillation, but failed to oscillate for a capacitive vocal tract load and was incapable of demonstrating vertical phasing (a mode of vibration in which the upstream and downstream parts of the cords vibrate with a phase difference).

1.3.4 Two Mass Models

To incorporate these details, multiple mass models were developed with one of the most successful being a two mass model, (Ishizaka and Matsudaira, 1968, 1972), where the masses were aligned in the direction of flow so as to represent the upper and lower margins of the medial, or inner surface. The separation of the upper and lower margins enabled the model to simulate the travelling mucosal wave.

Ishizaka and Matsudaira's work was extended by Ishizaka and Flanagan (1972) who incorporated glottal impedance equations that included the inertial impedance of the air in the glottis. Fig 1.1 is a schematic, in frontal cross section, of their two mass approximation. The parameters in the diagram represent various mass, spring, and damping constants, as well as pressures at various points (P_{ij}) in the system.

1.3.5 Longitudinally Distributed Masses

Ishizaka and Flanagan's work was further extended by Titze (1973, 1974) who introduced longitudinally distributed masses to create a cord-like coupled mass system.

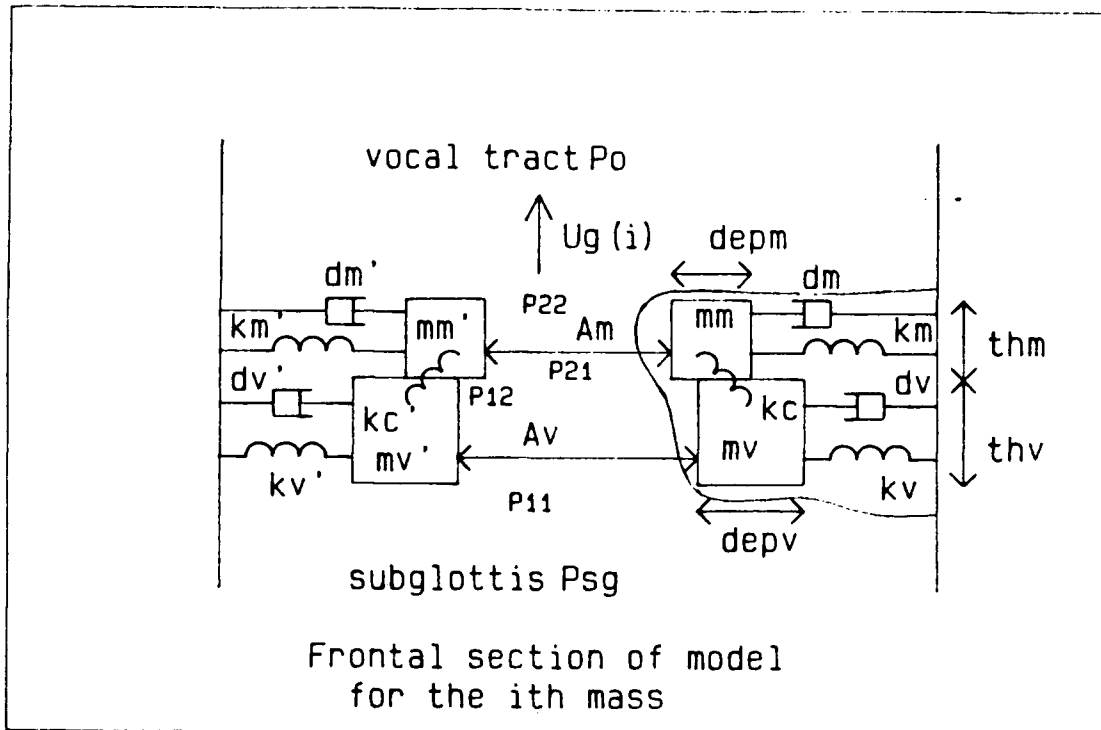


Fig1.1 Ishizaka-Flanagan two mass model. After Ishizaka and Flanagan, (1972).

According to Titze each vocal fold could be thought of as two strings in the longitudinal direction (normal to the flow, see Fig1.2) with unequal tensions and mass densities, coupled together, and fixed to the walls of the larynx. (In comparison, Ishizaka and Flanagan used single masses for each string and did not attach the string end points to the larynx).

The cords were physiologically associated with the epithelial mucous membrane and the vocalis muscle-ligament combination respectively. The longitudinal tensions in the strings restricted movement by acting in the lateral (x) and vertical (z) direction but not in the longitudinal (y) direction.

Titze observed, however, that for normal phonation the upper margin (mucosal string) started at the same vertical level as the vocalis string but moved to the region above the vocalis as soon as phonation started and remained above throughout phonation - indicating that the vertical degree of freedom was unnecessary (ie for

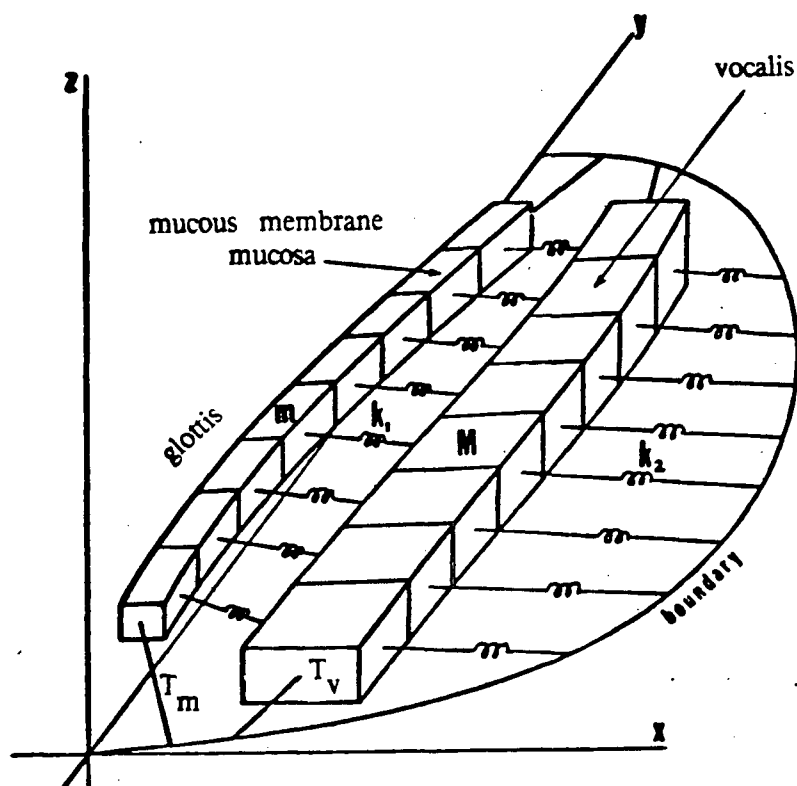


Fig1.2 Multiple Mass Model Used by Titze. After Titze, (1973)

normal phonation the behaviour of Titze's model was exactly the same as Ishizaka and Flanagan's).

The distribution of mass in the longitudinal dimension provided the flexibility to examine higher "string" modes of vibration that have been observed in films of the larynx. It also enabled localised parameter values to be used so that localised pathologies could be simulated. Unfortunately very little experimentation with pathology simulation was performed by Titze.

The only well documented study that has examined pathologies through the use of a computational model has been by Isshiki and Ishizaka (1976), using Ishizaka and Flanagan's two mass model. Their study was essentially one of observation with very little analytical results. They achieved some perturbation phenomena but could not account for its presence beyond stating that it was a nonlinear phenomena requiring asymmetric parameters. The study was also limited because there was no mass

distribution in the longitudinal direction.

1.3.6 Continuum Models

Two other models of note have been developed in recent years - one using finite difference techniques (Titze 1979), the other using finite element techniques (Titze and Haghighi, 1983). The finite difference model was three dimensional and used normal and shear stresses in a three-layered structure as biomechanical parameters. The parameters were made linear, however, and not cubic as in the discrete element models. Another drawback was that it was computationally expensive, requiring the solution to large matrix equations due to large regular mesh sizes.

The finite element model is the present state of the art, but it has not yet been refined, existing only as a two dimensional planar model. It is also computationally more expensive than the discrete model, although the ability to use irregular mesh sizes makes it faster than the finite difference technique.

1.3.7 The Proposed Model

In this study, we restricted ourselves to a discrete element model because it was felt that:

- finite difference and finite element models were computationally expensive
- the finite element model was only two dimensional and no software packages for three dimensional finite elements were available
- the existing finite difference model only used linear stress strain equations
- the discrete element model had already successfully produced some perturbations (Isshiki and Ishizaka, 1976).

Since the Isshiki-Ishizaka (1976) and Titze (1973) models were both partially successful, it was felt that a contribution could be made to the understanding of

pathologies by developing a model which combined features from both models – with attention focused on the examination of asymmetric and localised pathologies through the use of mass and stiffness changes. It was also felt that a mathematical analysis could be made using the simple lumped element model that would identify the physical reasons for self oscillation and perturbation phenomena.

1.4. Aims of the Study

The goals that have been chosen for this study are:

- to establish a computer model of the vocal folds which is reasonably close to physiology
- to make the model flexible enough so that asymmetric and localised abnormalities of biomechanical parameters can be simulated
- to explain phenomena that have occurred in these and other results, using nonlinear oscillation theory.

1.5. Thesis Organisation

The following chapters discuss various aspects of vocal fold modelling such as physiology, choice and implementation of the model, choice of parameters, analysis and validation of the model, and discussion of the results with respect to expected mathematical behaviour, as well as in comparison to observed results.

Chapter two introduces the physiology of the larynx and examines the vocal fold's role as a vibrator during phonation. Pathologies are briefly discussed so that the link between organic changes and analogous mechanical changes is established.

Chapter three discusses the phonatory system in terms of the physics involved and presents the mathematical relationships required to achieve a continuous time model.

Chapter four presents the results of the simulations and discusses them in comparison to results obtained from other models and to clinical experiments.

Chapter five analyses the results mathematically to show why one might deterministically expect perturbations to occur under certain conditions.

Chapter six, the concluding chapter, presents the major findings and suggests future work.

CHAPTER 2

VOCAL FOLD PHYSIOLOGY AND ANATOMY

In this chapter we examine the anatomy and physiology of the phonatory organs, the vocal fold's structural suitability for vibration, and the process through which phonation is achieved. By reviewing the physiology, the formulation of the proposed model will be made concrete.

The following description is adapted from Hirano (1980). It is abridged so as to include only the basics required for understanding the form and function of the musculature.

2.1. Brief Overview

The production of continuous speech requires complex and precise control of the peripheral phonatory organs by the central nervous system and other self-regulated subsystems. The phonatory organs include the pulmonary system, intrinsic and extrinsic laryngeal musculature, and jaw and lip articulators.

In this discussion, the speech signal is limited to voiced sounds since these sounds utilise the vocal folds heavily without complications due to movement of the vocal tract or to nasal tract coupling through the velum (though coupling to the vocal tract and subglottal tract are expected). The vibratory pattern of the folds can then be observed in relative isolation under steady state conditions.

2.2. Control

During speech, the speech centre of the cerebral cortex determines the sequence of sounds to be produced. Commands from this high order centre are transmitted to the motor cortex, which are linked by the motor nerves to the spinal cord. Nerves in the spinal cord distribute the excitations they receive to the respiratory, laryngeal and articulatory muscles. Sensors in the phonatory and other peripherally involved organs help in finely regulating the muscles so that the desired sounds are achieved. These feedback sensors are auditory as well as mechanoreceptive in nature.

2.3. Descriptors of Vibration

According to Hirano (1981) the parameters involved in phonation can be divided into three major groups:

1. Parameters regulating the vibratory pattern
2. Parameters describing the vibratory pattern
3. Parameters specifying the nature of the sound generated.

1. Parameters Regulating the Vibratory Pattern

Regulatory parameters can be either physiological or physical descriptors. Physiology alludes to those factors describing phonatory, articulatory and respiratory muscle activity. Physical descriptors are those factors used in equations to describe the motion eg. expiratory force, fold position, fold shape, elasticity, mass, and tract shape, among others. These primary physical factors determine secondary features resulting from the motion eg. the pressure profile between the glottal folds, the volume velocity, and the glottal impedance. These secondary factors are aerodynamic parameters.

2. Parameters Describing the Vibratory Pattern

The vibratory pattern is described by the fundamental period or frequency, measures of irregularities in amplitude or period (shimmer and jitter), wave

motion descriptors, contact area at glottal closure, and glottal area waveforms (the area through which air flows).

3. Parameters Specifying the Nature of the Sound Generated

The nature of the resulting sound output wave is largely determined by the vocal fold vibration pattern, which determines the flow through the vocal tract. The speech output is specified in acoustic or psychoacoustic terms. The acoustic parameters are fundamental frequency, intensity, and frequency spectrum, and the psychoacoustic parameters are pitch, loudness and voice quality.

2.4. Macro Anatomy of the Laryngeal Musculature

The human larynx is located near the base of the neck, linking the trachea to the vocal tract. Two fleshy folds align longitudinally along the posterior–anterior direction in the throat to form the larynx, and their function is to control the glottal constriction. The dominant function of the larynx is to act as a valve to control air flow to and from the lungs during respiration. A secondary function is phonation.

During phonation the folds act like a membranous constriction across the air passage with a slit at the centre. Steady pressure from below the larynx sets the flexible folds into vibratory motion, and the surrounding musculature fine tunes the control. The larynx consists of soft tissues supported by a skeleton of semirigid cartilages able to undergo limited movement. Movement of the cartilages is controlled by the intrinsic and extrinsic laryngeal muscles which align the cartilages in certain positions suitable for specific tasks such as swallowing or phonation. The intrinsic muscles affect the biomechanical properties of the folds, adjusting the shape, position and elasticity. Once this is achieved, the intrinsic muscles will then interact passively with the intraglottal air pressure to maintain phonation. The cartilages remain fixed as

long as steady phonation is desired, but once new sounds are required, the extrinsic muscles readjust the cartilage skeleton.

2.4.1 Intrinsic Muscle Functions

The following descriptions refer to Fig 2.1

1. Cricothyroid (CT) - When the CT contracts the folds are brought paramedially into the centre to line up with the posterior cricoarytenoid (PCA) and the anterior commissure. The entire vocal fold is stretched, elongated and thinned, causing the medial edge to be sharp. Mechanically, all layers are stiffened and the airway is opened in a triangular shape.
2. Vocalis (VOC) - Contraction of the vocalis muscle adducts the folds making them shorter and thicker, with a rounded edge. The body of the fold is actively stiffened while the covering layer is passively slackened. The airway is completely sealed.
3. Lateral Cricarytenoid (LCA) - Activation of this muscle brings the fold in medially, but lowers the tip of the vocal process of the arytenoid cartilage, causing the fold to elongate and become thinner. All layers are passively stiffened. There is a very small glottal chink present below the arytenoid cartilage.
4. Interarytenoid (IA) - Contraction of the IA adducts the arytenoid cartilage without changing the mechanical properties of the folds. The airway is not fully sealed.
5. Posterior Cricarytenoid (PCA) - The tip of the vocal process of the arytenoid cartilage is abducted and elevated, causing the fold to be lifted at the posterior end. The fold is thinned but the edge is rounded. Mechanically, contraction of the PCA causes all layers to be passively stiffened resulting in poor abduction.

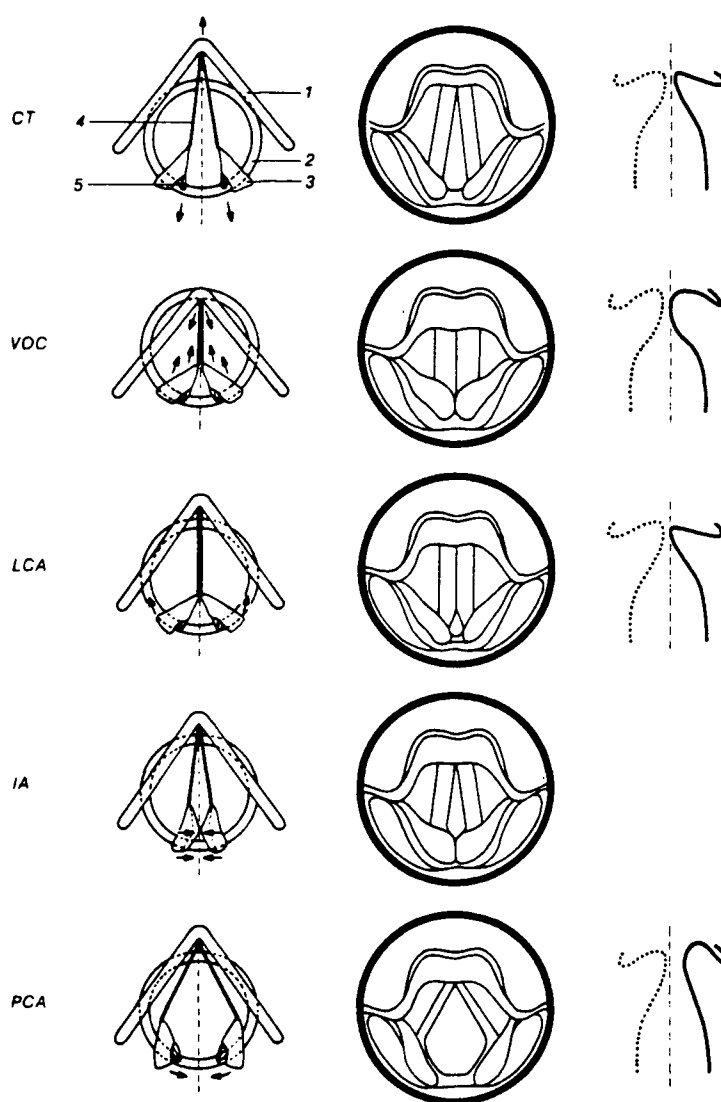


Fig2.1 A schematic of laryngeal muscle cartilage function (After Hirano, 1981); 1. thyroid cartilage; 2. cricoid cartilage; 3. arytenoid cartilage; 4. vocal ligament; 5. posterior cricoarytenoid ligament.

Before phonation, the combined activity of all these muscles controls the initial position of the fold, and during phonation the muscles regulate the tissue stiffness.

2.5. Tissue Structure

Humans possess the ability to phonate over a wide, continuous range of fundamental frequencies, intensities, and qualities. This is achieved in musical instruments by using multiple strings, whereas in humans it occurs with just a single pair of folds.

One reason for this remarkable ability is the intrinsic muscle control of the cartilage frame of the larynx. Because so many muscles are involved, a large variety of pre-phonatory shapes can be established.

Another more important reason, however, is the layered structure of the internal tissue of the cord. Hirano (1977, 1979) found that the human vocal fold consists of five layers - a vocalis muscle, and a mucous membrane (mucosa) consisting of four sublayers, which covers the vocalis. The layers are oriented longitudinally, making it easier for vibrations to occur laterally in string, or cord like modes, and each layer can have different biomechanical parameter values for stiffness, damping, mass, and volume, thus enabling the fold to exhibit a wide variety of properties.

The vocalis, which can be likened to stiff rubber bands oriented longitudinally, makes up the bulk of the mass of the vocal fold.

The mucosa is subdivided into the epithelium and three other layers, which are known together as the lamina propria.

- The epithelium, which consists of stratified squamous cells, is very thin and stiff (longitudinally), and its purpose is to maintain the shape of the fold.
- The three layers which make up the lamina propria are; the superficial layer

which contains loose fibrous components and has a consistency of soft gelatin; the intermediate layer which contains elastic fibers and can be compared to a bundle of soft rubber bands oriented longitudinally; and the deepest layer which consists almost entirely of collagenous fibers which are less flexible than the more superficial layers and mechanically comparable to bundles of cotton thread, again oriented longitudinally. The intermediate and deep layer of the lamina propria are commonly grouped together as the vocal ligament (see Fig2.2).

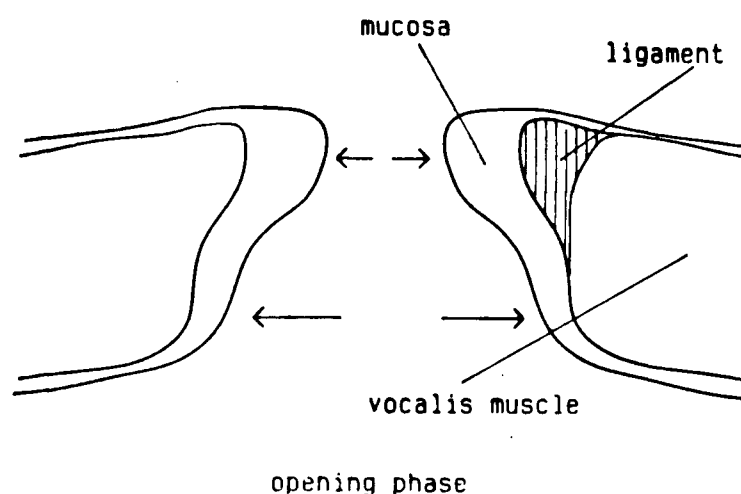


Fig2.2 Vocal fold tissue structure

The continuum model of Titze and Talkin (1979) treats the five layers as three: the cover - made up of the epithelium and the superficial lamina propria; the transition or ligament consisting of the intermediate and deep layers; and the body, consisting of the vocalis muscle. Each layer can have different biomechanical parameter values for stiffness, damping, mass, and volume. This enables the vocal fold to exhibit a wide variety of properties.

The longitudinal orientation of the various layers aids in vibration, making it easier for vibrations to occur laterally in string, or cord-like modes. The layered structure, however, is not uniform along the length of the vocal fold. At the anterior and posterior endpoints, the fold is strengthened by masses of collagenous fibers and other elastic fibers. These fibers protect against the stresses that occur during vibration, while the relative lack of collagenous fibres in the centre (longitudinally speaking) makes the centre part the most pliant.

Control of these properties is passive in the epithelial and lamina propria layers, and both active and passive for the vocalis. This means that the vocalis is able to control its stiffness by its own contraction, whereas the epithelium and lamina propria are externally controlled by the action of the intrinsic and extrinsic muscles as the cartilages are manoeuvred about. In summary, we observe that the tissue structure is conducive to vibration. The parallel orientation of the fibres permits string-like vibration modes, and the multiple layered structure, with control by a number of muscles, allows a wide range of frequencies to be produced. The fact that the outermost mucosal layer is jelly-like permits travelling waves, which move vertically from the sublaryngeal surface to the supralaryngeal surface of the larynx, to occur. Work by Baer (1975) in the area of high speed cinematography and Stevens (1977) in laryngeal modelling has shown this to be the case.

2.6. Pathologies

Pathologies of the vocal folds can be placed in two broad classifications - organic and functional.

An organic disease is a pathology which results in a vocal fold with tissue properties that are abnormal. This classification is further divided into mass lesions where growths are present, and neurological pathologies where the local nerve fibers and motor units are abnormal, resulting in disturbed feedback control.

For functional dysphonia, the origin is not organic. The cause is usually psychological, indicating that control of the voice is altered abnormally at a higher level of control (ie. the brain).

In sections 2.6.1 and 2.6.2 we discuss a few of the major pathologies, their etiology, and the likely mechanical changes caused by them. The discussion is limited to organic pathologies.

2.6.1 Mass Lesions

Mass lesions of the vocal folds can occur at any age, but here we limit ourselves to adult disorders. Any of the following changes in the folds can be produced by a mass lesion –

- an increase in the mass of the fold and surrounding tissue
- an alteration in the shape of the fold and the airway
- restricted mobility – possibly preventing complete closure
- a change in the tension in the fold – possibly causing excessive tightness at closure.
- a change in elasticity and viscosity

Examples of pathologies with mass lesions are:

- (a) Nodules – bilateral fleshy masses arising on the outer mucosal layer. They are usually one symptom of muscular tension dysphonia (MTD) (Morrison et al, 1983), the other signs being excessive tension of the arytenoid,

para-laryngeal and suprahyoid (extrinsic) muscles, and a posterior glottic chink (incomplete closure at one end even when the folds are fully adducted). The lesion often consists of fibrous matter. It remains on the surface layer and does not penetrate the vocal ligament. In biomechanical terms, mass and stiffness of the outer layer around the lesion is increased, with other layers unaffected. The glottal chink is a position of the folds where the posterior section, from the nodules to the arytenoid cartilages, never closes, even during phonation. This leads to a breathy and harsh quality in the voice. In general, MTD and the resulting nodules will occur in people who use the voice excessively in stressful conditions eg. singers and teachers. Nodules arise from the reaction of the tissue to abnormally forceful collisions of the folds. Chronic laryngitis is one specific manifestation of MTD. Because the cause of the nodules is not organic, removal from stress usually aids in recovery.

- (b) Carcinoma (cancer) - in almost all cases (Aronson, 1980), carcinoma originates in the epithelium. It is not restricted to this area however, and being malignant, it rapidly invades the ligament and vocalis, eventually consuming the entire vocal fold and beyond, increasing the mass and stiffness of any layer it invades. Hoarseness is usually the initially perceived symptom. Mechanically, a notable feature of carcinoma is that its effects are asymmetric and it may develop as a local lesion or it may consume the whole fold.

2.6.2 Neurologic Diseases

As discussed earlier neurophysiological control is necessary for maintaining phonation. Abductor-adductor muscle contractions maintain the tone of the vocal fold, keeping tension relatively constant, optimal and bilaterally symmetric.

Abnormal voice is inevitable when a disease attacks the neural system responsible for fine tuning phonation. The manifestations of abnormal voice may vary widely depending on the location of the lesion on the nervous system.

Neurological disorders can produce a consistent or a variable response in the acoustic signal, depending on whether the disease produces a constant or a fluctuating abnormal vocal fold oscillation. Here we discuss two examples where the effects are always present. For those disorders with a consistent response, a certain voice quality, loudness, and intensity, while abnormal, is always present.

An example of this is Parkinson's disease, or hypokinetic dysphonia. The patient's voice is breathy (due to possible incomplete bilateral closure), reduced in intensity, and reduced in pitch range. Nerve damage inhibits lower motor neurons causing rigidity and slowness of response in the laryngeal and respiratory musculature, explaining why the patient cannot achieve a wide frequency range or vary stress patterns. Biomechanically we might assume increased longitudinal stiffness over the entire length of both of the folds.

Another example of a constant disorder is flaccid dysphonia or laryngeal paralysis. With this disorder, the lesion can occur anywhere along the vagus nerve, with the exact position determining the degree of impairment. The disease causes bilateral or unilateral paralysis of the laryngeal muscles. The vocal fold becomes fixed, often resulting in incomplete closure during vibration. Other symptoms of paralysis are breathy voice, reduced intensity, and poor maximum phonation times. (An interesting feature which commonly occurs in unilateral paralysis is diplophonia. This is the simultaneous perception of two pitches. Less common is triplophonia, the perception of three pitches.) The mucosa is left unaffected by the muscle paralysis. Mechanically, the mass and stiffness of the body is decreased as the muscle atrophies. The cover and ligament, however, are not affected.

2.7. Summary

The process of phonation is controlled by a number of systems;

- the brain
- the auditory feedback system
- the laryngeal cartilage system
- and the mechanoreceptor feedback system.

Parameters describing phonation and vibration can be classified into three major groups;

- vibratory regulators
- vibratory descriptors
- voice output descriptors

The macro anatomy of the laryngeal musculature and the vocal fold tissue's structure is conducive to vibration.

The following pathologies have been described and possible biomechanical changes noted;

- nodules
- cancer
- Parkinson's disease
- and laryngeal paralysis.

CHAPTER 3

MODELLING THE VOCAL FOLD AND ITS VIBRATIONS

In this chapter we discuss the physics through which oscillation is maintained and present the proposed model. The first section describes the vibration, the next seven present the equations involved and their implementation, and the last section briefly describes the underlying phenomena that maintains oscillation.

Equations are drawn from fluid dynamics, mass-spring dynamical systems, and network theory, based on work by Fant (1960), Ishizaka and Flanagan (1972), and Titze (1973).

3.1. A Description of Vibration

During phonation, air leaves the lungs as a constant pressure source and is forced between the vocal folds. The interaction between the air as it passes through the glottis and the internal tensile forces in the vocal fold tissue results in sustained vibration. The amplitude of vibration causes the opposing folds to collide and seal the airway periodically, resulting in a train of flow pulses. The pressures between the folds at any instant are a function of the flow, and the external forces on the tissue are generated from these pressures.

Farnsworth (1940) produced some of the first photographs of vibrating vocal cords using high speed cinematography, capturing features of the vibratory pattern lost to stroboscopy. For phonation in the chest register (normal male phonation) he observed the medial surface of the folds to be relatively blunt and rounded, allowing contact with the opposing fold to occur over a vertical surface of several millimeters.

In a relaxed state, with closure occurring over several millimeters, the cords begin to open from below, and the opening progresses upward and outward. This vertical phase difference in upper and lower margins occurs simultaneously with the larger, dominant transverse motion, permitting the passage of air. The phase difference is the result of the mucosal fluid, which moves from the medial glottal to the supraglottal surfaces.

The motion of the vocal folds is a semi-passive response to the air flow. This concept is known as the myoelastic-aerodynamic theory of phonation, first discussed by Ferrein (1741) and later formulated in physical terms by Van den Berg (1958).

3.2. Structure of the Model

The proposed model is a hybrid of the Titze (1973) and Ishizaka-Flanagan (1972) models - distributing masses longitudinally, but permitting only one degree of freedom per mass (in the lateral direction) as in the Ishizaka-Flanagan model. Vertical motion is not permitted.

Fig 3.1 is an elevated view of the vocal cords, the frontal section appearing as for Fig 1.1. Masses m_m and m_v represent the upper and lower margins of the medial surface. They are connected to the wall by nonlinear springs k_m and k_v , representing the stiffness of the tissue, and dampers d_m and d_v , representing the viscous resistance of the tissue. The masses are coupled together by a nonlinear spring stiffness k_c .

Aerodynamic pressures (indicated as P_{ij} in Fig 1.1) are time and area dependent functions. The areas A_v and A_m are cross sectional and provide passage for the flow. T_v and T_m are the longitudinal tensions for the vocalis (lower) and mucosa (upper) regions respectively.

The rest of the transverse couplings in Fig 3.1 are not shown in order to preserve clarity. Large particle displacements are permitted and the end points of the

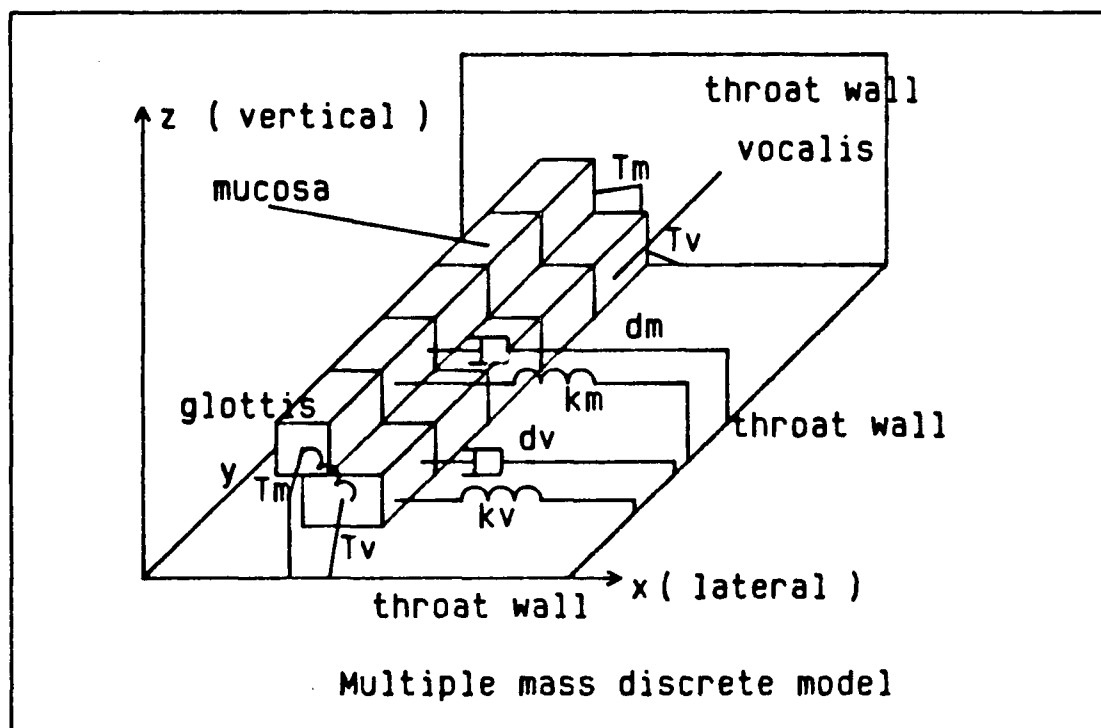


Fig3.1 Elevated view

longitudinal tension "strings" are fixed. Five masses per row has been chosen to minimise the computation effort, yet it permits localisation of parameter values. The number of longitudinal eigen modes has been reduced from infinity (for a continuum model) to a maximum of five. This is justifiable since the energy in higher modes should be negligible. With five masses, it is expected that at least the first three modes will be represented well.

3.3. Nature of the Tensile Forces

3.3.1 Lateral springs

Referring to Fig 3.1, k_c , the coupling spring between the upper and lower margins, represents the attachment between the epithelium holding the mucosal fluid and the underlying muscle. The springs k_v and k_m are an equivalent lateral tension caused by the contraction of the intrinsic muscles. They are given a nonlinear cubic characteristic to match the stiffnesses measured by Kaneko (1968). The lateral forces are thus of the form (Ishizaka and Flanagan, 1972):

$$\begin{aligned} F_{kv} &= k_v (\Delta x_v + \eta_v \Delta x_v^3) \\ F_{km} &= k_m (\Delta x_m + \eta_m \Delta x_m^3) \\ F_{kc} &= k_c ((x_v - x_m) + \eta_c (x_v - x_m)^3) \end{aligned} \quad \dots(3.1)$$

where Δx is the relative displacement from initial positions. These stiffness functions are known as cubic hardening functions for positive η .

During closure, the masses collide together, generating extra forces that push the masses back to equilibrium. This tensile deformation is represented as an extra spring force occurring only during collision ie.

$$\begin{aligned} F_{vcol} &= h_v (\Delta x_{cv} + \eta_v \Delta x_{cv}^3) \\ F_{mcol} &= h_m \Delta x_{cm} + \eta_v \Delta x_{cm}^3 \end{aligned} \quad \dots(3.2)$$

where Δx_{cv} and x_{cm} are the relative distances past the point of initial collision contact.

3.3.2 Longitudinal Forces

The forces caused by the longitudinal strings are more complicated, and depend on the relative positions of adjacent masses. Consider the free body diagram of the

ith mucosal mass in Fig3.2

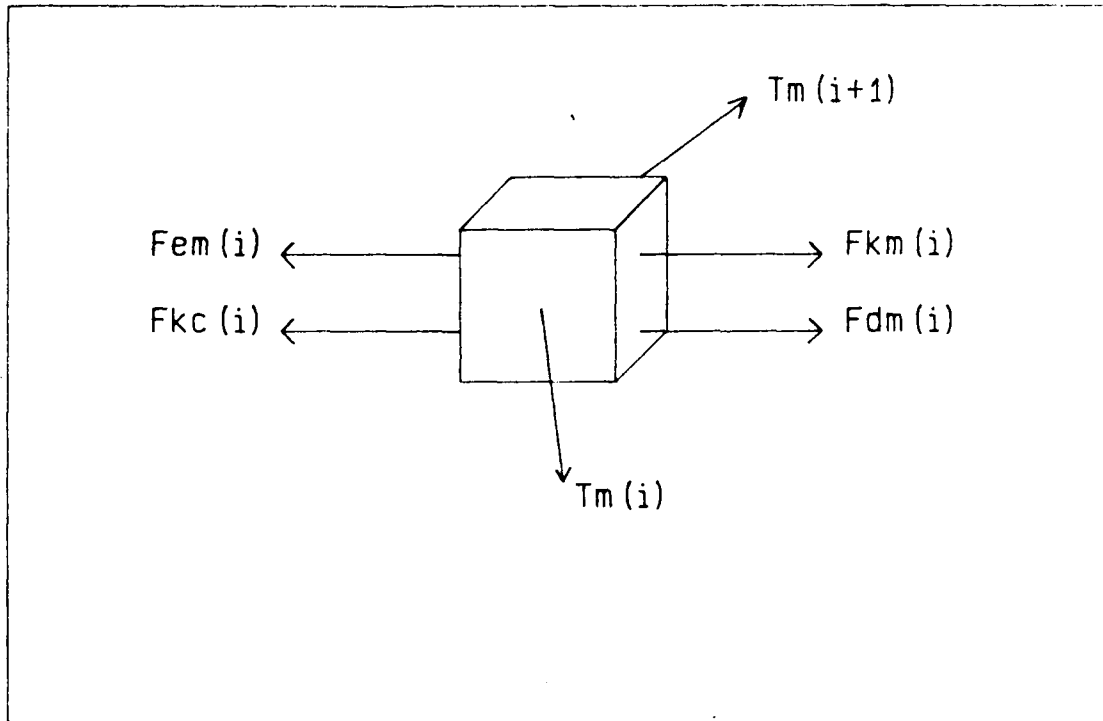


Fig3.2 Free body diagram of ith mucosal mass

$T_m(i)$ and $T_m(i+1)$ are the forces generated by the relative movement between adjacent masses. They produce components in the y and x dimension. As in Titze (1973), and Titze (1979) the masses have no motion in the y direction, so the effective contribution from $T_m(i)$ (and $T_m(i+1)$) is the sinusoid of the angle created between the mass positions (Fig3.3). Thus,

$$F_{tma}(i) = T_m(i) [x_m(i-1) - x_m(i)] / r_m(i)$$

$$F_{\text{tmp}}(i) = T_m(i+1) [x_m(i+1) - x_m(i)] / r_m(i+1) \quad \dots(3.3)$$

$r_m(i)$ is the shortest distance between the masses $m(i)$ and $m(i-1)$ (w, the longitudinal width of each mass is constant). Subscripts a and p indicate anterior and posterior positions relative to the i^{th} mass.

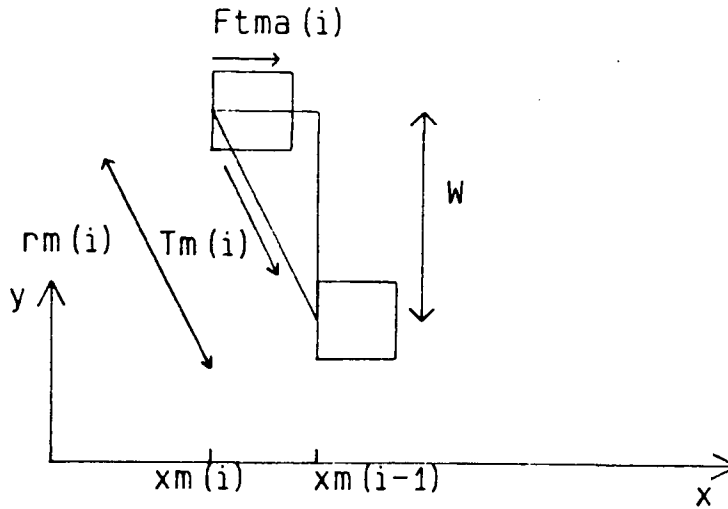


Fig3.3 components of the longitudinal forces

$T_m(i)$ is a nonlinear function of the relative strain between the adjacent masses. Van den Berg (1960), Hirano (1979), and Kaneko (1968) have made studies of the elastic properties of the various types of tissues in excised human larynges, and have shown that the general stress-strain characteristic is of the form

$$S = S_{\text{max}} (1 - e^{-T/\tau}) \quad \dots(3.4)$$

where T is the tension in g/cm^2 , S_{max} is the maximum strain, τ is a stress constant in g/cm^2 , and S is the unknown strain (Titze, 1973). The above equation can be

rearranged to calculate T , giving

$$T = \pm g \tau \ln(1 - |S|/S_{\max}) \text{th.dep} \quad \dots(3.5)$$

Where T is now measured in dynes, g is gravity in cm/s^2 , and th and dep are thickness and depth dimensions which give the cross sectional area through which T normally acts. The \pm sign indicates that compression and expansion can occur, depending on the strain. When strain S is positive, the negative sign is used to make T positive since the logarithm will always be negative. Conversely, when S is negative, the positive sign is used.

For the mucosa then,

$$T_m(i) = \pm g \text{th}_m(i) \text{dep}_m(i) \tau_m \ln(1 - |S_m(i)|/S_{m\max}) \quad \dots(3.6)$$

where $i = 1, \dots, 5$ masses

$T_v(i)$ is more complicated because the effects of the ligament and the muscle are combined, along with a constant tension T_{act} which is present regardless of the strain. T_{act} can be thought of as the stress due to vocalis muscle contraction, varying from about zero to 1000 g/cm^2 (10 g/mm^2) (Van den Berg, 1958). For steady phonation, T_{act} is constant. T_v is thus

$$T_v(i) = \pm g \text{th}_l(i) \text{dep}_l(i) \tau_l \ln(1 - |S_v(i)|/S_{l\max}) + g \text{th}_v(i) \text{dep}_v(i) T_{\text{act}} \\ \pm g \text{th}_v(i) \text{dep}_v(i) \tau_v \ln(1 - |S_v(i)|/S_{v\max}) \quad \dots(3.7)$$

Fig 3.4 graphs the tensions from equation 3.7 for the following parameter values;

$$\text{th}_l = 0.1 \text{ cm}$$

$$\text{th}_v = 0.3 \text{ cm}$$

$$\text{dep}_l = 0.1 \text{ cm}$$

$$\text{dep}_v = 0.3 \text{ cm}$$

$$\tau_l = 800 \text{ g/cm}^2$$

$$\tau_v = 300 \text{ g/cm}^2$$

$$g = 981 \text{ cm/s}^2$$

$$T_{\text{act}} = 150 \text{ g/cm}^2$$

$$S_{l\text{max}} = 0.4$$

$$S_{v\text{max}} = 0.8$$

The strains are calculated from nearest neighbour distances, $r(i)$. Hence, from Fig 3.3

$$\begin{aligned} r_m(i) &= \sqrt{[(x_m(i-1) - x_m(i))^2 + w^2]} \\ r_v(i) &= \sqrt{[(x_v(i-1) - x_v(i))^2 + w^2]} \end{aligned} \quad \dots(3.8)$$

$$\begin{aligned} \text{and } S_m(i) &= (r_m(i) - r_{mo}(i)) / r_{mo}(i) \\ S_v(i) &= (r_v(i) - r_{vo}(i)) / r_{vo}(i) \end{aligned} \quad \dots(3.9)$$

where $r_{mo}(i)$ and $r_{vo}(i)$ are the original nearest neighbour distances. The pre-phonatory displacements for each of the masses need not be of the same value, permitting triangular, rectangular, or elliptical longitudinal shapes to be used. It should be noted from Fig 3.4 that the vocalis strain S_v will never exceed 0.4 as the tension T_v will reach infinity at 0.4.

3.4. Damping Forces

The damping forces arise because of the viscous resistance to movement within and between tissues. In a similar fashion to Titze (1973), the damping coefficients are calculated from the equations:

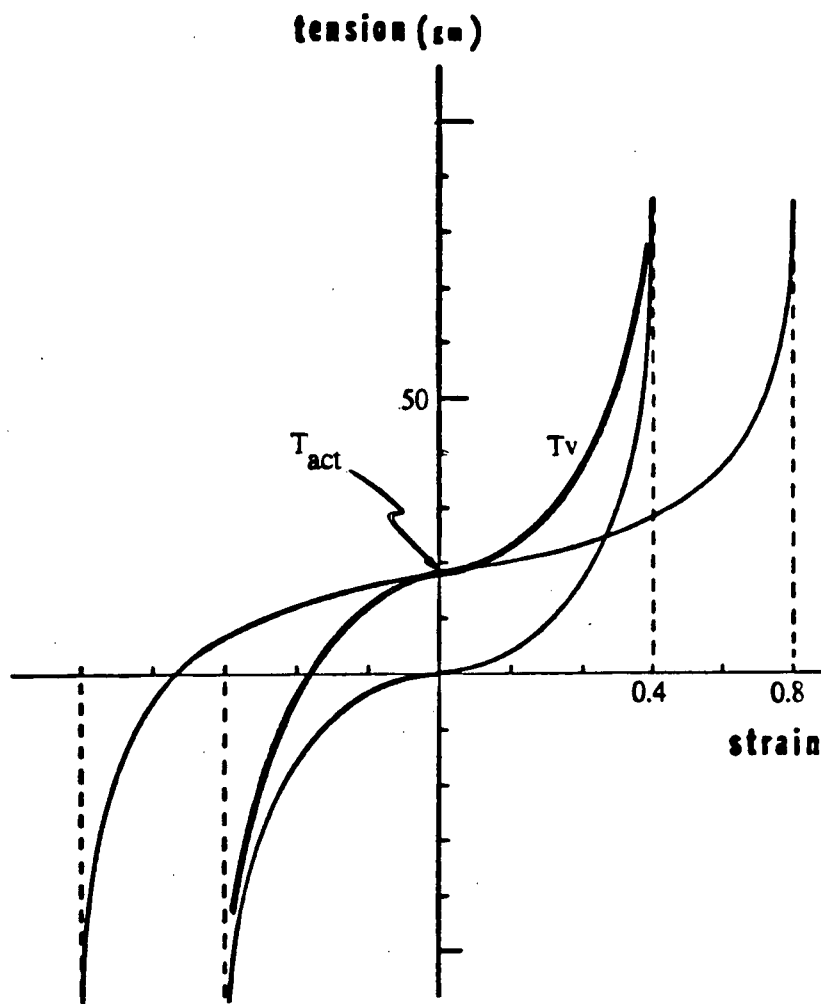


Fig3.4 Stress-strain characteristic of *Tv*. After Titze, (1973)

$$d_v(i) = 2\rho_v(i) / [m_v(k_v(i) + k_c(i) + |T_v(i)|/r_v(i) + |T_v(i+1)|/r_v(i+1))]$$

$$d_m(i) = 2\rho_m(i) / [m_m(k_m(i) + k_c(i) + |T_m(i)|/r_m(i) + |T_m(i+1)|/r_m(i+1))]$$

...(3.10)

Equations 3.10 are derived from the equation for critical damping for a simple mechanical vibrator ie.;

$$d = 2\sqrt{(mk)}$$

except that for a body surrounded by springs the equivalent k is the sum of the spring constants, regardless of the direction in which they act. Thus, $T_v(i)/r_v(i)$ and $T_v(i+1)/r_v(i+1)$ are the equivalent spring stiffnesses in the tension string, k_v and k_c are the linear components of the lateral stiffness constants, and ρ_v and ρ_m are the damping factors for the vocalis muscle and for the mucosal layer respectively.

3.5. Summary of Tissue Forces

Let us now collate the tensile and damping forces.

From equation 3.10, the damping forces are:

$$F_{dv}(i) = -d_v(i)\dot{x}_v(i)$$

$$F_{dm}(i) = -d_m(i)\dot{x}_m(i)$$

...(3.11)

From equation 3.7, 3.8, and 3.9, the longitudinal contributions to the tensile forces are:

$$F_{tva}(i) = T_v(i) [(x_v(i-1) - x_v(i))/r_v(i)]$$

$$F_{tvp}(i) = T_v(i+1) [(x_v(i+1) - x_v(i))/r_v(i)]$$

$$F_{tma}(i) = T_m(i) [(x_m(i-1) - x_m(i))/r_m(i)]$$

$$F_{tmp}(i) = T_m(i+1) [(x_m(i+1) - x_m(i))/r_m(i)]$$

...(3.12)

From equation 3.1 the lateral spring forces are:

$$\begin{aligned}
 F_{kv}(i) &= k_v(i)[(x_v(i) - x_{vo}(i)) + \eta_v(x_v(i) - x_{vo}(i))^3] \\
 F_{km}(i) &= k_m(i)[(x_m(i) - x_{mo}(i)) + \eta_m(x_m(i) - x_{mo}(i))^3] \\
 F_{kc}(i) &= k_c(i)[(x_m(i) - x_v(i)) + \eta_c(x_m(i) - x_v(i))^3] \quad \dots(3.13)
 \end{aligned}$$

The additional forces during collision are, from equation 3.2 :

$$\begin{aligned}
 F_{vcol}(i) &= h_v(i)[(x_v(i) - x_{vcol}(i)) + \eta_{vcol}(x_v(i) - x_{vcol}(i))^3] \\
 F_{mcol}(i) &= h_m(i)[(x_m(i) - x_{mcol}(i)) + \eta_{mcol}(x_m(i) - x_{mcol}(i))^3] \quad \dots(3.14)
 \end{aligned}$$

where:

- $x_{vcol}(i)$, and $x_{mcol}(i)$ are the displacements when the medial surfaces first collide.
- $x_{vo}(i)$ and $x_{mo}(i)$ are the initial pre-phonatory displacements

The dynamic equations for any mass pair (vocalis and mucosa) are then:

$$\begin{aligned}
 m_v \ddot{x}_v(i) &= F_{dv}(i) + F_{tva}(i) + F_{tvp}(i) + F_{kv}(i) + F_{kc}(i) + F_{vcol}(i) + F_{ev}(i) \\
 m_m \ddot{x}_m(i) &= F_{dm}(i) + F_{tma}(i) + F_{tmp}(i) + F_{km}(i) - F_{kc}(i) + F_{mcol}(i) + F_{em}(i) \quad \dots(3.15)
 \end{aligned}$$

where $F_{ev}(i)$ and $F_{em}(i)$ are the external forces generated by the flow. These forces are discussed in the next section.

3.6. Derivation of Flow Induced Pressures

The external forces arise due to air flow from the lungs. Expressions can be derived for the pressures on the medial surfaces using fluid mechanics. There are

several significant phenomena dictating the form of the expressions and these are;

1. The size of the glottis and fold.

Typically the dimensions of the vocal folds are 1.5 cm long, 0.3 cm thick, and 0.3 cm deep. The glottal opening is 1.5 cm long and up to 0.3 cm wide. These dimensions are much smaller than the fundamental frequency wavelengths in phonation (for $f < 4000$ Hz, the wavelength is > 10 cm), and so one dimensional planar waves are commonly assumed.

2. Gas compressibility.

The compressibility of a gas flowing through an orifice can be neglected if the ratio of orifice to atmospheric pressure is less than 0.01 (Ishizaka and Matsudaira, 1972). Based on Ishizaka and Matsudaira's work, this ratio is 0.008 for the vocal fold constriction.

3. Viscous resistance.

Schiller (1922) calculated that for smooth pipes, the critical Reynolds number for turbulence is

$$Re = rv/\nu = 1160$$

where r is the pipe radius, v is the average velocity of flow, and ν is the kinematic viscosity of air, or $\nu = \mu/\rho$ where μ is the shear viscosity and ρ is the density of the fluid. Ishizaka and Matsudaira established a Reynolds number of $Re = 1150$, and assumed that laminar flow occurred within the constriction. The flow was analogous to that between two parallel plates separated by a small distance h (the distance between the left and right medial surfaces). A pressure drop due to viscous resistance was calculated as

$$= R_v(i)U_g(i) + R_m(i)U_g(i) \quad \dots(3.16)$$

where $R_v(i) = [12\mu th_v(i)w^2] / A_v(i)^3$

and $R_m(i) = [12\mu th_m(i)w^2] / A_m(i)^3$

4. Presence of a Vena Contracta.

When the cross-sectional area from the subglottal to the glottal region contracts rapidly a vena contracta is produced, causing a greater loss in energy and pressure than would be expected from a less rapid transition. Van den Berg, Zanteema and Doornenbal (1957) empirically determined a loss factor for the drop in pressure between the subglottal cross-section and the entrance to the glottis to be

$$= R_c(i)U_g(i)^2 \quad \dots(3.17)$$

where $R_c(i) = 1.37\rho / (A_v(i)^2)$.

ρ is the air density, $U_g(i)$ is the glottal flow for the i_{th} branch, $A_v(i)$ is the cross sectional area at the entrance, and 1.37 is an empirical constant. According to Ishizaka and Matsudaira, the vena contracta is indicative of turbulent flow.

5. Pressure Recovery

At the exit of the upper medial surface, there is a pressure recovery as the flow leaves the constriction and enters the first supraglottal chamber. Ishizaka and Matsudaira have determined an expansion coefficient that is dependent on the area ratio ie

$$a = (1 - 5A_m(i)/A_0)^2$$

where A_m is the constriction area, and A_0 is the supraglottal chamber area.

The increase in pressure due to expansion is thus

$$= R_e(i)U_g(i)^2 \quad \dots(3.18)$$

where $R_e(i) = (1 - 5A_m(i)/A_0)^2(\rho/2)/A_m(i)^2$

6. Pressure Discontinuity

At the juncture of the upper and lower medial surfaces, the discretisation from a continuously diverging or converging glottis to one with just two areas causes a discontinuity in the pressure and velocity. Assuming a continuous flow,

$$U_g(i) = A_v(i)v_v(i) = A_m(i)v_m(i)$$

where $v_m(i)$ and $v_v(i)$ are the velocities of flow in the mucosal and vocalis chambers. The change in kinetic energy due to the change in pressure (Ishizaka and Matsudaira, 1968) is:

$$= \rho (v_v(i)^2 - v_m(i)^2)/2 = R_t(i)U_g(i)^2$$

$$\text{where } R_t(i) = \rho (1/A_m(i)^2 - 1/A_v(i)^2)(i)^2 \quad \dots(3.19)$$

7. Flow Inertia

Because the cords oscillate, the flow will vary with time. We can thus calculate a change in pressure due to the inertance of the air mass within the cord and the time varying flow (Ishizaka And Matsudaira, 1968). The change due to air inertia is:

$$= L_v(i)dU_g(i)/dt + L_m(i)dU_g(i)/dt$$

where $L_v(i) = \rho th_v(i)/A_v(i)$ and

$$L_m(i) = \rho th_m(i)/A_m(i) \quad \dots(3.20)$$

Depending on $dU_g(i)/dt$, the pressure change in equation 3.20 may be a gain or a

loss. Because the areas at the entrance and exit of the glottis are large, the inertances at the contraction and expansion are small enough to neglect.

3.6.1 Summary of Flow Induced Forces

Collating the pressure changes due to the contraction, internal glottis and expansion, a pressure distribution can be found (Ishizaka and Flanagan, 1972)

The resulting pressure profile is thus:

$$\begin{aligned}
 P_s - P_{11}(i) &= R_c(i)U_g(i)^2 \\
 P_{11}(i) - P_{12}(i) &= R_v(i)U_g(i) + L_v(i)dU_g(i)/dt \\
 P_{12}(i) - P_{21}(i) &= R_t(i)U_g(i)^2 \\
 P_{21}(i) - P_{22}(i) &= R_m(i)U_g(i) + L_m(i)dU_g(i)/dt \\
 P_{22}(i) - P_1 &= R_e(i)U_g(i)^2 \quad \dots(3.21)
 \end{aligned}$$

where the pressure points are shown in Fig 1.1. From this,

$$\begin{aligned}
 F_{ev}(i) &= (P_{11}(i) + P_{12}(i))w_{th_v}(i)/2 \\
 F_{em}(i) &= (P_{21}(i) + P_{22}(i))w_{th_m}(i)/2 \quad \dots(3.22)
 \end{aligned}$$

When closure occurs, there is no flow. This is modelled by increasing the glottal impedances to a very large value. There are, however, aerodynamic forces generated during collision in addition to those due to the collision of the tissues (F_{col}). These arise due to the transglottal pressure. The subglottal and post glottal pressures are present regardless of closure, building up when the glottis is closed and helping to force the masses apart. Thus, for different closure conditions (see Fig 3.5)

we present alternative equations for F_{ev} , F_{em} ...

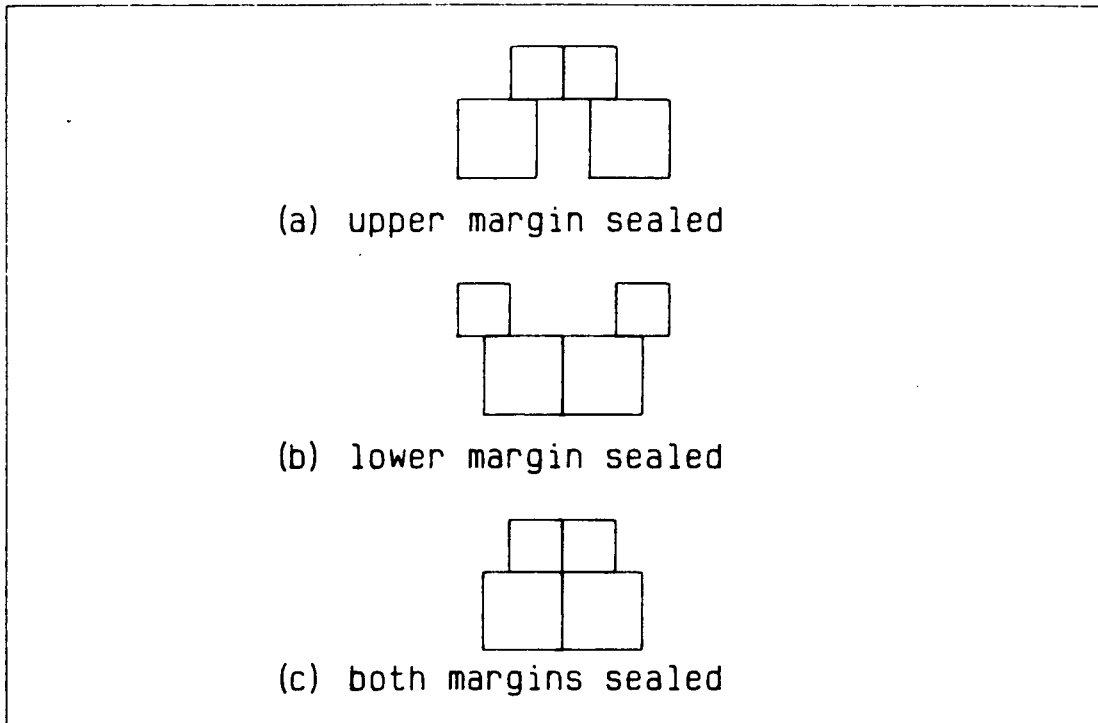


Fig3.5 Closure conditions

- (a) when only the upper medial edge is sealed

$$F_{em}(i) = (P_s - P_o)/2 \cdot th_m(i)w$$

$$F_{ev}(i) = P_s th_v(i)w \quad \dots(3.23)$$

- (b) when only the lower medial surface is closed

$$F_{em}(i) = P_o th_m(i)w$$

$$F_{ev}(i) = (P_s - P_o)/2 \cdot th_v(i)w \quad \dots(3.24)$$

- (c) when both surfaces are closed

$$\begin{aligned}
F_{em}(i) &= (P_s - P_o)/2 \left[th_m(i)/(th_v(i) + th_m(i)) \right] th_m(i)w \\
F_{ev}(i) &= (P_s - P_o)/2 \left[th_v(i)/(th_v(i) + th_m(i)) \right] th_v(i)w
\end{aligned}
\quad \dots(3.25)$$

3.7. Vocal Tract and Subglottis Models

The vocal tract is represented in this model as a transmission line of 10 cylindrical, hard walled acoustic tubes, of variable cross-sectional areas and lengths. The acoustic impedances are derived by Flanagan (1960) and Ishizaka and Flanagan (1972). To account for viscous losses due to friction along the walls, the acoustic resistance R_n is

$$R_n = 2\pi\sqrt{(A_n\rho\mu W/2\pi)} \alpha l_n A_n^2 \quad \dots(3.26)$$

where

A_n is the cross-sectional area

l_n is the tube length

ρ is the air density

μ is the kinematic viscosity

W is the fundamental frequency of the vocal folds (approximated as $\sqrt{(k_v/m_v)}$)

α is an empirical attenuation coefficient broadening formant bandwidths, usually equal to 25

The air inertance is approximated as

$$L_n = \rho l_n / (2A_n) \quad \dots(3.27)$$

and air compressibility is

$$C_n = l_n A_n / (\rho c^2) \quad \dots(3.28)$$

where c is the speed of sound. C_n is a significant impedance since A_n and l_n are much larger in the tract than in the glottis.

The vocal tract is terminated by the lips, which act as an integrator. The load is equal to that for a circular piston in an infinite baffle, and is represented by a resistance R_{rad} and inductance L_{rad} in parallel. Flanagan approximates these impedances by

$$R_{rad} = 128\rho c/(9\pi^2 A_{rad}) \quad \dots(3.29)$$

$$\text{and } L_{rad} = 8\rho/(3\pi v/(\pi A_{rad}))$$

where A_{rad} is the tube area created by the lips. A network representation of the entire system appears in Fig 3.6

The glottal constriction is coupled to the lungs via the tracheal subglottal tube. Ishizaka et al. (1976) estimated the cross-sectional area and length of a tube possessing the same acoustic impedance as the lung-trachea subglottal system to be equivalent to a simple acoustic tube 20 cm long and 2.5 cm² in area. In the proposed model, the subglottal impedances are defined in the same way as for the vocal tract i.e. transmission line resistance, inertance and compliance.

3.8. Network Representation

As there are five masses in the longitudinal dimension the glottal branch is repeated in parallel five times. Vocal tract and subglottal tract models are then concatenated to the glottal model to complete the speech system (see Fig 3.6). It should be noted that the glottal impedances are time varying and nonlinear functions of the constriction areas while the vocal tract and subglottis impedances are fixed.

3.9. Implementation Notes

1. The network in Fig 3.6 was solved using loop analysis. A continuous

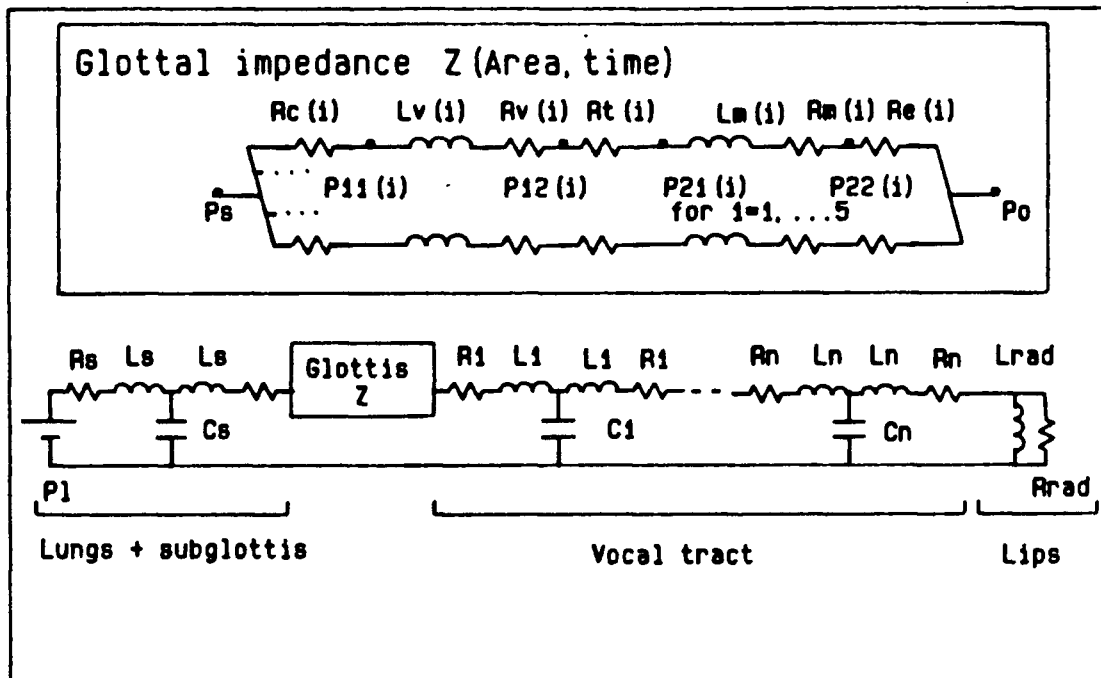


Fig3.6, Network model of phonation

language, ACSL, was used with the VAX 11/750 in the Electrical Engineering Department to model the system.

2. As the flow pulses have discontinuities due to closure of the glottis, the Adams-Moulton variable step integration method was used to maintain accuracy and speed.
3. Solution of the currents in the glottal loop (the five parallel branches) requires a matrix solution as instabilities occur when the differential equations are solved serially. The following differential equations were derived from Fig3.6:

(a) The subglottal loop;

$$dU_{sg}/dt = [P_L - R_{sg}U_{sg} - (1/C_{sg})\int_0^t (U_{sg} - U_{gtot})dt]/L_{sg} \quad \dots(3.30)$$

$$\text{where } U_{gtot} = \sum_{i=1}^5 U_g(i)$$

(b) The glottal loop;

$$\begin{aligned} [L][d\bar{U}_g] = & ([R][\bar{U}_g] + [R''][\bar{U}_g'] + [1/C]\int_0^t \bar{U}_g] \\ & + [\bar{T}](1/C_{sg})\int_0^t U_{sg}dt + [\bar{T}](1/C_1)\int_0^t U_1dt)/[L''] \end{aligned} \quad \dots(3.31)$$

$$= \bar{B}$$

where

$$\begin{aligned} b(i) = & [-(R_{sg} + R_1) \sum_{j=1}^5 U_g(j) - (R_v(i) + R_m(i))U_g(i) \\ & + (1/C_{sg})\int_0^t U_{sg}dt + (1/C_1)\int_0^t U_1dt - (1/C_{sg} + 1/C_1) \sum_{j=1}^5 U_g(j)dt \\ & + (R_c(i) + R_e(i) + R_t(i))U_g(i)^2] / (L_v(i) + L_m(i) + L_{sg} + L_1) \end{aligned} \quad \dots(3.32)$$

The right hand side of 3.31 consists of values known from the previous time step. The equation is thus of the form $[L]d\bar{U} = \bar{B}$. The $[L]$ matrix elements are calculated from the loop equations and are;

$$[L] = \begin{bmatrix} 1 & a1 & a1 & a1 & a1 \\ a2 & 1 & a2 & a2 & a2 \\ a3 & a3 & a3 & a3 & a3 \\ a4 & a4 & a4 & 1 & a4 \\ a5 & a5 & a5 & a5 & 1 \end{bmatrix}$$

where

$$a_i = (L_{sg} + L_1)/(L_{sg} + L_1 + L_v(i) + L_m(i)) \quad \dots(3.33)$$

(c) Vocal tract loops are solved using network loop analysis for $k=1, \dots, 10$

If $k=1$

$$\begin{aligned} dU_1 = [& -(R_1 + R_2)U_1 - (1/C_2)\int_0^t (U_1 - U_2)dt \\ & - (1/C_1)\int_0^t (U_1 - \sum_{i=1}^5 U_g(i)dt)] / (L_1 + L_2) \end{aligned}$$

If $k= 2, \dots, 8$ then

$$\begin{aligned} dU_k = [& -(1/C_k)\int_0^t (U_k - U_{k-1})dt \\ & - (1/C_{k+1})\int_0^t (U_k - U_{k+1})dt \\ & - (R_k + R_{k+1})U_k] / (L_k + L_{k+1}) \end{aligned}$$

If $k= 9$

$$\begin{aligned} dU_9 = [& -(R_{rad} + R_9)U_9 + R_{rad}U_{10} \\ & - (1/C_9)\int_0^t (U_9 - U_8)dt] / L_9 \end{aligned}$$

If $k= 10$

$$dU_{10} = - (U_{10} - U_9)R_{rad}/L_{rad} \quad \dots(3.34)$$

The R_i , L_i and C_i are the impedances for the i th vocal tract tube. Once the above differential equations are solved, the newly found currents are used to calculate new pressure drops and aerodynamic forces. The tissue displacement equations in equation 3.15 are solved using the aerodynamic forces, and from the displacements the new

cross-sectional areas are found. Now the new glottal impedances can be calculated, thus completing the cycle.

3.10. Asymmetric Modelling

To allow for distinct left and right cords, the equations are duplicated so that the cords are modelled separately, but coupled together through the common flow. The area through which the flow occurs is calculated as:

$$A_v(i) = w(x_{vl}(i) + x_{vr}(i))$$

$$A_m(i) = w(x_{ml}(i) + x_{mr}(i))$$

where the displacement is measured from the midline between the folds.

3.11. The Underlying Nature of the Aerodynamic Forces

The pressures generated in the glottis arise from the air flow. Unfortunately, it is difficult from the acoustic circuit analogy to appreciate the significance of the interaction and timing of the components of the aerodynamic forces because they are not explicitly derived in circuit equations 3.30 to 3.34. The aerodynamic forces occurring under certain conditions during the oscillation period are:

1. subglottal pressure during closure
2. impulsive negative Bernoulli forces generated when the folds are open but narrowly constricted (ie. just before closure and just after opening).
3. velocity dependent forces created by the vertical phasing between the upper and lower margins
4. velocity dependent forces created by the glottal and vocal tract load impedance

3.11.1 Subglottal Pressure During Closure

At closure, a buildup of pressure due to the constant lung pressure source helps to force the folds apart. The lower margin leads the upper margin, and air flows through once the folds open completely. Once air flows, the subglottal pressure decreases and the pressure profile becomes the dominant driving force.

3.11.2 Negative Bernoulli Forces

The Bernoulli force is a "sucking" force which arises when fluid enters a narrow orifice. The constriction causes a rapid drop in pressure and an increase in velocity. Often the pressure in the orifice can be negative relative to atmospheric pressure, sucking the medial margins inwards. The force is impulsive - greatest just before closure and just after opening, when the orifice is smallest.

Early researchers assumed that the dominant force required to sustain oscillation was due to the Bernoulli effect (Van den Berg, 1958), (Fant, 1960). However, the Bernoulli effect creates impulsive pressures of the same magnitude and sign, without regard to the direction of tissue movement. This means that the energy supplied to the tissue by the Bernoulli force just before closure is cancelled by the Bernoulli force opposing motion that arises just after opening. There is thus no net gain of energy to overcome tissue damping and help sustain oscillation.

Ishizaka and Matsudaira (1968, 1972), Stevens (1977), and Titze (1980, 1983, 1985) have suggested that other mechanisms, such as vertical phasing and inertial loading by the glottis and vocal tract, contribute to the driving force and input the energy necessary to overcome tissue damping into the system.

3.11.3 Vertical Phasing

Experimental work by Ishizaka and Matsudaira (1968) (cit: Stevens (1977)) on steady state converging, diverging, and uniform glottal shapes has shown that the lower

margin's pressure is dependent on the relative width between the left and right upper margins. For instance, when the upper width is very small and the lower width is much larger (a convergent glottis), the pressure on the lower margin is positive, aiding outward movement (this situation occurs when the fold is in its opening cycle).

Conversely, when the lower width is small and the upper width is much larger (a divergent glottis shape), the pressure on the lower margin is negative, causing the lower margin to be sucked into closure. Again, this aids the direction of movement as the situation only occurs when the fold is in its closing cycle. For a uniform glottis (upper and lower widths the same), the pressure is always close to zero, providing no driving force. Phase lag is thus a necessary component for net energy input from the aerodynamic flow.

Ishizaka and Matsudaira showed that the direction of the aerodynamic force on the lower margin depended on the relative displacements of the margins and aided the direction of motion

$$\text{ie. } F_{\text{ext}} = \phi(x_v - x_m)$$

where v is the vocalis, or lower margin; m is the mucosal, or upper margin; and F_{ext} is the vertical phasing contribution to the aerodynamic force.

The magnitude of the aerodynamic stiffness ϕ is inversely proportional to the average glottal width of the margins. Thus ϕ is relatively large for situations near closure, and small when the glottal width is large.

F_{ext} helps to sustain oscillation by providing some of the energy required to overcome the losses due to tissue damping. By adding to the Bernoulli forces when the glottal width is small, the total aerodynamic force is large and negative just before closure (aiding the direction of motion) and small and negative just after opening. Thus while the force just after opening opposes motion, it does so with less energy

than just before closure and so the resulting asymmetric force provides a net energy input.

3.11.4 Inertial Loading Forces

Titze (1983) demonstrated that inertial lag due to glottal and vocal tract inductance could create a velocity dependent force similar to the force generated by phasing, hence reinforcing the effects of the vertical phasing force. Depending on the conditions, the resulting summative force just after opening could be either slightly positive or negative.

3.12. Summary

In chapter three we have proposed a nonlinear multiple mass-spring system as a representation of the vocal folds.

Vocal tract and subglottis models have been concatenated using a transmission line representation. Implementation of the model has required a matrix representation to ensure stability.

The proposed model simplifies the Titze (1973) model by eliminating the z degree of freedom, thus decreasing computation time. The frontal (side) view is the same as that for the Ishizaka-Flanagan (1972) model. All equations may be found in Titze (1973), and Ishizaka and Flanagan (1972).

The underlying forces maintaining oscillation are not explicitly represented in the model equations, but analyses by Ishizaka and Matsudaira (1972), and Titze (1983, 1985) have shown that -

- the Bernoulli effect
- vocal tract loading

- subglottal pressure during closure
- and vertical phase difference

play a part in sustaining oscillation. The terms related to the above phenomena and other perturbatory phenomena will be derived in chapter five.

CHAPTER 4

SIMULATION OF THE VOCAL FOLD

In this chapter, the simulations are presented and analysed. The chapter is divided into two major sections - the normal fold, and the pathological vocal fold.

In the first section, time domain signals such as the speech pressure wave, the glottal volume velocity, and the tissue displacement waves are introduced. All plots are made using cgs units. The characteristics of these signals are compared to equivalent signals obtained from the Ishizaka-Flanagan model and from the database at the Voice Lab at Vancouver General Hospital.

In the second section, various asymmetric configurations are examined. The resulting waves are analysed using jitter, shimmer and harmonics-to-noise ratio measures (Yumoto, Gould and Baer, 1982) and then compared to:

1. the VGH database
2. analytical results from speech data gathered by Monsen (1979 NIH)
3. Isshiki and Ishizaka's asymmetric model (Isshiki and Ishizaka, 1976)

4.1. Selection of Parameter Values

The tissue parameter values used in this model are derived from work by Kaneko (1968), Titze (1973), and Hirano (1979). Since the model is a hybrid between the two discrete mass models previously discussed, the values do not agree exactly with those used by the original authors, but are generally within the same order of magnitude. The parameter values used in the studies mentioned above have been

"finetuned" to establish a set of parameters more amenable to the hybrid nature of the model, the overall objective being to achieve speech, volume velocity, and displacement waves consistent with experimental evidence, with less regard to matching specific parameter values.

The modelling of pathologies has proven to be an arduous task. While some parameters for the normal fold have been published, (Kaneko, 1968), there have been very few studies on pathologies. Hirano (1979) suggests trends or relative changes from the normal case for some parameters, but specific values are not given. Only Isshiki and Ishizaka's study (1976) using the two mass model has provided any concrete figures (the parameter ranges they chose were probably determined by trial and error, however). Since their work was performed using a model (as in this study) rather than with real speech, the perturbations achieved were due to the characteristics of the model and not the real pathological larynx, making the study less conclusive. Due to the lack of data for pathologies, there is a danger of assuming a physically unrealisable set of values which still achieves the desired results. This problem cannot be solved until experimental data for the pathological cord becomes available.

4.1.1 Values for the Normal Vocal Fold

The following parameters represent the normal vocal cord in this model. All values apply to both left and right cords.

1. Mass - Typical values for the lower and upper masses are the same as in Ishizaka and Flannagan's model ie.

$$m_v = 0.025 \text{ g}$$

$$m_m = 0.005 \text{ g for } i = 1, \dots, 5$$

2. Lung pressure P_L - Chosen to be 10cm H₂O or 10 kdyne/cm². This is

within the typical range suggested by Hirano (1981).

3. Dimensions - The longitudinal length, vertical depth and lateral thickness are taken from Titze (1973) to be:

$$w = 0.28 \text{ cm}$$

$$\text{dep}_v(i) = 0.25 \text{ cm}$$

$$\text{dep}_m(i) = 0.1 \text{ cm}$$

$$\text{dep}_l(i) = 0.1 \text{ cm}$$

$$\text{th}_v(i) = 0.25 \text{ cm}$$

$$\text{th}_m(i) = 0.05 \text{ cm}$$

$$\text{th}_l(i) = 0.1 \text{ cm respectively.}$$

4. Lateral stiffness - The equivalent linear parameters used by Ishizaka and Flanagan were:

$$k_v = 16 \text{ kdynes/cm}$$

$$k_c = 5 \text{ kdynes/cm}$$

$$\text{and } k_m = 1.6 \text{ kdynes/cm.}$$

As the proposed model has additional stiffnesses due to the longitudinal tensions, the lateral stiffnesses chosen by Ishizaka and Flanagan have been reduced to compensate. The appropriate values have been determined by trial and error to be

$$k_v = 5 \text{ kdynes/cm}$$

$$\text{and } k_m = 0.5 \text{ kdynes/cm.}$$

The criterion for selection was based on the achievement of speech, volume velocity flow, and edge displacement waves consistent with published speech

data and existing models. In conjunction with the lateral stiffnesses, the longitudinal tensions permit the maximum lateral excursion of the mucosal and vocalis masses to be approximately 0.12 cm and the glottal flow to be between 300 and 500 cm³/s, which is typical for the healthy cord (Hirano, 1983).

$k_c = 3.5$ kdynes/cm gives a phase lag of 60° between the upper and lower masses. This phase lag agrees with Baer's results estimating a 1 m/s vertical travelling wave velocity, and a 60° lag for a glottis of similar dimensions to the proposed model (Baer, 1975).

5. Nonlinear lateral coefficients are -

$$\eta_v, \eta_c, \eta_m = 100$$

$$h_v = 3k_v = 15 \text{ kdyne/cm}$$

$$h_m = 3k_m = 1.5 \text{ kdyne/cm}$$

$$\eta_{vcol}, \eta_{mcol} = 500$$

These values are based upon measurements of static stress-strain measurements for an excised human larynx by Ishizaka and Kaneko (1968). Ishizaka and Flanagan (1972) demonstrated that the given values for η produced a curve of lung pressure versus fundamental frequency with a rise of 2.5 Hz/cm H₂O in their model, which is in good agreement with measurements by Baer (1975).

6. Longitudinal stiffness The stress constants and maximum strains have been estimated by Hirano (cit: Titze 1980) using canine tissue. The values are approximately the same for humans, and estimates from the curves yield

$$\tau_m = 800$$

$$\tau_l = 800$$

$$\tau_v = 300$$

$$S_{mmax} = 0.3$$

$$S_{lmax} = 0.4$$

$$S_{vmax} = 0.8$$

$$T_{act} = 1.50 \text{ g/cm}^2$$

These values differ significantly from the data used in Titze's discrete model, but the resulting vibratory patterns in our results agree well with the Ishizaka-Flanagan model, suggesting that Hirano's experimental data is more accurate than that used by Titze (1973), who used 5 g/mm^2 for T_{act} , equivalent to the chest register of phonation. The proposed hybrid model used values between 0.25 and 2.0 g/mm^2 , rather than 5 g/mm^2 since abduction was difficult to achieve using the greater value.

7. Damping - ρ_v and ρ_m have been experimentally determined by Kaneko (1968) and Flanagan and Landgraf (1968) to be of the order of 0.1 and 0.6 respectively when not in collision, and 1.1 and 1.6 (close to critical damping) during collision. The increased damping during collision represents adhesion between the medial surfaces, causing energy to be lost as the surfaces stick together.
8. Initial configuration - The prephonatory position of the masses is controlled by muscular adjustment (section 2.4). Ishizaka and Flanagan estimate the typical initial glottal cross sectional area to be 0.05 cm^2 for both the upper and lower margins. In the proposed model, an oval constriction shape is used with the following initial displacement values.

$$x_{vo}(1), x_{vo}(5), x_{mo}(1), x_{mo}(5) = 0.012 \text{ cm}$$

$$x_{vo}(2), x_{vo}(3), x_{vo}(4), x_{mo}(2), x_{mo}(3), x_{mo}(4) = 0.021 \text{ cm}$$

The total cross-sectional areas A_v and A_m each sum to 0.05 cm^2

9. Vocal tract - As previously mentioned in section 3.6, the vocal tract consists

of 10 cylindrical acoustic tubes of variable cross-sectional area and length. For these simulations the vowel /a/ was adapted from Fant (1960) (see Appendix A).

4.2. Simulation of the Normal Vocal Fold

The normal configuration of the vocal fold was simulated and various parameters recorded for the vowel /a/. The following discussion interprets the results.

4.2.1 Speech, Flow, and Area Waves

In the proposed electrical/acoustical analogy, the speech (SPEECH) and glottal volume velocity (UTOT) waves are equivalent to voltage/pressure and current/flow respectively. The area waves represent the total glottal area open to the flow measured at the upper (A_{mtot}) (AMTOT) and lower (A_{vtot}) (AVTOT) margins. Points to note in Fig 4.1 are:

1. The negative areas indicate closure of the glottis and continued displacement of the centre of mass. The area waves are asymmetric when negative due to conditions during, just before and just after collision when the driving forces are large and asymmetric.
2. The constant delay of approximately 60° between the lower and upper margin displacements (with A_{vtot} leading A_{mtot}) is indicative of vertical phasing and occurs in spite of the fact that both margins are initially at the same width. The two masses synchronise with a constant phase delay because of mutual synchronisation (also observed by Minorsky (1962) and Ishizaka and Matsudaira (1972)).

A qualitative explanation has been made by Titze (1976, 1980) which suggests that the masses of Ishizaka and Matsudaira's system vibrate with a

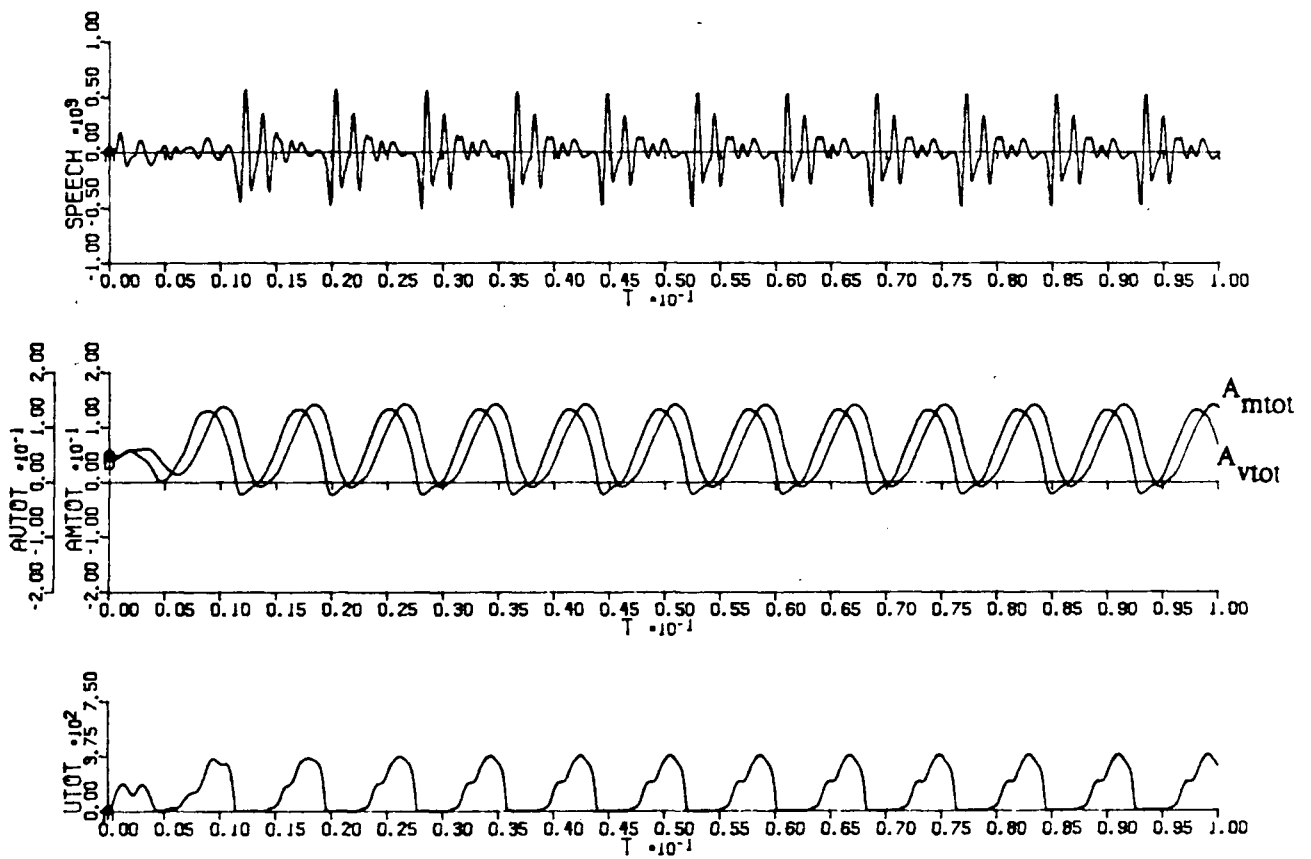


Fig4.1 Speech, area and flow waves for normal cords.

combination of the first two modes of the coupled system. The first mode requires the masses to move in phase, while the second mode requires them to vibrate 180° out of phase at a higher frequency than the first mode. Proportions of both modes are excited by the impulsive, flow induced aerodynamic forces, causing the masses to vibrate at a phase difference of 60° . Titze (1976) concluded that the masses would vibrate at the frequency of the first mode.

Ishizaka and Matsudaira (1972) and Ishizaka (1980), however, analytically showed that the two natural modes of their two-mass system (the same modes discussed by Titze (1976)), when coupled to the flow, degenerated into a single frequency midway between the two non-flow-coupled modes. They also showed that the synchronised phase delay was part of the phenomenon.

Minorsky (1962) derived a similar explanation to Ishizaka and Matsudaira's for two coupled, detuned, Van der Pol oscillators as part of a treatise on nonlinear oscillatory behaviour. He analytically demonstrated that the oscillators synchronised together at a frequency midway between the autonomous frequencies of each separate system, with a constant phase lag. It is probable that Ishizaka and Matsudaira's explanation is more likely than Titze's, although a definitive answer has not yet been found.

3. The glottal flow wave U_{tot} is skewed to the right with a steep falling slope, in contrast to the area waves which have symmetrical slopes. This is due to the inertia of the air in the glottis and the vocal tract. As the glottis opens, the transglottal pressure tries to accelerate the air through but the inertia of the air mass causes a time lag, resulting in a less steep initial slope. At closure the airway seals, causing the flow to fall rapidly to zero, producing a

steep negative slope.

4. The steep falling slope of the flow followed by the discontinuity at closure provides the energy input to the vocal tract. Just after the flow goes to zero, there is a large impulse in the speech wave indicating large instantaneous energy content (the observed 0.5 ms delay is due to the time taken for the pressure wave to traverse from the glottis to the lips). After the initial excitation at closure, the energy of the pressure (SPEECH) wave exponentially decays as the impedances in the vocal tract attenuate the signal, since the tract is not excited again until the beginning of the next period.
5. The flow wave shows temporal detail not observed in the area waves due to the interaction between the supraglottal (vocal tract), the subglottal, and the glottal systems. This is because the flow is dependent on the transglottal pressure and the glottal impedance. The glottal impedance is approximately inversely proportional to the areas $A_{v_{tot}}$ and $A_{m_{tot}}$, which are nearly triangular when positive. If the transglottal pressure is constant, as in a simple model with fixed sub and supraglottal pressures, the glottal flow is also nearly triangular. However, the vocal tract, with its varying areas and lengths reflects pressure waves back toward the glottis, causing the supraglottal, and consequently the transglottal pressure to vary. This variation is reflected in the flow but not in the area waves because of the inertia of the tissue masses.
6. The fundamental pitch, marked at the excitation points in the speech, is 124 Hz. This is within the expected 80 to 160 Hz range for normal adult males.
7. The maximum current is about 375 cm³/s, which is within the normal range for phonation.

4.2.2 Velocity Wave

Fig 4.2 plots (i) the medial edge displacements of left and right vocal cords and (ii) the velocity of the flow (VEL) and its relationship to the area A_{vtot}

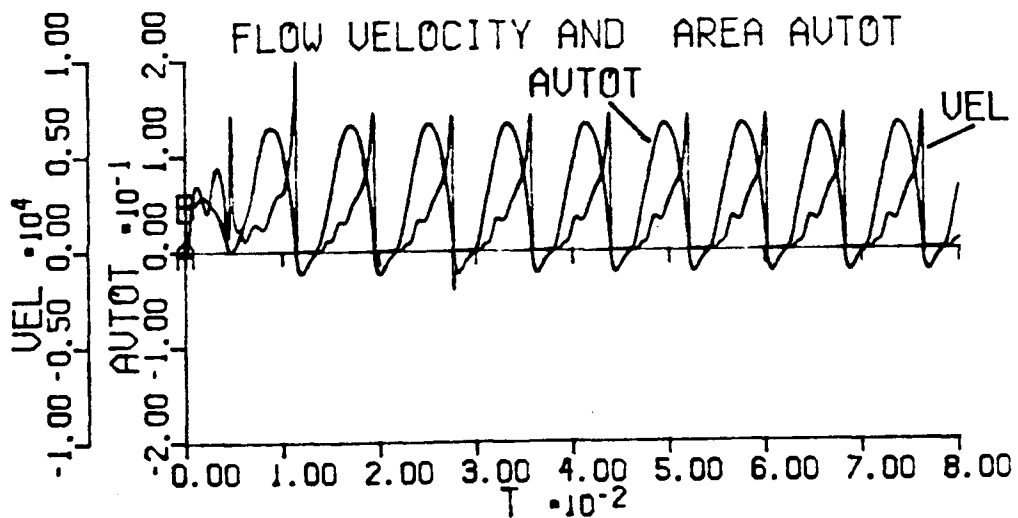
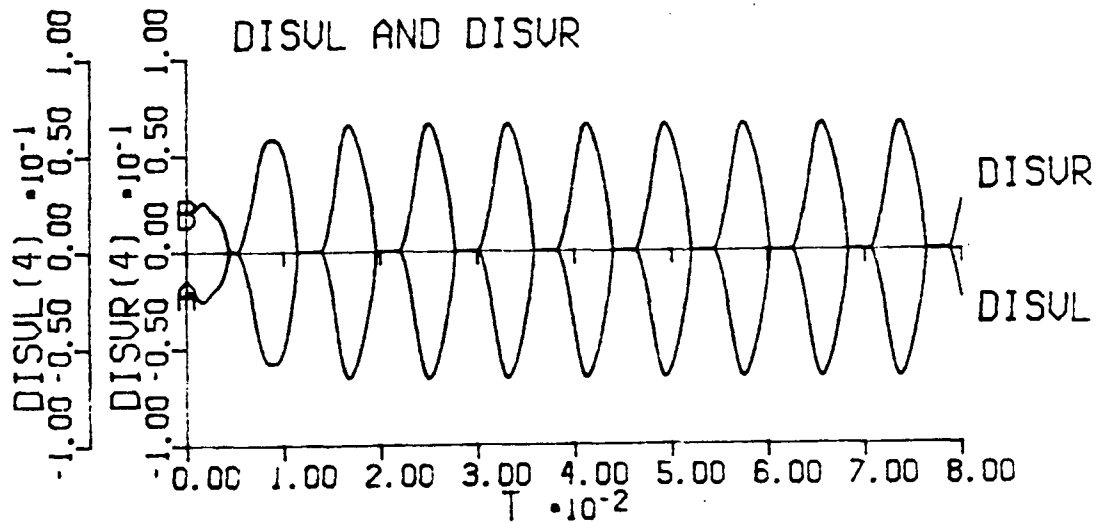


Fig4.2 (i) displacements and (ii) flow velocity and area, for normal fold

Points to note are:

1. Since the vocal cords are symmetric, the phase and magnitude of left and right cord displacements are also symmetrical.

2. The flow velocity has a saw-tooth shape due to closure of the cords. When both the lower and upper margins open, the velocity begins to increase along with the displacements and the flow, but lags the displaced area due to inertial (inductive) lag as v is a function of U_g ($v = U_g/A_v$). When the lower margins close, the velocity immediately falls to zero since the flow stops. Note that if the inductive loading from the tract is not large, and if collision does not occur, the flow velocity is approximately in phase with the displaced area. A Fourier series approximation is used in chapter five to show that a $1/2$ subharmonic is possible.

4.2.3 Phase and Energy Plots

Fig 4.3 is the phase plane plot of the lower vocalis margin of the right vocal cord. The system reaches a stable but asymmetrical limit cycle. The asymmetrical character is due to the changing nature of the system during collision.

Note that the trajectory crosses over itself, indicating that the system is not a simple autonomous system (Minorsky, 1962). In chapter five we present the argument that the crossed trajectory can occur only if the system is not second order ie third or higher order since the system is autonomous (not forced).

Fig 4.4 is a plot of the equivalent restoring force vs the displacement of the upper medial edge (assuming that the centre of gravity can enter the negative displacement region). The equivalent restoring force is a summative force which includes the tissue damping and stiffness forces, and the aerodynamic forces. It may be thought of as an acceleration vs displacement plot. Note that the forces are summed up as the total contribution from both the upper and lower masses. The displacement tracks the medial edge as it enters collision, but follows the continued movement of the centre of mass. By summing the forces on both masses, we can compare our

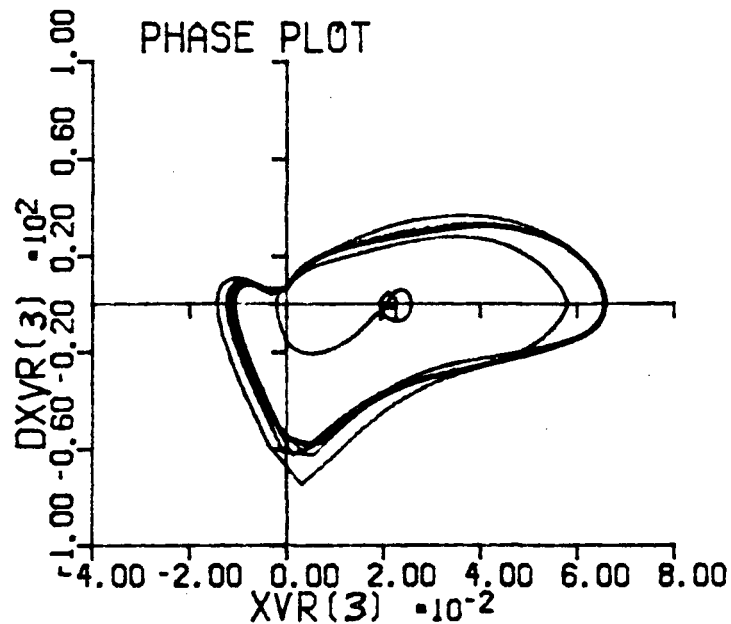


Fig 4.3 Velocity phase plot showing stable limit cycle

result with Titze (1983) where the force and displacement was measured midway between the upper and lower margins. Points to note are:

1. The upper left enclosed area is traversed anticlockwise, indicating that energy is dissipated when the displacement is negative. The lower right enclosed area is traversed clockwise, indicating energy absorption. When steady state oscillation is reached, both areas should be equal. (To prove this, force displacement graphs for all masses in the system need to be acquired and the areas summed, but this requires too much memory).

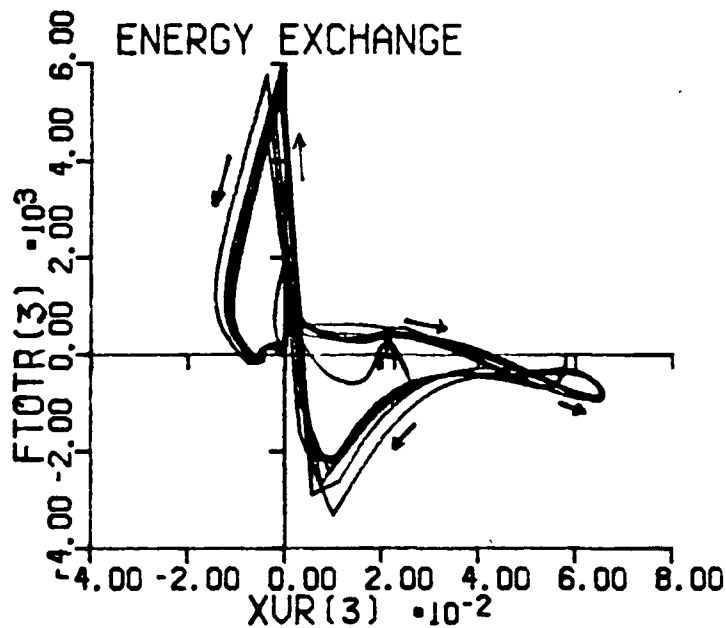


Fig 4.4 Equivalent restoring force vs displacement, showing energy exchange

2. The plot is very similar to that produced by Titze (1983) (see Fig 4.5) except that the loops of Fig 4.5 lie along the $-kx$ axis because Titze's model assumes linear stiffness. In Fig 4.4 however, a cubic stiffness characteristic suggesting a $-kx^3$ path is used, which causes a more upright "figure-8".
3. The two figures also differ when the displacement is negative because in the proposed model the mass is allowed to collide with its opposing margin. The extra spring forces introduced to simulate collision, and the impulsive subglottal pressure P_s result in a force during collision much larger than in Titze's case, which does not model the collision.

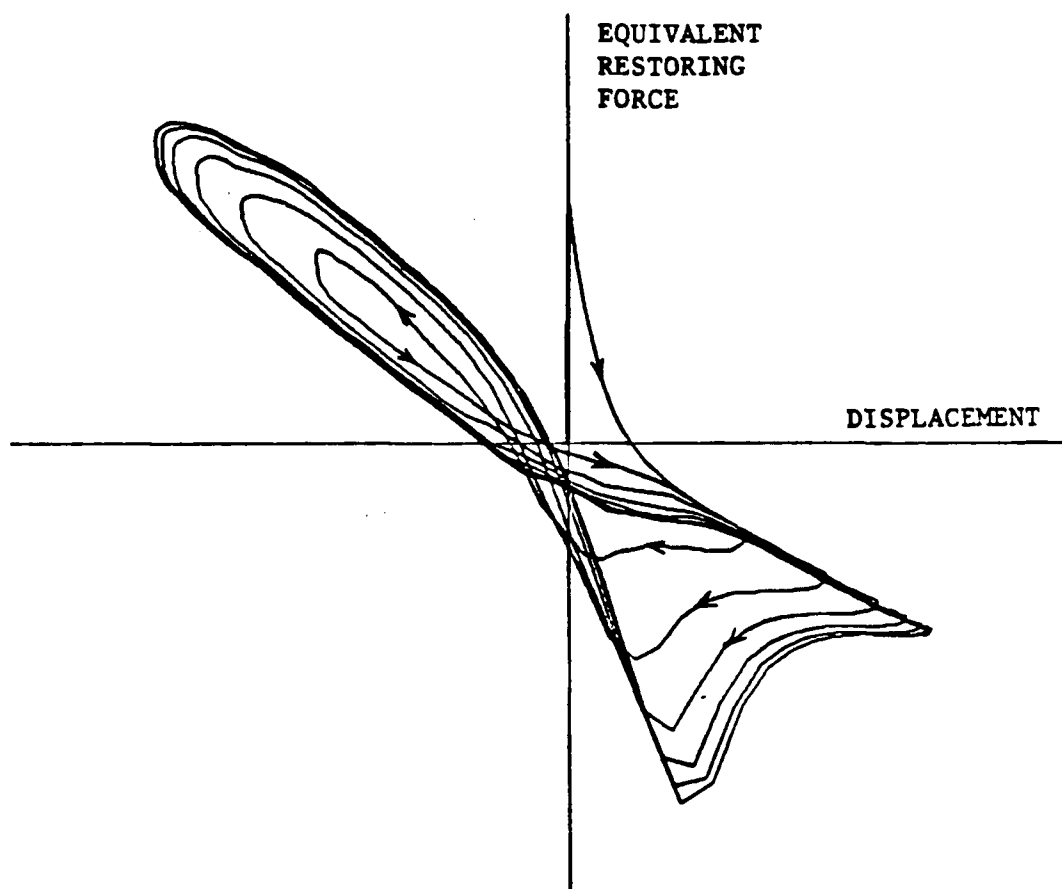


Fig 4.5 Equivalent restoring force. After Titze, (1983)

4.2.4 Force and Displacement

A time plot of the restoring force and the displacements reveals the changes occurring as the folds oscillate (see Fig 4.6). There are three regions of special note, and these are

1. Just prior to the lower margin closing (point (a)), the Bernoulli sucking force is large and negative, and with additional negative forces due to vertical phasing and inertial vocal tract loading the sum force produces a large,

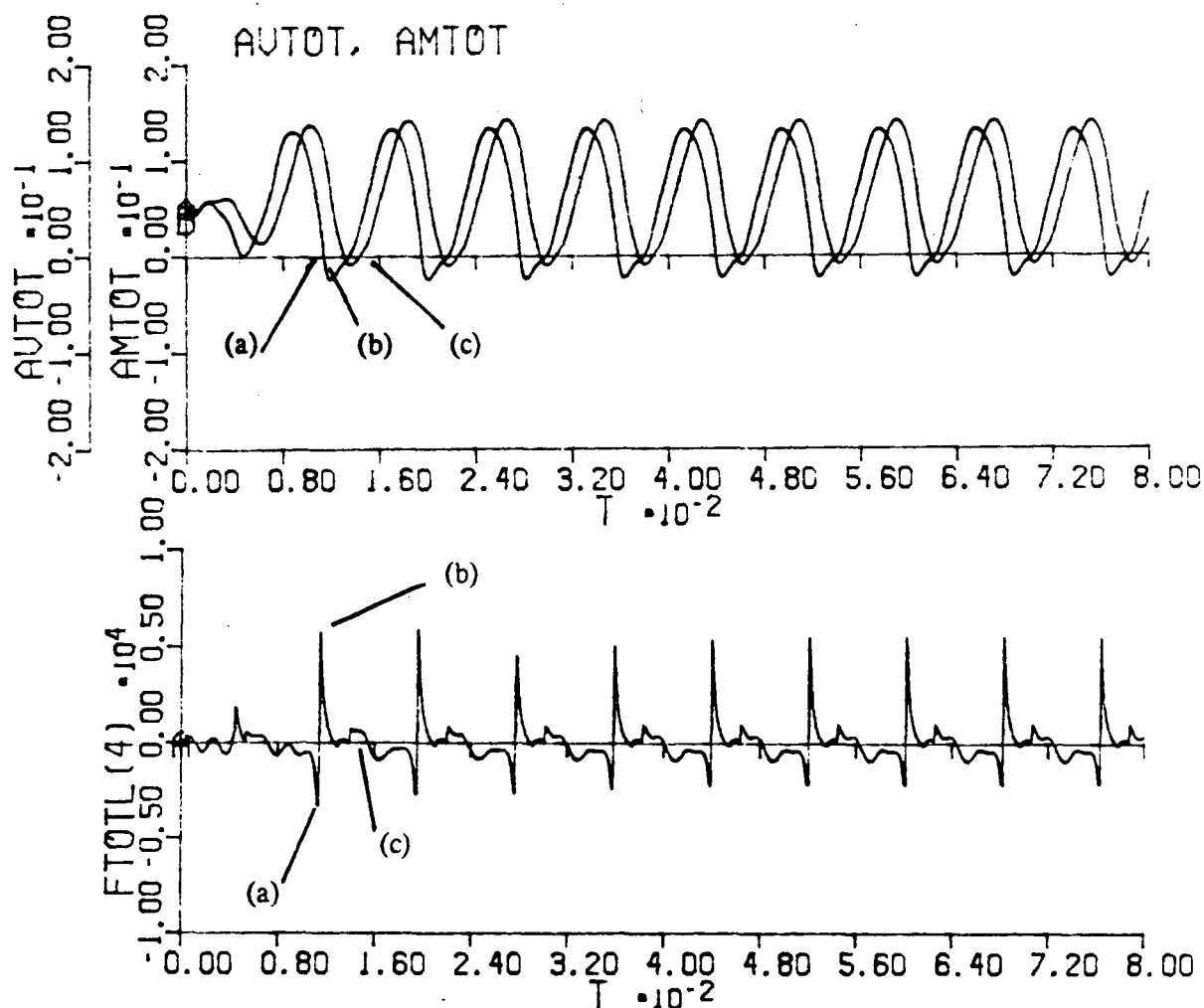


Fig 4.6 Areas and equivalent restoring force as functions of time

impulsive, negative force.

2. Just after closure of the lower margin (b), the vocalis mass is subjected to a buildup in pressure due to P_s , the subglottal pressure. P_s is momentarily very large (Koike, 1980) and consequently the aerodynamic force, proportional to $(P_s - P_0)/2$, increases to a very large value. At this time P_0 , the pressure just above the glottis, is negative due to air mass inertia (although its magnitude is relatively insignificant in comparison to P_s). The negative P_0 aids in bringing the upper margins together.

3. At point (c) the upper margin has just left the collision, and the air begins to flow again. At this time, the lower margin is wider than the upper margin producing positive vertical phasing and loading forces. Consequently the negative Bernoulli force is counteracted by these positive forces and the resultant force is much less negative, and possibly positive depending on the relative magnitudes.

4.2.5 Perturbation Analysis of the Normal Wave

Pitch period duration and largest-peak amplitude perturbation plots are presented in Fig 4.7

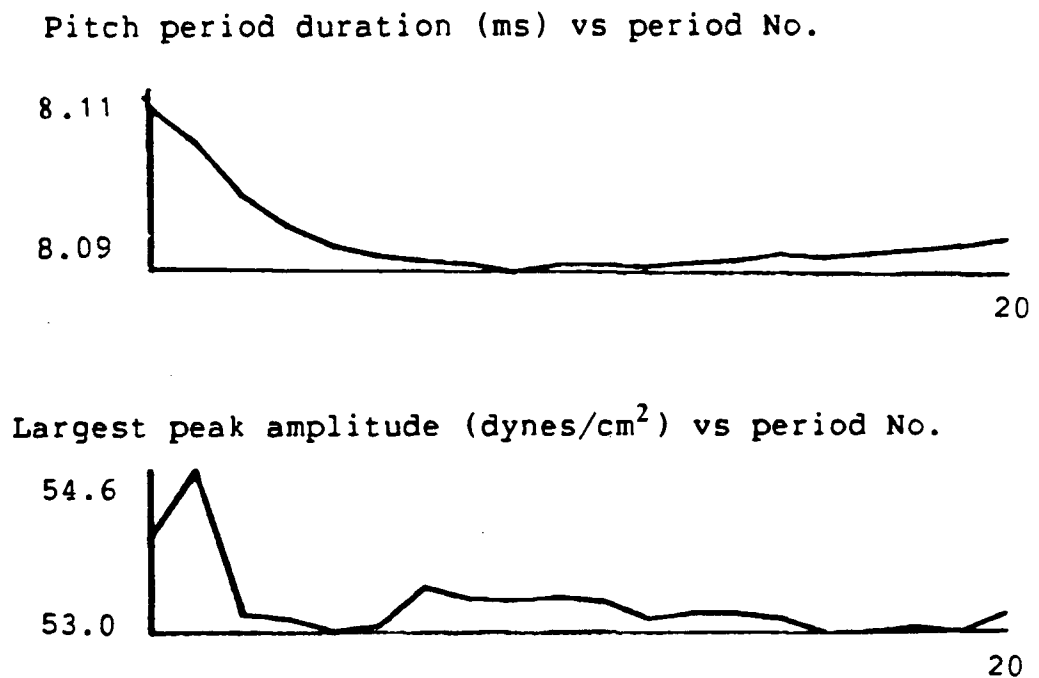


Fig 4.7 Perturbation plots, normal speech, from model

The pitch period duration versus period number plot appears relatively stable, indicating that very little jitter is present. The amplitude plot is a measure of the magnitude of the largest peak in each period. In this case the variation is extremely small (less than 0.2%).

The Harmonics-to-Noise ratio (HNR) is a measurement of the cycle to cycle consistency of the waveform, separating the periodic content from the assumed additive noise, assuming zero mean amplitude for the noise. (A derivation of the HNR appears in Appendix B). For the simulated normal case $\text{HNR} = 37.7 \text{ dB}$ which is very high when compared to real speech data (usually 15 to 25 dB for non-pathological voices. (Kojima et al, 1980)). This is not unexpected, however, since we have not modelled turbulence, which would contribute a significant random noise component.

4.2.6 Comparisons With Other Data

Fig 4.8 is a plot of the speech, flow and area signals from the Ishizaka-Flanagan model for the vowel /a/. In comparison with Fig 4.1 the results are very similar, although the fundamental frequencies are different (124 Hz for Fig 4.1 vs 160 Hz). The area and flow waves are also smaller than in Fig 4.8.

Fig 4.9 is from Davis (1976) using a patient with a normal voice. Both the speech wave (A) and the glottal volume velocity (flow, B) show very similar results to those obtained by the Ishizaka-Flanagan and the proposed models. The flow wave in Fig 4.9 was obtained by inverse filtering the speech and then low pass filtering the resulting residual signal.

Recordings from a subject with normal vocal cords obtained from Vancouver General Hospital are presented in Fig 4.10. The first plot is of 128 msec of speech for the vowel /a/, and shows periodic excitation followed by a period of attenuation. Again the appearance of the signal is consistent with simulated speech. The ppd and lpa plots appear to have a random nature, but the variation lies within relatively small bounds. The resulting HNR is 22.3 dB, which is well within the expected range for normal cases.

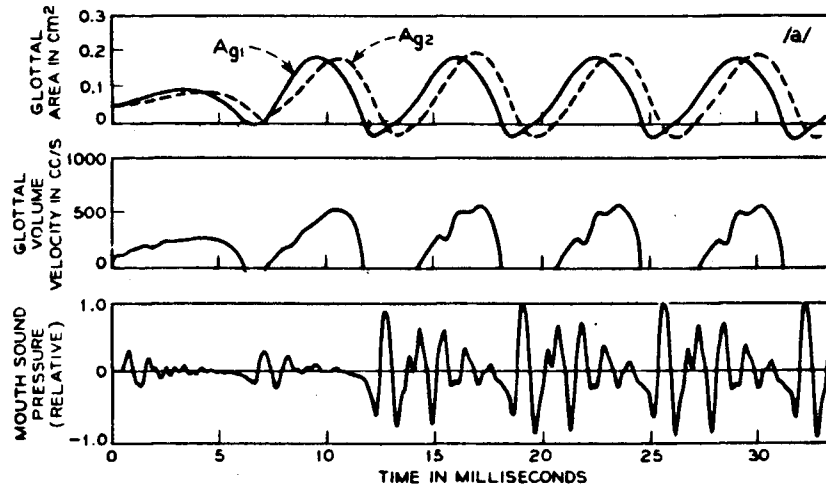


Fig 4.8 Results from the Ishizaka-Flanagan model using /a/. After Ishizaka and Flanagan, (1972)

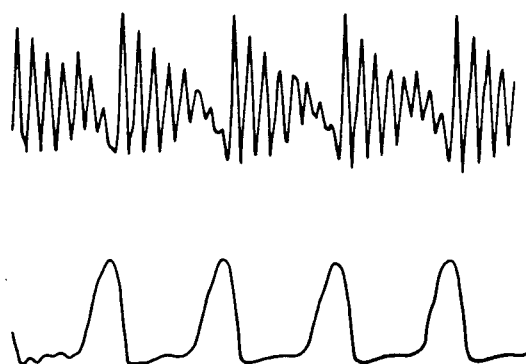
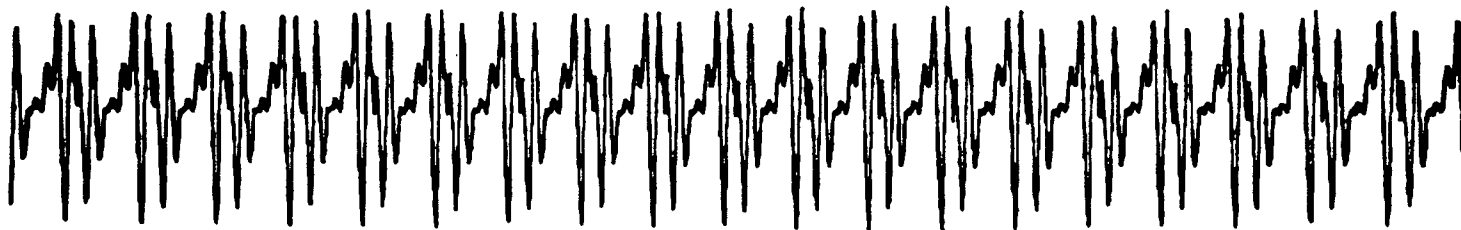
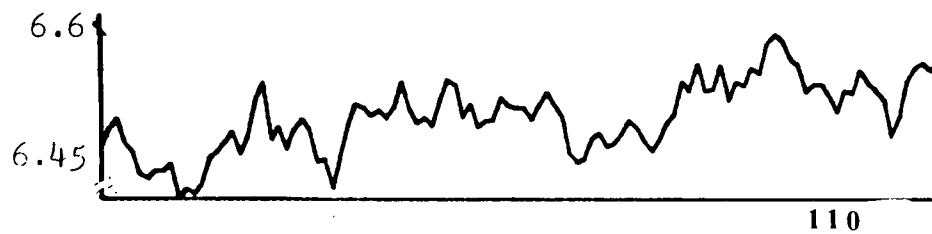


Fig4.9 Speech and flow, normal. After Davis, (1976)

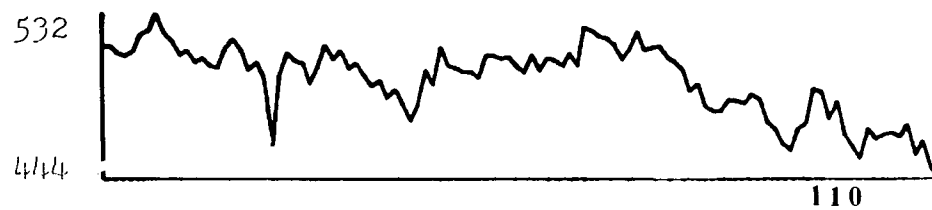
Fig4.10 Speech, ppd and lpa plots for normal case. From Vancouver General Hospital database.



Pitch period duration (ms) vs period No.



Largest peak amplitude (dynes/cm) vs period No.



4.3. The Pathological Vocal Fold

In this study we had great difficulty in associating specific parameter values with particular pathologies, since very little work has been performed prior to this time in collating the necessary data.

The approach taken here was to vary the mass, lateral stiffness and longitudinal constant tension values, either locally, or by changing an entire fold. The mass or stiffness was increased by a factor of 1.5 or 2.0, and the tension was varied between 150 (normal) and 25 (flaccid). The resulting speech waves were analysed using jitter, shimmer and harmonics-to-noise ratio measurements (Appendix B).

4.3.1 Unilateral Stiffness Change

In this series of simulations, the effect of increasing the lateral stiffness of an entire fold was investigated.

(1) The right fold's lateral stiffnesses k_v , k_c and k_m were increased by a factor of 1.5, with $T_{act} = 150$ (normal value).

- The resulting speech and flow waves presented in Fig 4.11 appear to be very regular, with little evidence of perturbation.
- Observation of the phase plot (Fig 4.12) for the third vocalis mass reveals a single limit cycle, and as a consequence, the ppd and lpa plots (Fig 4.13) are very stable.
- $HNR=35.9\text{dB}$, a very high value near that of the normal fold.

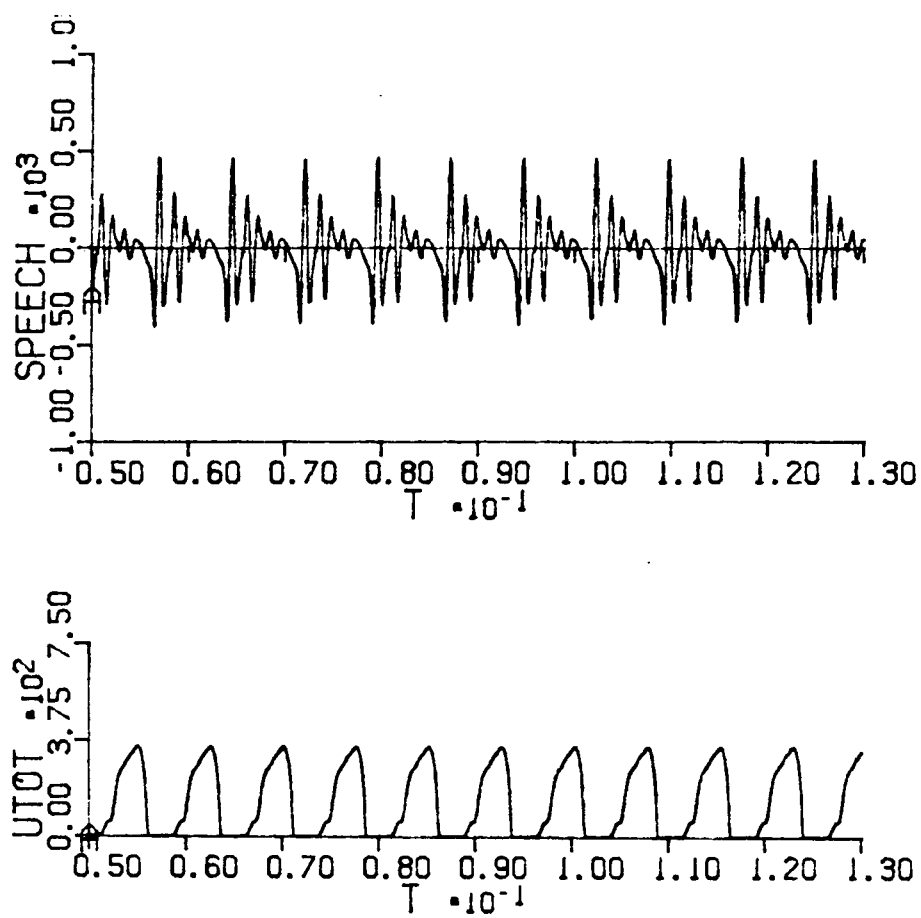


Fig4.11 Speech, Flow Waves for Unilateral stiffness change, $T_{\text{act}} = 150$

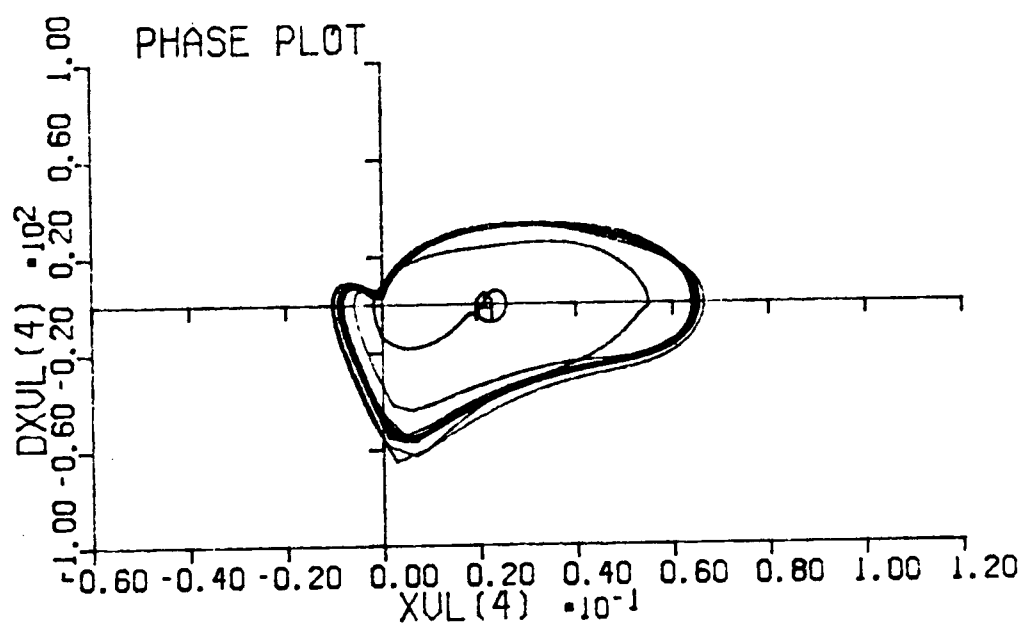


Fig4.12 Single Trajectory Limit Cycle Oscillation, Unilateral Stiffness Change, $T_{act} = 150$

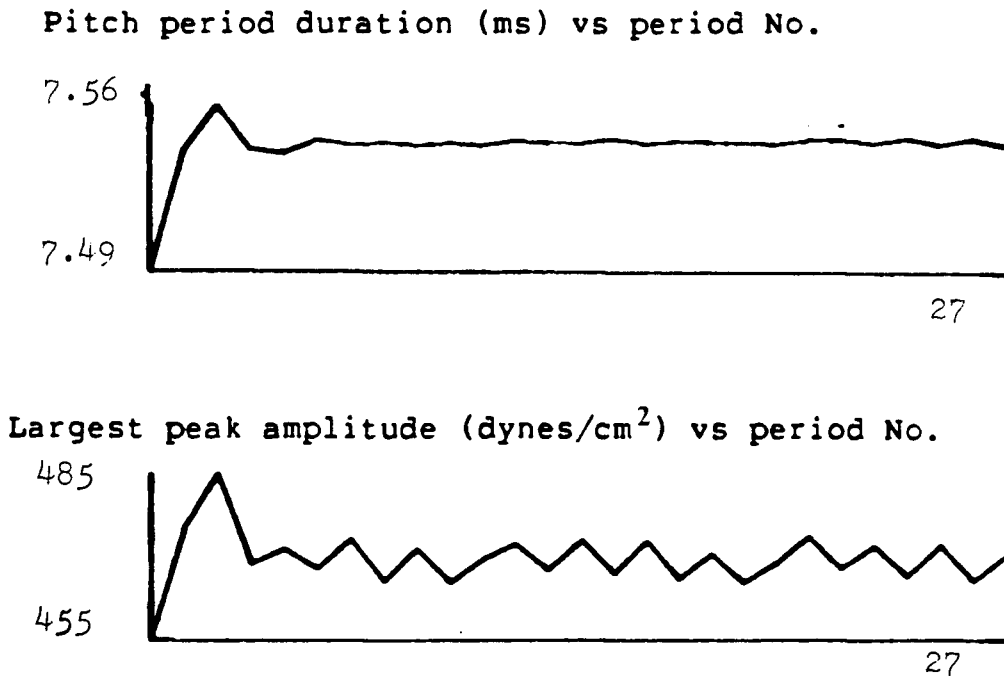


Fig4.13 ppd and lpa Plots for Unilateral Stiffness Change, $T_{act} = 150$

(2) Unilateral stiffness increase (as for (1)), $T_{act} = 25$

- The speech, flow and area waves appear in Fig4.14. From these waves, we can observe a slightly larger volume of flow, and significant changes in all waves on alternating cycles.
- The amplitude peaks in the speech signal are alternately attenuated, and the flow waves have a different shape at every second cycle. The cause of this phenomenon can be traced back to the area waves, where the amplitude reveals the presence of a $1/2$ subharmonic.
- The plot of left and right displacements (Fig4.15) shows that the stiffer right cord (B) is smaller in amplitude and leading the left cord, although the frequency of oscillation for each cord is exactly the same.
- The energy exchange plot (Fig 4.16) reveals two trajectories. Unequal areas are enclosed, indicating a net gain in energy for the mass pair examined. Since the oscillation for the system is stable, then the other masses in the

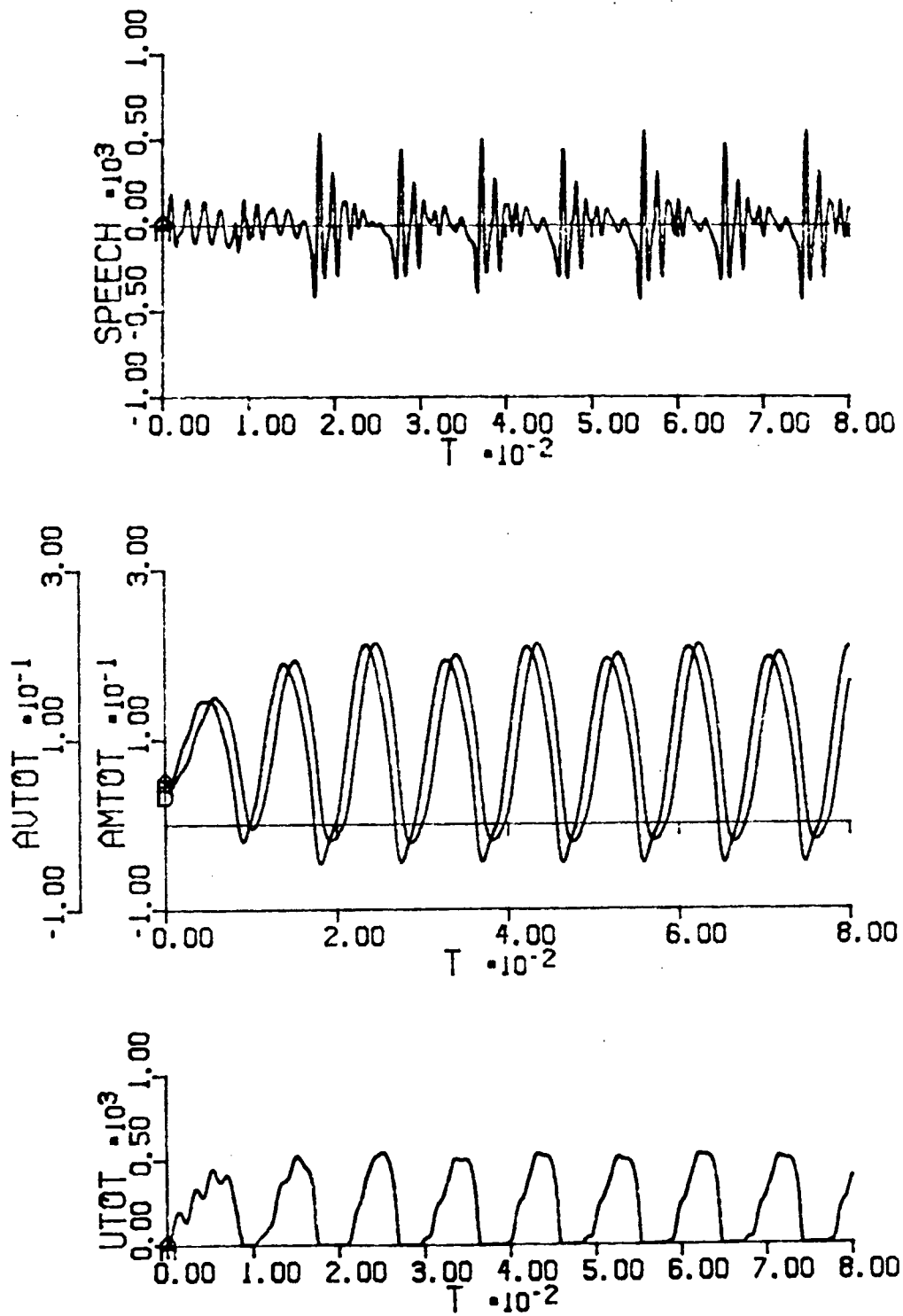


Fig4.14 speech, flow and area waves for unilateral stiffness change, $T_{\text{act}} = 25$

- system are dissipating energy to maintain energy equilibrium.
- The lpa plot is presented in Fig4.17. The variation in peak amplitude is clearly of a subharmonic nature.
 - HNR = 16.8 dB indicating a significant amount of perturbation.

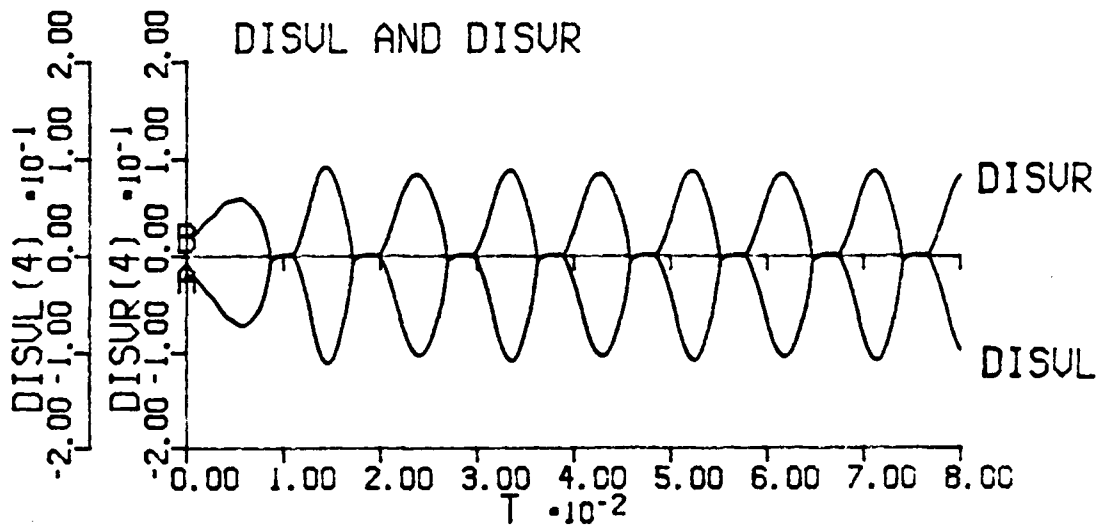


Fig4.15 Left and right displacements, unilateral stiffness change, $T_{act} = 25$

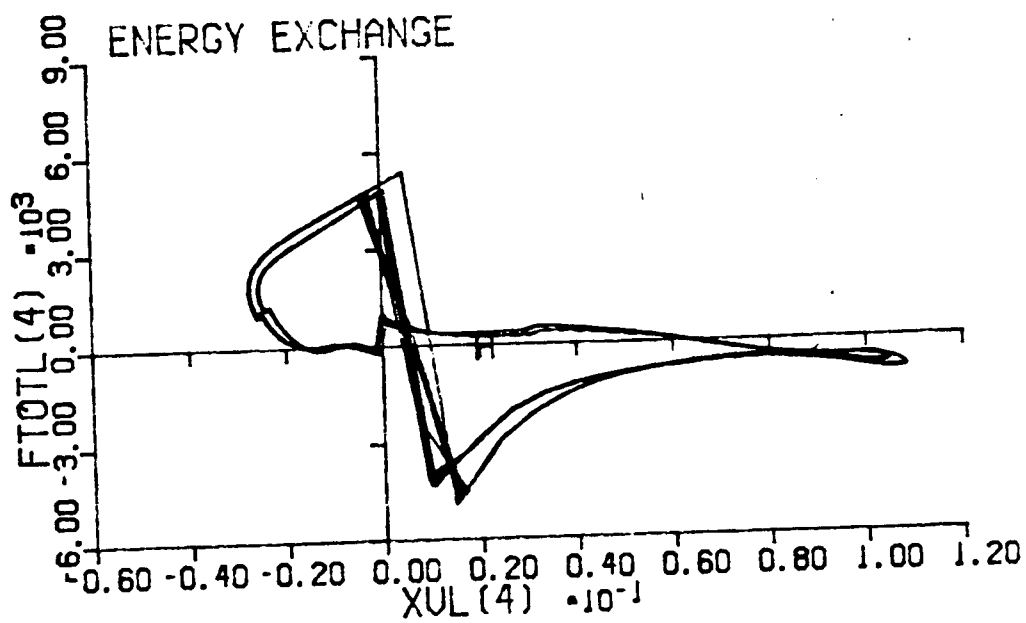


Fig4.16 Energy Exchange Plot, Unilateral Stiffness Change, $T_{act} = 25$

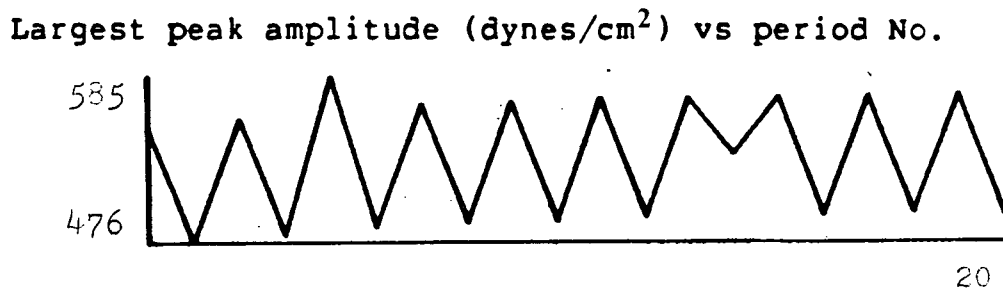


Fig4.17 lpa vs period #, Unilateral Stiffness Change, $T_{act} = 25$

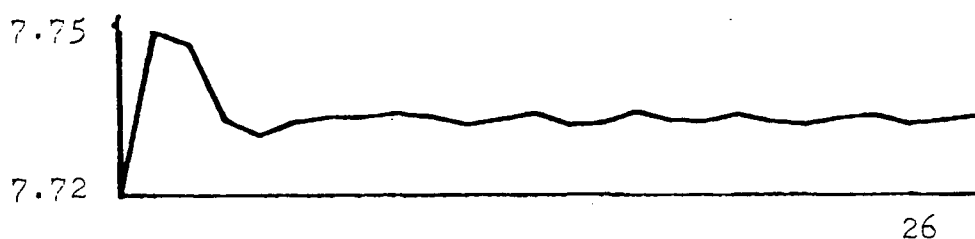
4.3.2 Localised Stiffness Changes

The lateral stiffnesses, k_v , k_c , and k_m , of the second and third elements of the right fold were increased by a factor of two.

(1) At $T_{act} = 150$

- The ppd and lpa plots show the speech signal to be very stable, although there is a slight, but perceptually insignificant 1/3 subharmonic in the lpa plot (Fig4.18)
- As a result of the steady tissue vibration, HNR = 36.1 dB, a very high value indicating that local stiffness changes have no effect at this value of T_{act}

Pitch period duration (ms) vs period No.



Largest peak amplitude (dynes/cm²) vs period No.



Fig4.18 ppd, lpa plots for local stiffness change, $T_{act} = 150$

(2) $T_{act} = 25$

- The phase plot (Fig4.19) shows a relatively stable system experiencing some slight variations in the trajectory compared to Fig4.3.
- The ppd and lpa plots (Fig4.20) have small jitter and shimmer components which are irregular.
- $HNR = 18.6$ dB, indicating that the irregular components have a significant effect on the amount of noise.

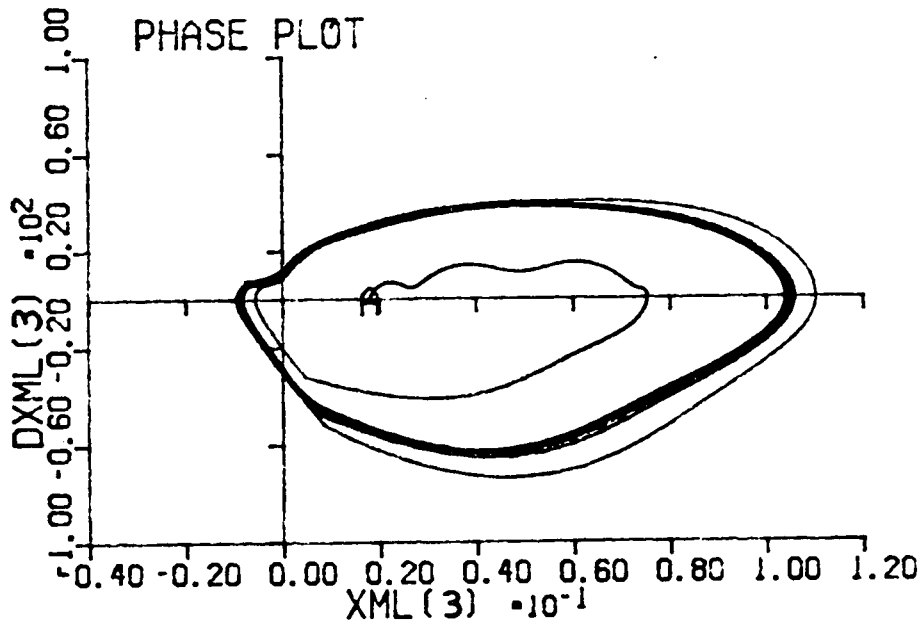
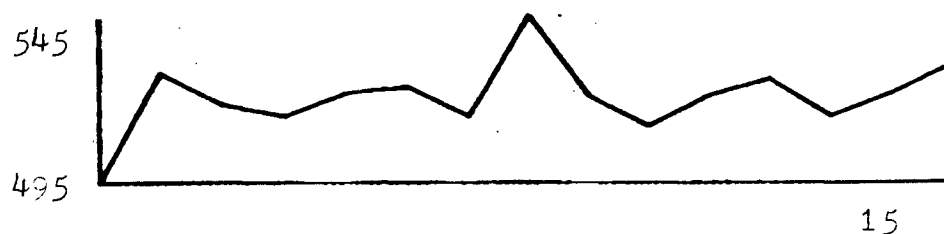


Fig4.19 Phase plot Localised Stiffness Change, $T_{act} = 25$

Pitch period duration (ms) vs period No.

Largest peak amplitude (dynes/cm²) vs period No.Fig4.20 ppd, lpa plots; Localised Stiffness Change, $T_{act} = 25$

4.3.3 Discussion of Results from Stiffness Changes

It appears from this set of simulations that stiffness changes are able to excite perturbations, but only when the constant tension T_{act} is very low. To determine whether low T_{act} or an asymmetrical change in stiffness produces the perturbations, a simulation was carried out with a normal configuration, but $T_{act} = 25$. The resulting flow and displacement waves (Fig4.21) and phase plot (Fig4.22) reveal the subharmonic nature to be still present, indicating that T_{act} is a controlling factor.

The pathology that the low T_{act} might represent is superior laryngeal nerve paralysis, a flaccid dysphonia in which the paralysed fold is characterised by a decrease in longitudinal tension. This may be bilateral or unilateral in its effect depending on where the lesion occurs. Gerratt and Hanson (1985) produced photoglottograms of patients with this paralysis and a regular perturbation was observed. They described the perturbation as possessing a regular component in which the third pulse was consistently larger. It is likely, based on the findings presented here, in

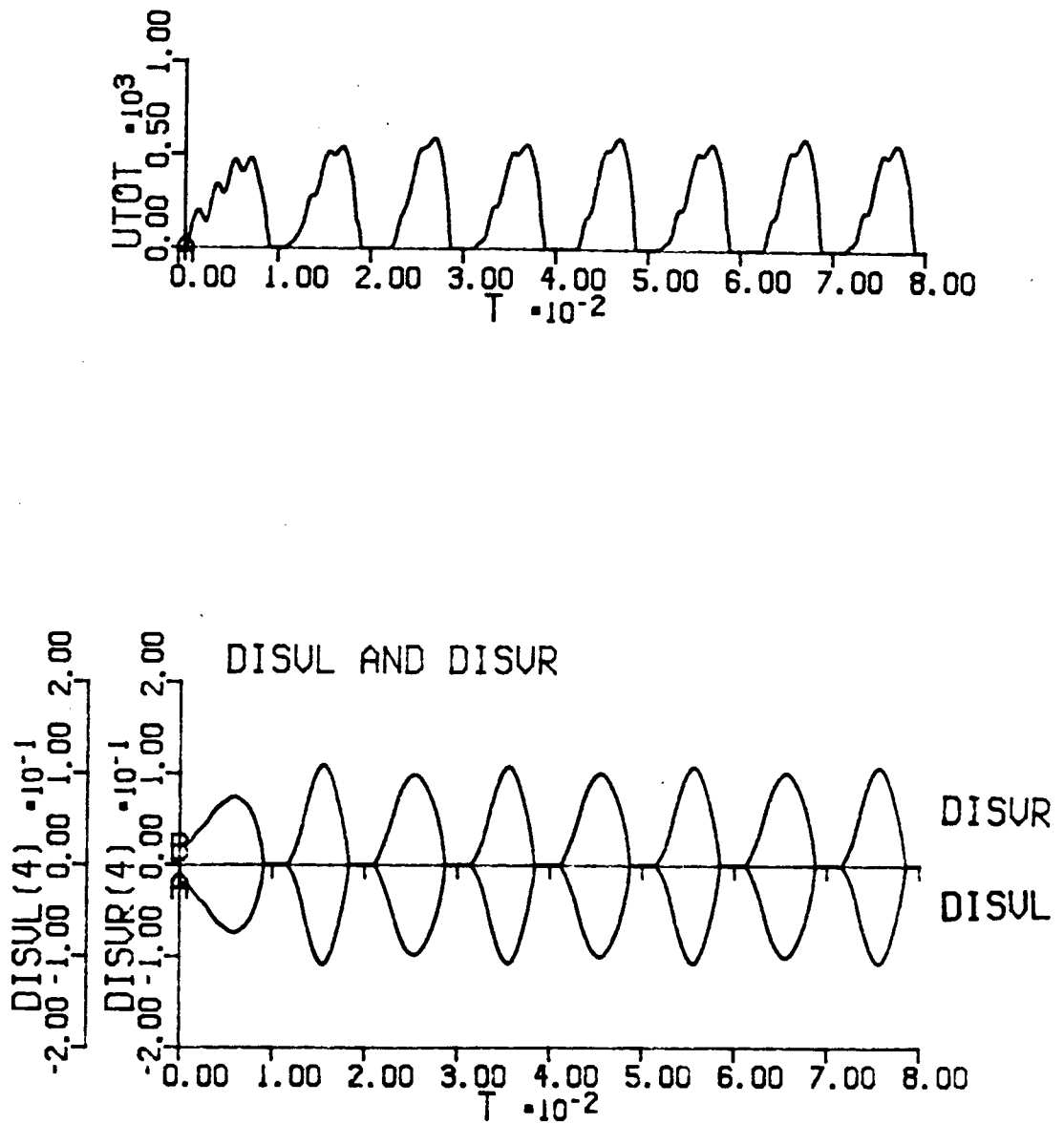


Fig4.21 Flow and Displacements, normal, $T_{act} = 25$

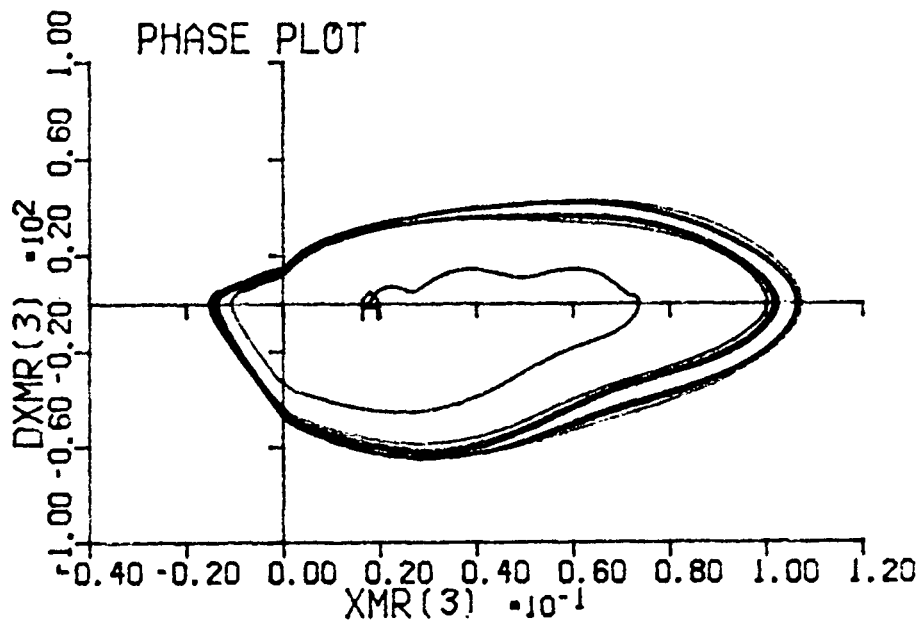


Fig4.22 Phase plot, normal, $T_{act} = 25$

chapter five, and by Isshiki and Ishizaka (1976) that it is a $1/3$ subharmonic (in our study the $1/3$ subharmonic did not manifest itself, although the dynamical equation proposed in chapter five indicates that it is possible).

The cause of the $1/2$ subharmonic is discussed in chapter five, but here the discussion is limited to its effect on the speech and flow waves. The $1/2$ subharmonic manifests itself by modulating the fundamental tissue displacement, causing the amplitude of vibration to be reduced on alternate cycles. The flow pulse, which is

proportional to the displacement, similarly alternates in magnitude and duration. The closing slope of the flow wave is steeper for the larger amplitude cycles, and consequently the speech receives more energy, being proportional to dU_g/dt . Thus, the modulation of the displacement amplitude manifests itself in the speech as alternately large and small peak amplitudes, and also in period duration, since the flow closure point will vary with the flow closing-slope.

By unilaterally increasing the stiffness, one would expect a smaller amplitude and increased frequency just for the affected cord. However, it is apparent from Fig 4.15 that while the stiffer cord's amplitude has decreased the same frequency is attained by both the left and right cords. This result agrees well with work by Tanabe et al. (1972) who experimented with live canine larynges and found the same phenomena. The frequency synchronisation occurs because the folds are coupled across the glottal gap by the common air flow, resulting in a mutually synchronised fundamental frequency between the expected left and right frequencies. We thus observe that the synchronisation between the upper and lower margins within a fold due to explicit spring coupling also occurs between the folds through the aerodynamic flow (although the coupling terms are obscure).

The energy exchange plot of Fig 4.16 has unequal enclosed areas indicating a net energy gain. This is not possible in an enclosed oscillatory system, so other masses in the system must be dissipating more energy than they receive so as to maintain the overall energy balance.

By localising the stiffness changes, the subharmonic behaviour is disturbed causing the perturbation to be more random, or perhaps a combination of $1/2$, $1/3$ and lower subharmonics (see chapter five).

It should be noted that HNR values have been reduced to half by decreasing the value of T_{act} . However, in comparison with HNR values measured by Kojima et al. (1980), 18 dB is at the lower range for nonpathological voices, and if judgement were made only on the results of the HNR analysis, the local stiffness changes would produce a borderline pathological case.

4.3.4 Unilateral Mass Changes

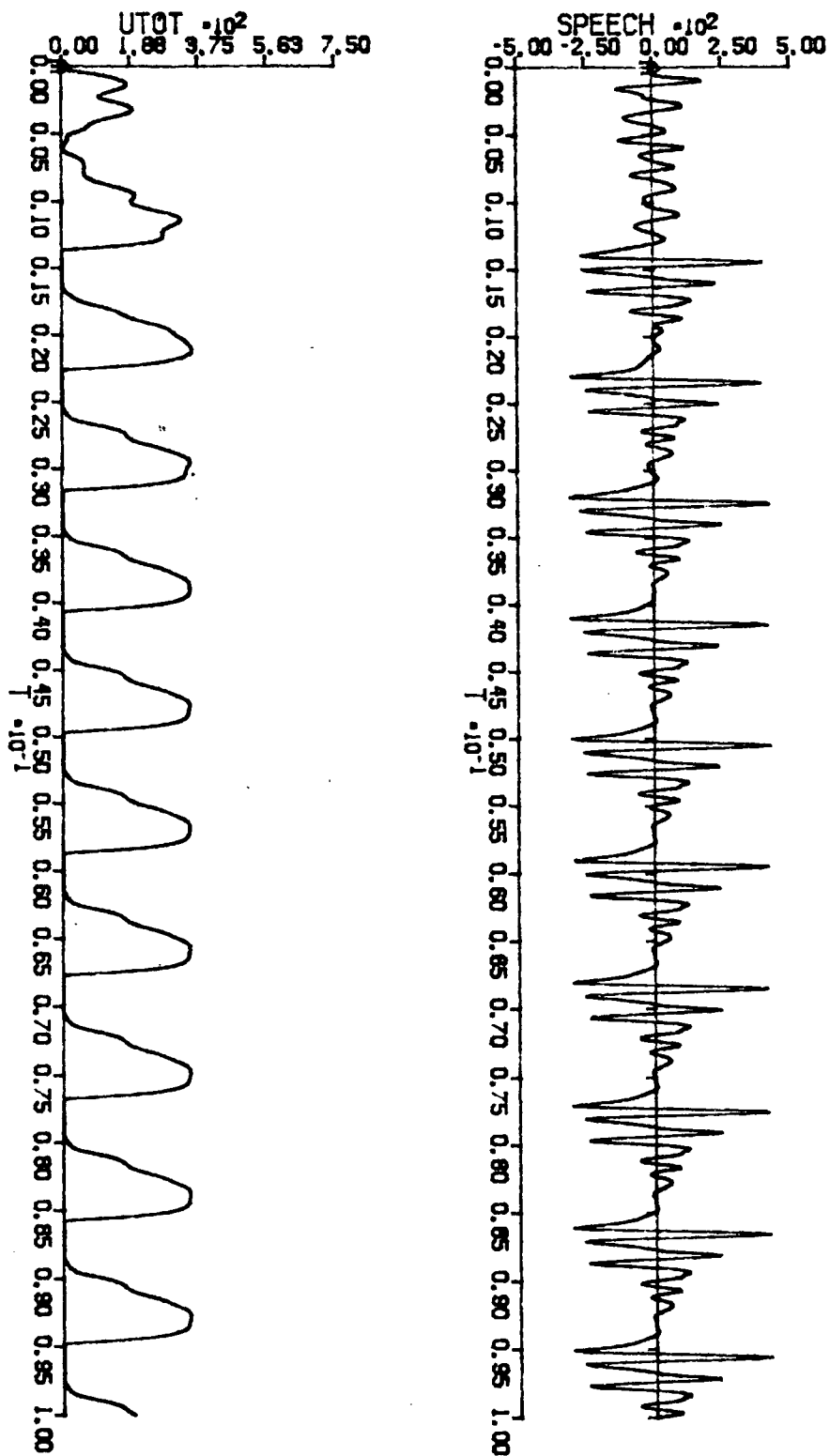
Changes in mass were simulated as a comparison to the stiffness changes. In this series of simulations the mass was varied unilaterally over the entire cord. In the following section localised changes are considered.

All masses on the right fold (both the vocalis and mucosal strings) were increased by a factor of two at T_{act} values of 150 and 25.

(1) $T_{act} = 150$

- The speech and flow waves in Fig 4.23 reveal a consistent, steady signal at 111 Hz with little evidence of perturbation.
- The displacement plot of Fig 4.24 shows the right fold mass element with a smaller amplitude than the left, and lagging. In comparison, increased stiffness caused the stiffer fold to lead the unaffected fold.
- The phase plot (Fig 4.25) shows a single steady state trajectory confirming the absence of jitter or shimmer.
- As a result, $HNR = 40.26$ dB

Fig4.23 Speech and Flow for Unilateral Mass Changes, $T_{act} = 150$



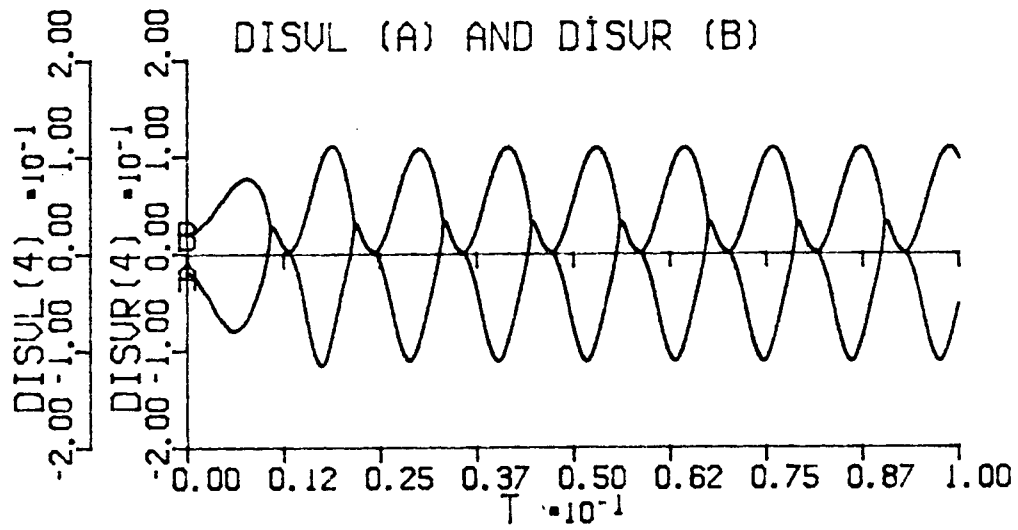


Fig4.24 Displacements for Unilateral Mass Changes, $T_{\text{act}} = 150$

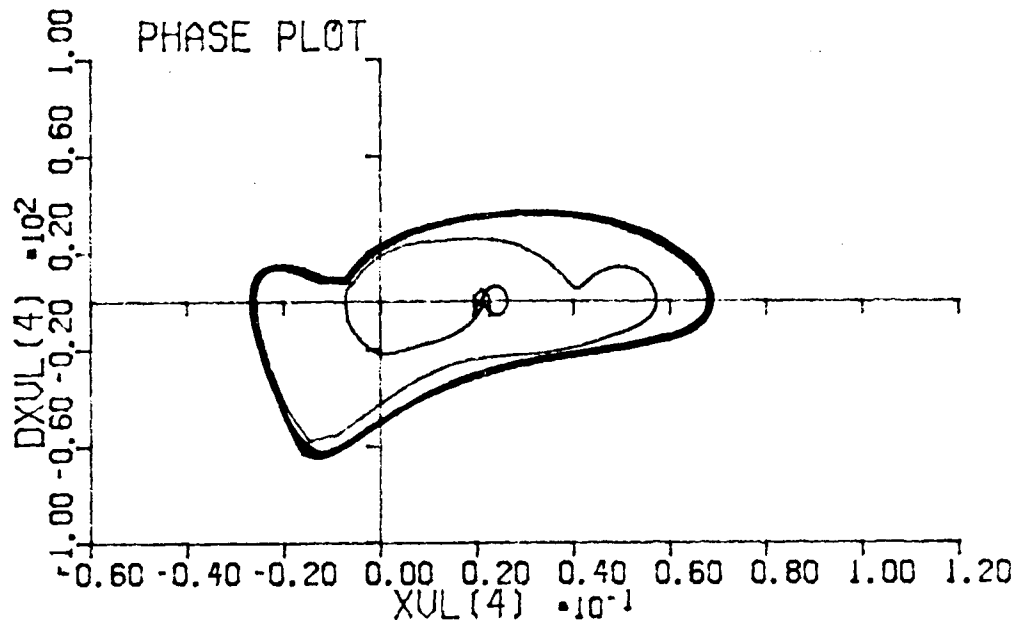


Fig4.25 Phase plot for Unilateral Mass Changes, $T_{act} = 150$

2. $T_{act} = 25$

- The speech, area, and flow waves of Fig4.26 do not reveal any observable perturbations.
- The frequency of oscillation is 87.7 Hz, significantly less than for $T_{act} = 150$. This is due to the relaxed longitudinal tension.
- The phase plot (Fig4.27), after an initial buildup, settles into a stable trajectory.
- $HNR = 40.8$ dB

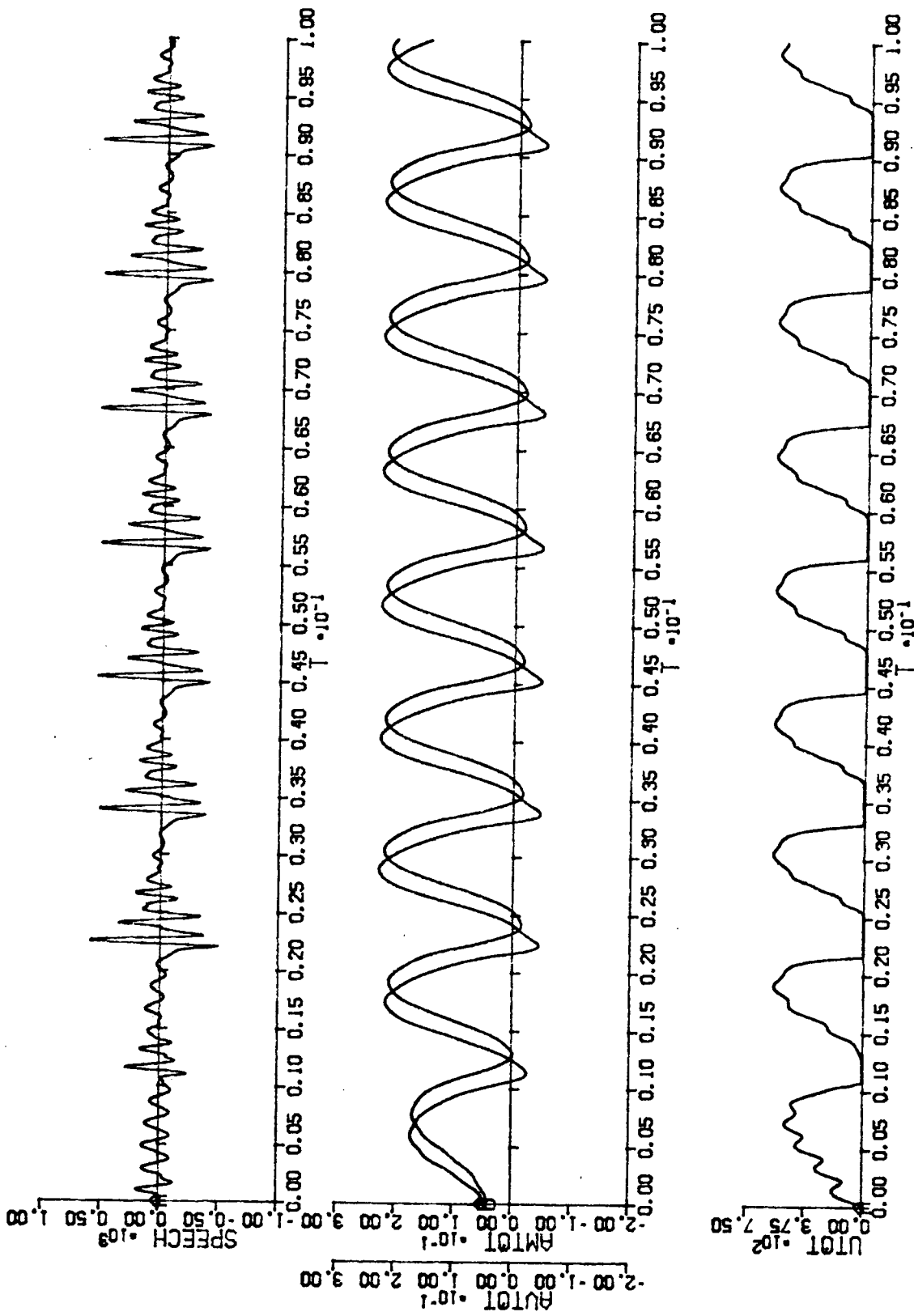


Fig4.26 Speech, flow and area waves for Unilateral Mass Changes, $T_{act} = 25$

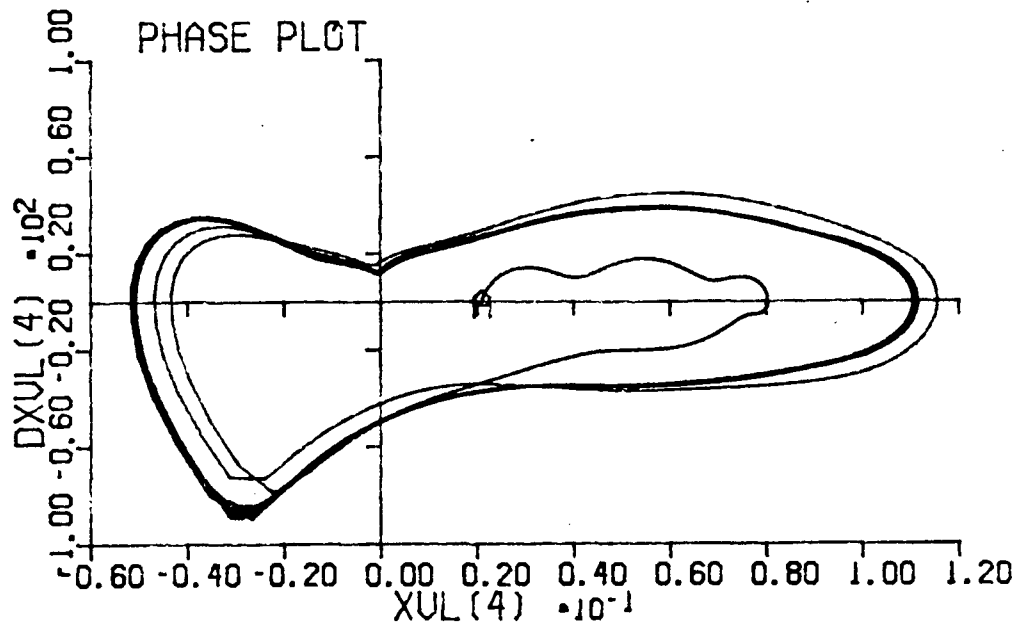


Fig4.27 Phase plot for Unilateral Mass Changes, $T_{act} = 25$

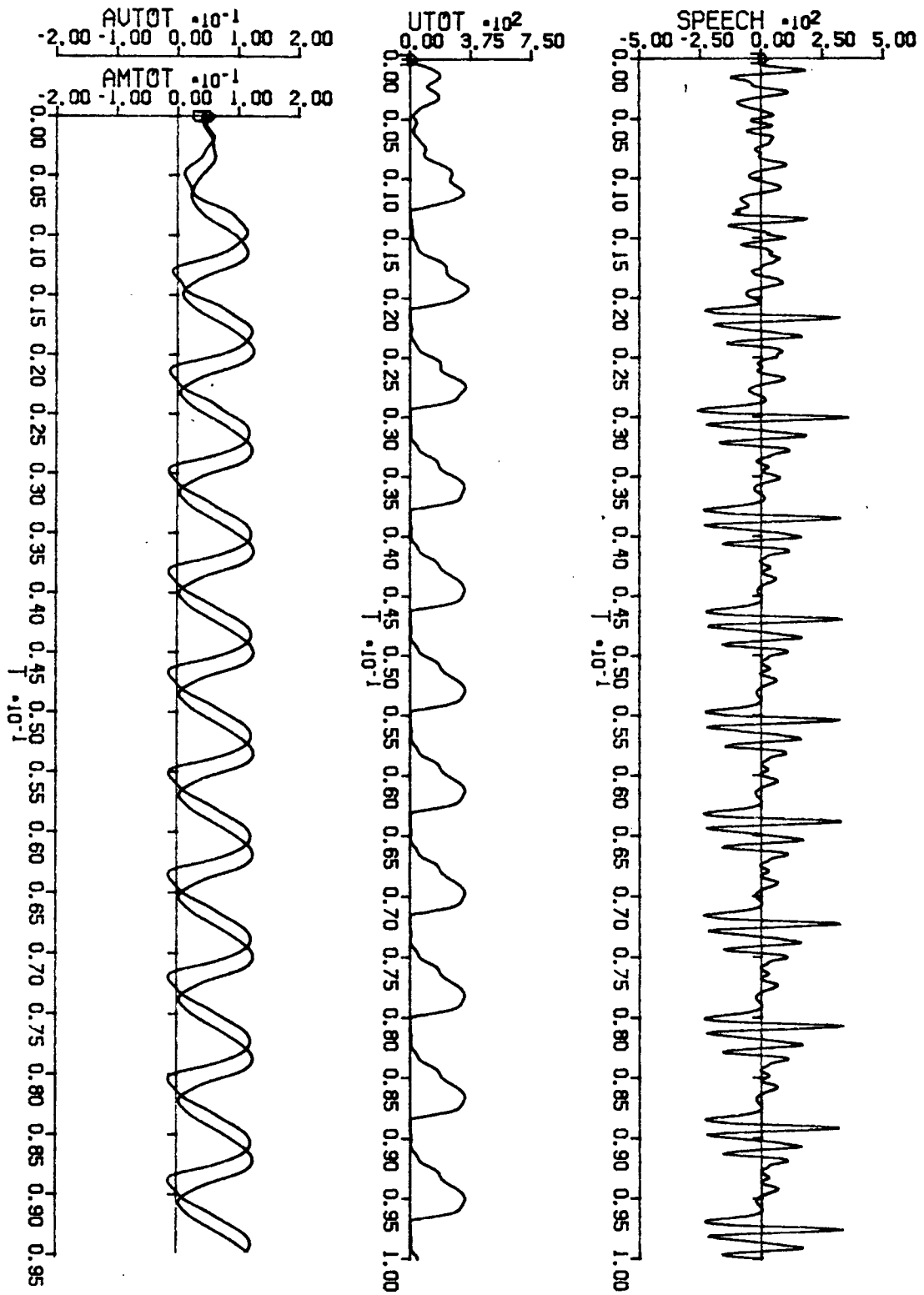
4.3.5 Localised Mass Changes

The second and third masses of the right vocalis and mucosal strings were increased by a factor of two for $T_{act} = 150$ and 25.

(1) $T_{act} = 150$.

- Speech, flow and area waves (Fig4.28) appear to be very stable.
- The phase plot (Fig4.29) shows a single stable trajectory.
- $HNR = 35.7$ dB

Fig4.28 Speech, flow and area waves for Localised Mass Changes, $T_{act} = 150$



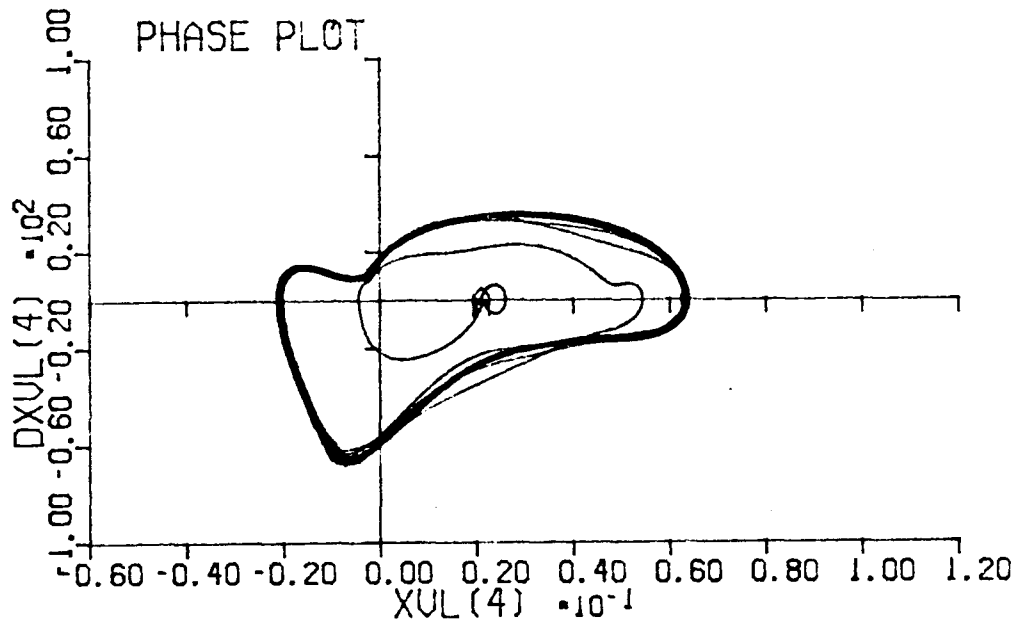
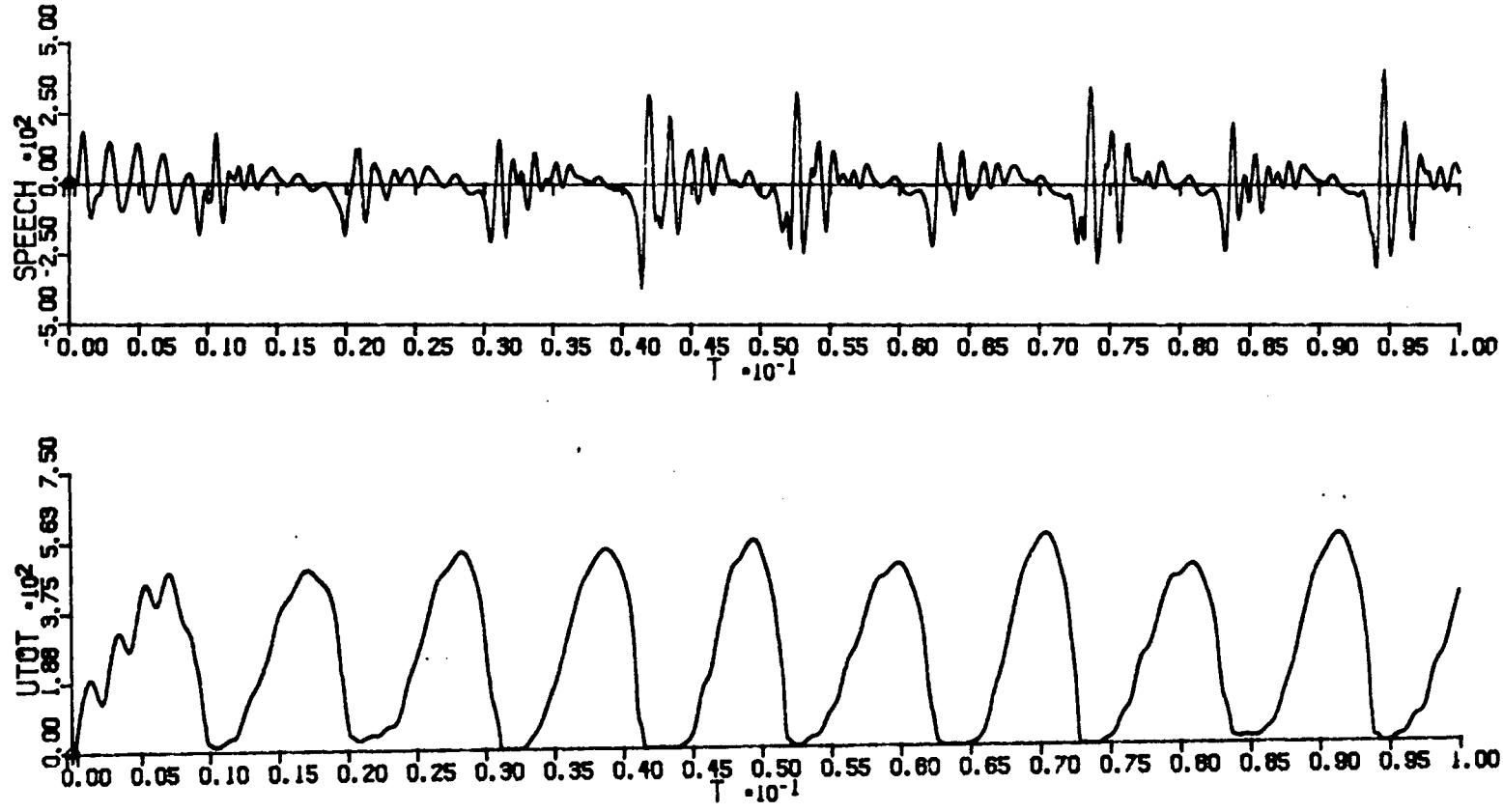


Fig4.29 Phase plot for Localised Mass Changes, $T_{act} = 150$

(2) $T_{act} = 25$

- The speech and flow waves reveal (Fig4.30) that the $1/2$ subharmonic is present, but there is an unstable random component affecting the waves.
- The displacement waves (Fig4.31) also have a random component in the amplitude, indicating that tissue displacement variation causes the jitter and shimmer in the speech.
- The phase plot (Fig4.32) of the third mass on the right fold shows a highly unstable trajectory.
- Fig4.33 and Fig4.34 shows energy exchange diagrams for both the left and right fold. The trajectory shape in each plot is different, and they both possess the random component that modulates the steady state trajectory. This

Fig4.30 Speech and flow for localised mass changes, $T_{act} = 25$



is surprising, since only the right fold has undergone local mass changes.

- Ppd and lpa plots in Fig4.35 again show a significant random jitter and shimmer component, with about 10 Hz variation in frequency, between 89 and 97 Hz
- As a result, $HNR = 7.9$ dB, a value that lies in the range for pathologies (Kojima et al, 1980)

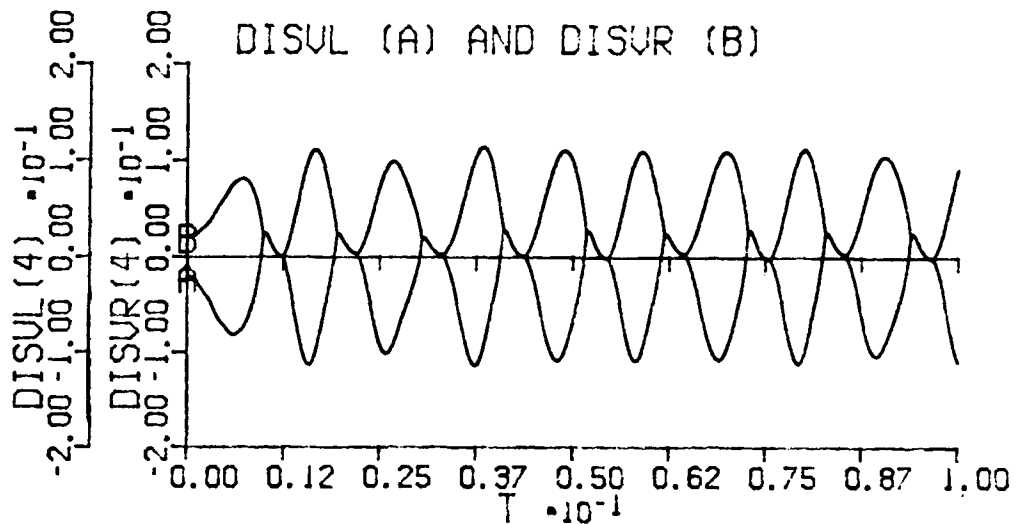


Fig4.31 Displacements for localised mass changes, $T_{act} = 25$

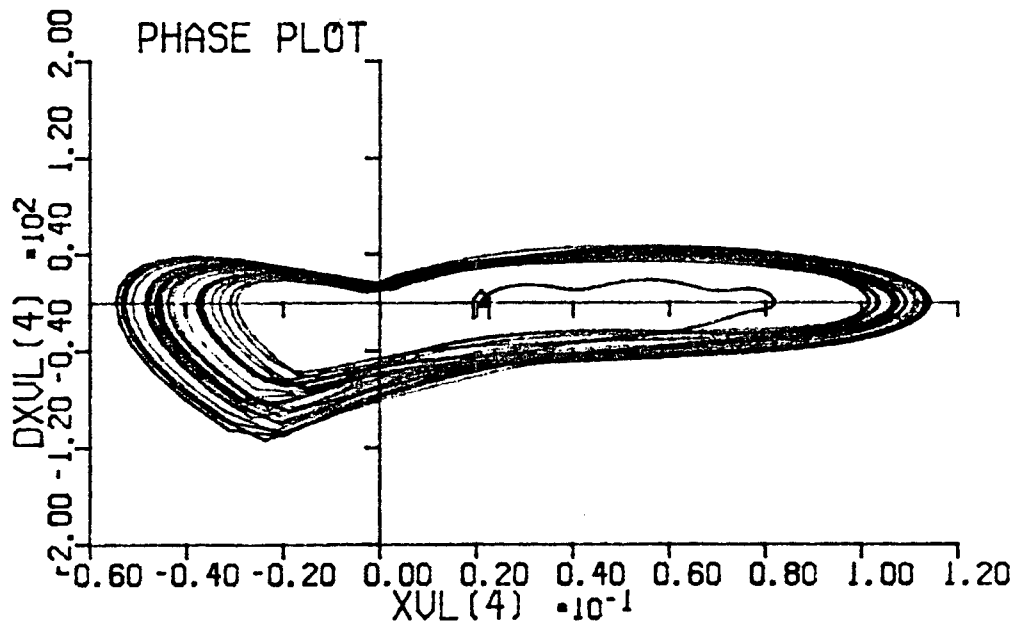


Fig4.32 Phase plot for localised mass changes, $T_{act} = 25$

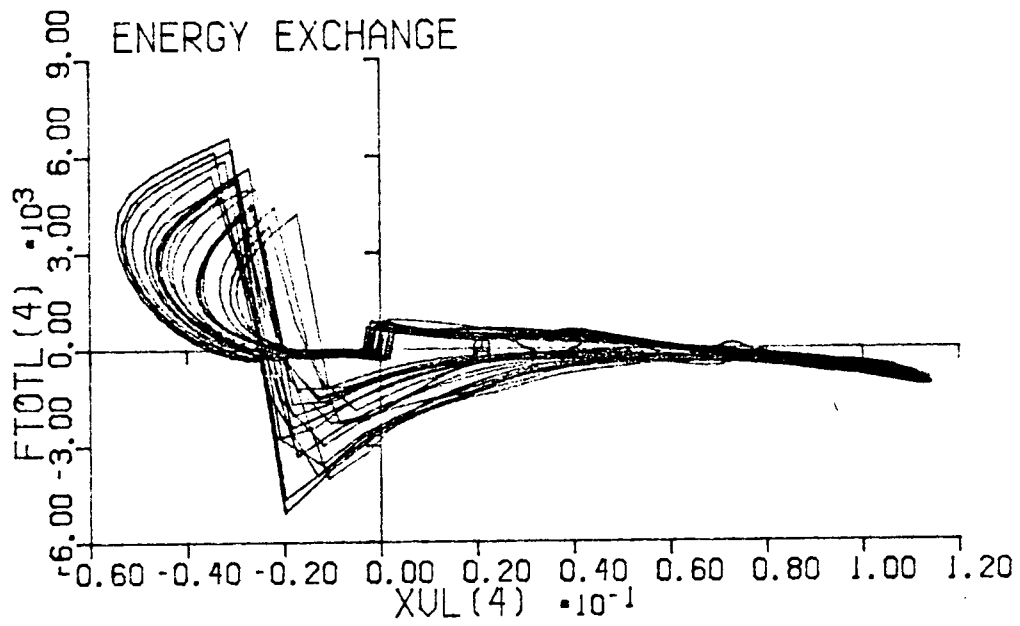


Fig4.33 Energy exchange plot for left fold, localised mass changes, $T_{act} = 25$

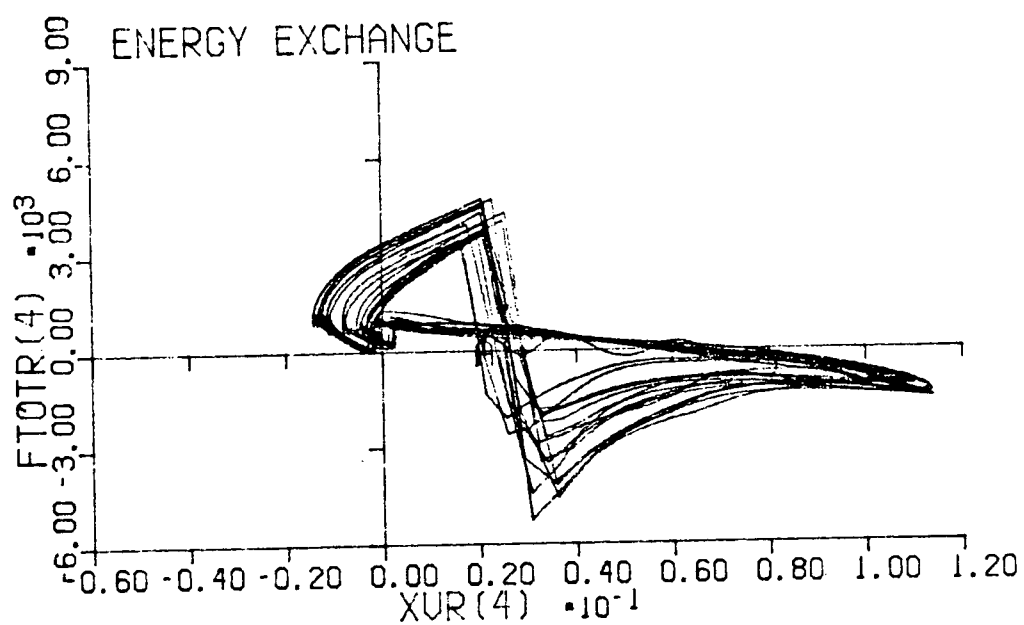


Fig4.34 Energy exchange plot for right fold, localised mass changes, $T_{\text{act}} = 25$

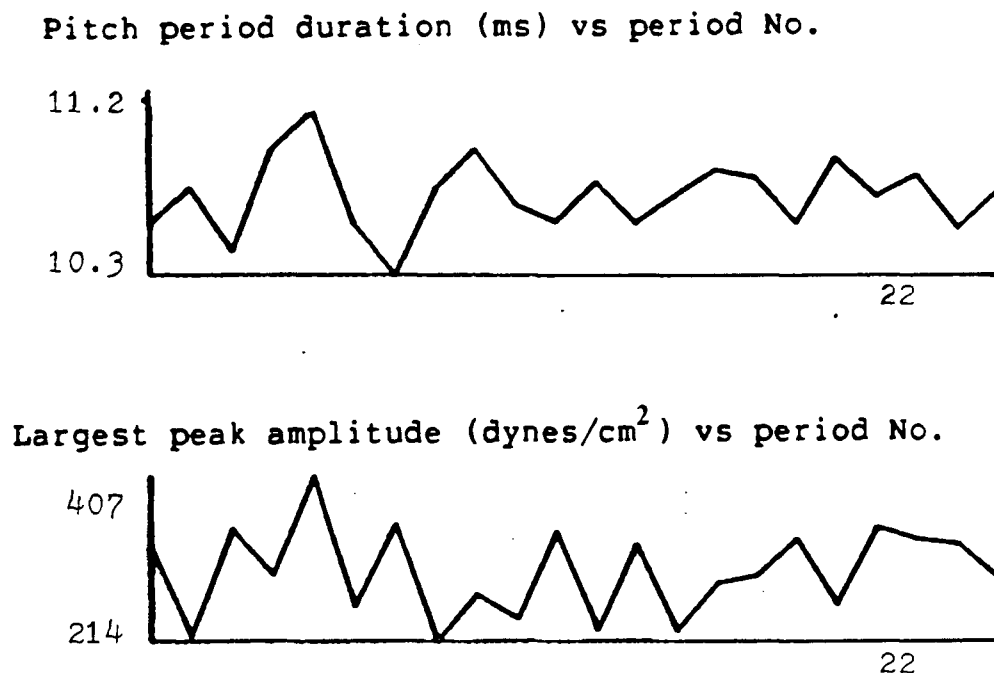


Fig 4.35, ppd and lpa plots, localised mass changes, $T_{act} = 25$

4.3.6 Discussion of Results from Mass Changes

As for the stiffness changes, no perturbations were observed for any variation in the masses for $T_{act} = 150$. The high value of T_{act} inhibited the appearance of any subharmonics or other phenomena, probably because T_{act} lay in a region of parameter space that did not allow subharmonics to manifest themselves.

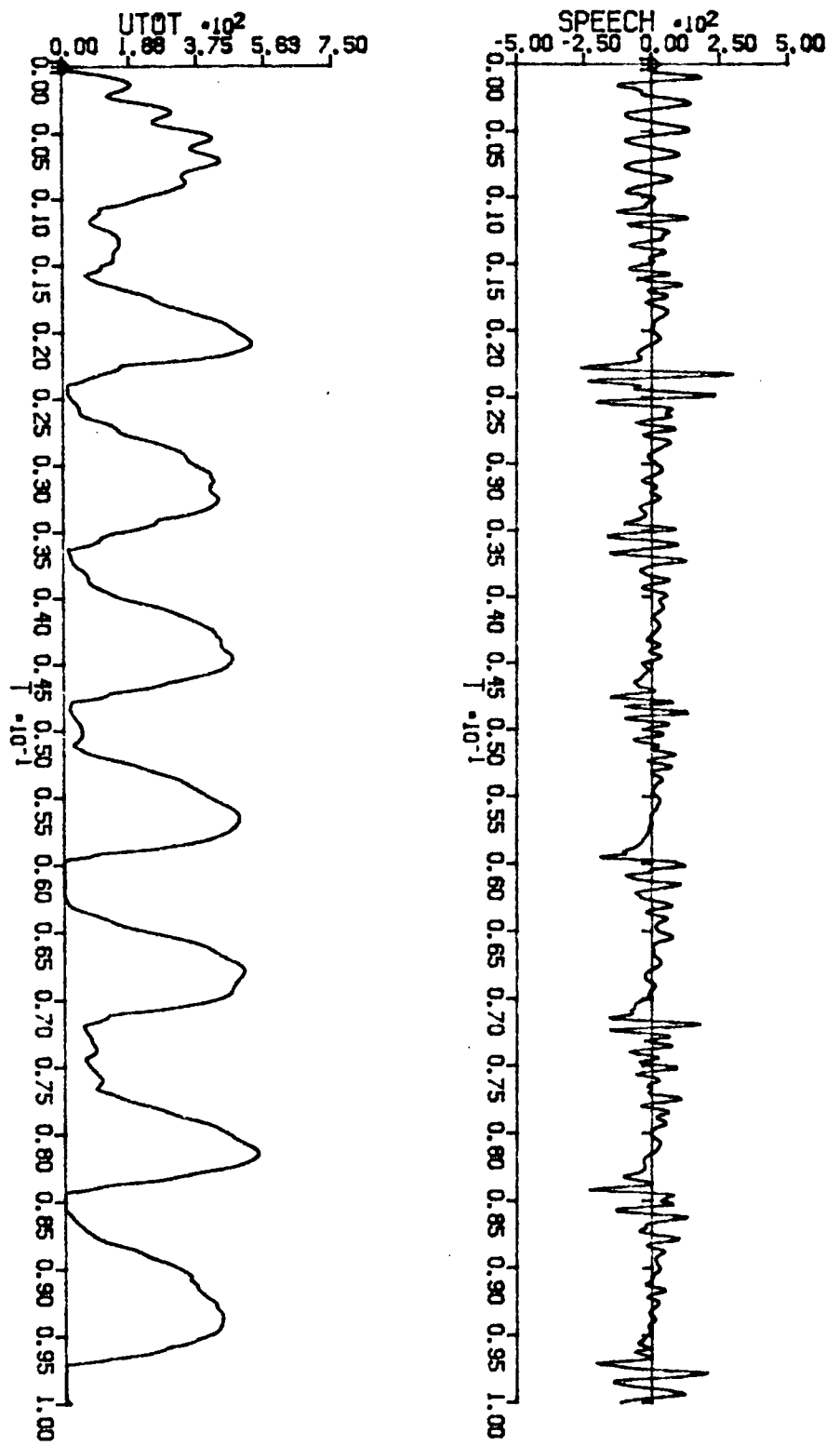
At $T_{act} = 25$ however, perturbations appeared for certain configurations. For nonlocalised mass changes, no perturbations resulted. Note, however, that a subharmonic appeared in the simulation of Fig 4.21 where there were no mass changes and a decrease in T_{act} . Clearly the increased masses have an effect on the presence of perturbation, as do the stiffnesses. The effects oppose each other, with unilateral

stiffness increases enhancing the subharmonic and unilateral mass increases inhibiting the subharmonic.

In the simulations using localised mass changes, the objective was to model cancer, which can be localised and invasive, and causes significant hoarseness. The local mass changes resulted in a phase trajectory with a large random component as well as the $1/2$ subharmonic. The combined effect significantly decreased the harmonic content of the speech and was more dramatic than for localised stiffness changes, which randomised the trajectory to a small extent but inhibited the $1/2$ subharmonic.

To determine whether the random components were due to the asymmetry of the model, local mass changes were made to both folds. The resulting speech and flow waves in Fig 4.36 show much the same results as for Fig 4.30, although the speech wave amplitude was reduced due to the shallow negative slope of the flow pulse. The energy diagram (Fig 4.37) still shows the random component, and as a result $HNR = 0.38$ dB, indicating very poor harmonic structure. Fig 4.38 displays the ppd and lpa plots, the frequency ranging from 73 to 98 Hz. This simulation would be analogous to vocal nodules, polyps or bilateral cancer, in which both folds are affected at the same location. Note, however that this simulation did not attempt to simulate the glottal chink, nor keep the nodule on the surface.

Fig4.36. Speech and flow for bilateral local mass changes, $T_{act} = 25$



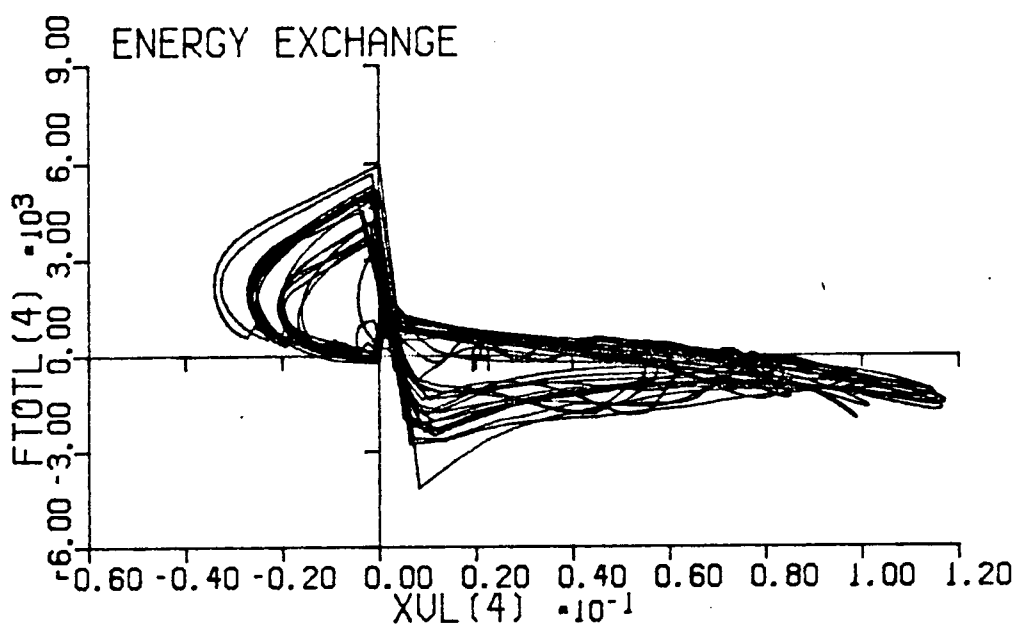
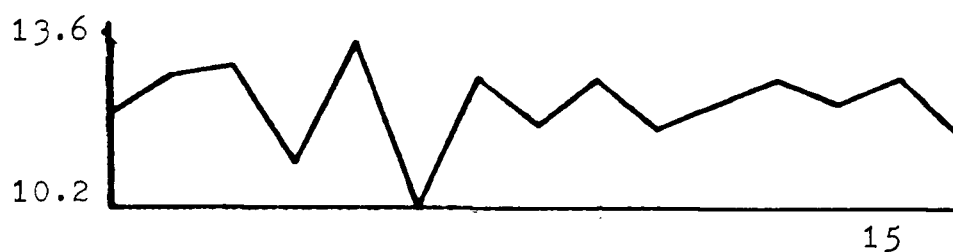


Fig4.37, Energy diagram for bilateral local mass changes, $T_{act} = 25$

Pitch period duration (ms) vs period No.



Largest peak amplitude (dynes/cm²) vs period No.

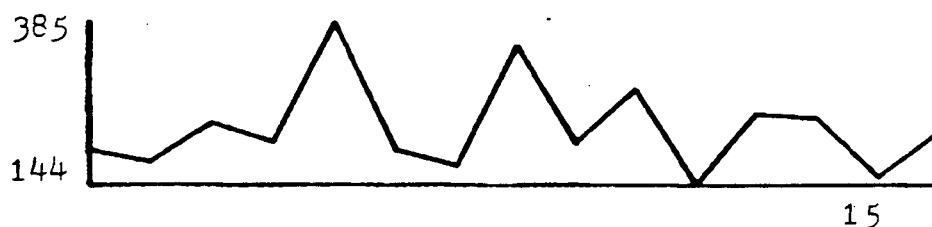


Fig4.38, ppd and lpa plots for bilateral local mass changes, $T_{act} = 25$

4.4. Comparisons with Other Pathological Data

The proposed model has demonstrated that it can produce some interesting perturbation phenomena, but we cannot conclude that the model is valid unless there is evidence that the human vocal fold, under pathological conditions, produces similar phenomena.

In this section we present data from two sources that appear to confirm the validity of the model. These sources are

- the database of pathologies at Vancouver General Hospital
- data from a study by Monsen (1979) using real pathological speech

4.4.1 Patients from Vancouver General Hospital

The first case is for a 63 year old male suffering from chronic laryngitis, which is similar in its effects to laryngeal paralysis, the vocal cords becoming weak and flaccid. Fig 4.39 shows the speech, ppd and lpa plots for this case. The speech wave is very similar to the speech waves produced in the simulation of laryngeal paralysis except that the lpa plot has a random component. The ppd plot however shows a $1/2$ subharmonic with a frequency range of 98 to 107 Hz. The resulting $\text{HNR} = 4.5$ dB, a figure indicating severe hoarseness.

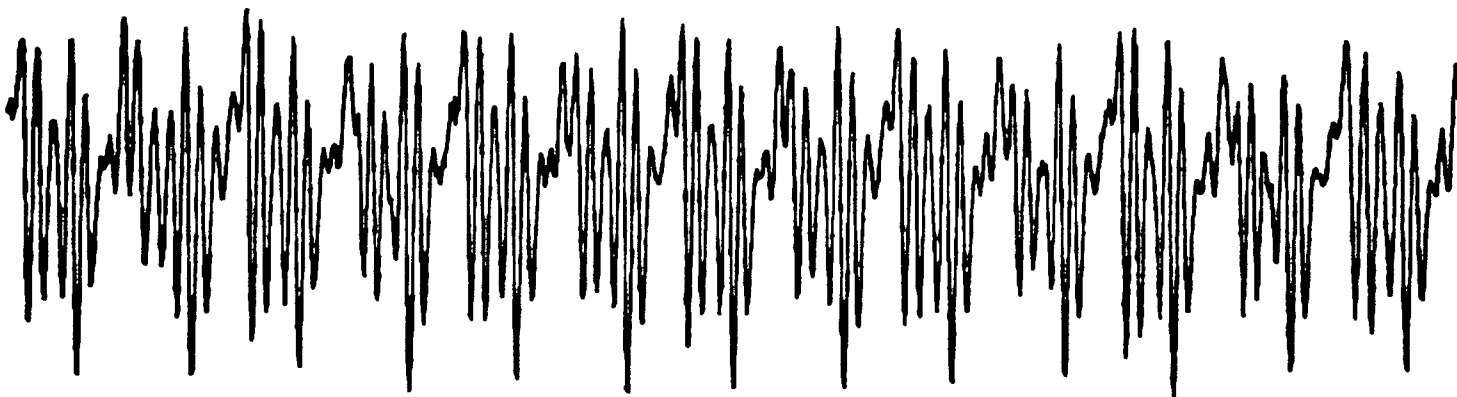
The second case is for a 65 year old male with T1 Glottic cancer. The speech wave in Fig 4.40 does not look like the wave normally associated with /a/, but this could be due to the effects of the cancer. The ppd plot reveals a strong $1/3$ and $1/4$ subharmonic structure, ranging from 132 to 181 Hz. The lpa plot again shows a random shimmer effect. $\text{HNR} = 5.5$ dB, indicating severe hoarseness.

The third case, a 62 year old man, also has T1 Glottic cancer. Fig 4.41 shows the speech, ppd and lpa plots. The ppd and lpa plots are almost completely random, with a frequency range from 107 to 114 Hz. $\text{HNR} = 12.0$ dB, again well within the range for pathologies.

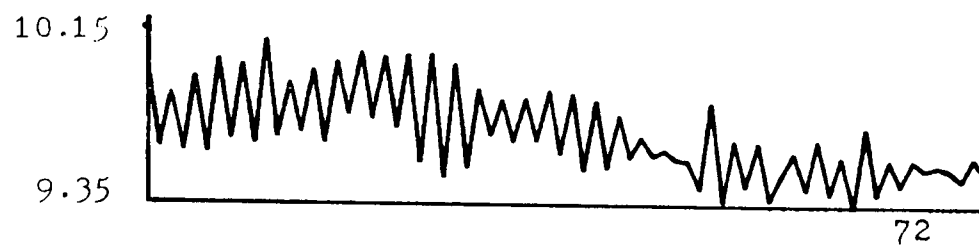
Each of the patients presented in this section clearly illustrate the phenomena predicted in chapter five, although the effect appears to be larger in real data than in the simulations.

As final evidence, we present the frequency versus period number plots found by Monsen (1979) in his study on the vocal function of hoarse voices (Fig 4.42). It

Fig4.39 Speech, ppd and lpa plots, a 63 year old male with chronic laryngitis



Pitch period duration (ms) vs period No.



Largest peak amplitude (dynes/cm²) vs period No.

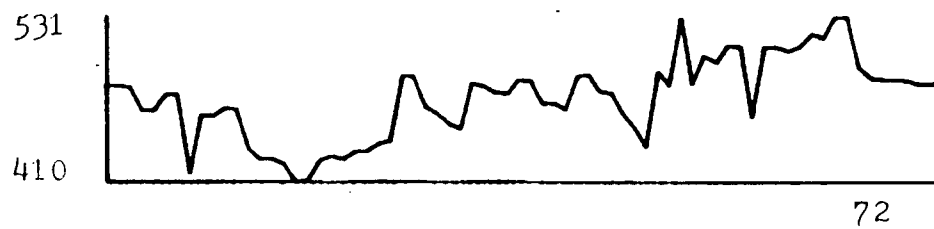
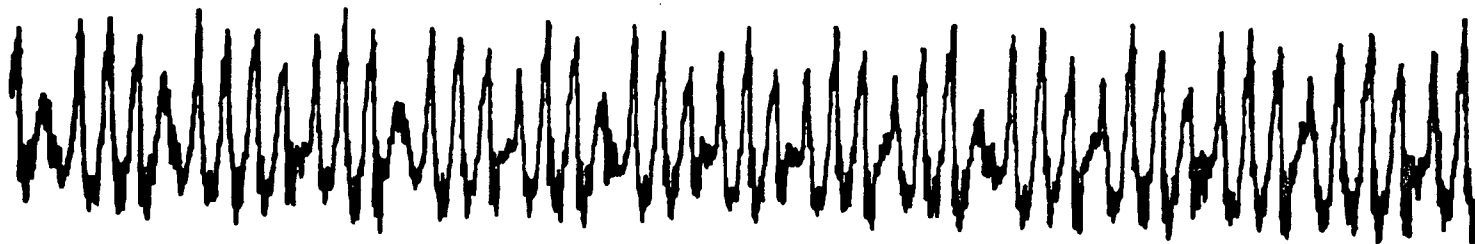
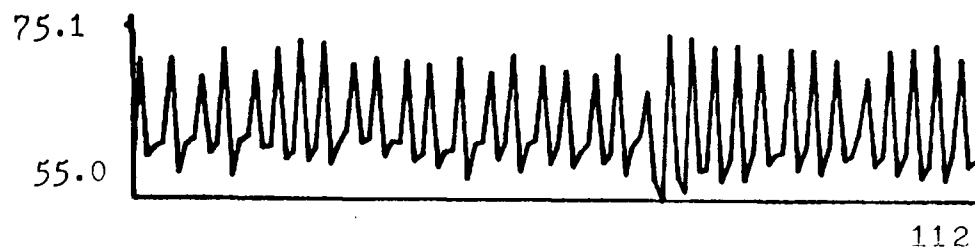


Fig.4.40 Speech, ppd and lpa plots, a 65 year old male with T1 Glottic cancer



Pitch period duration (ms) vs period No.



Largest peak amplitude (dynes/cm²) vs period No.

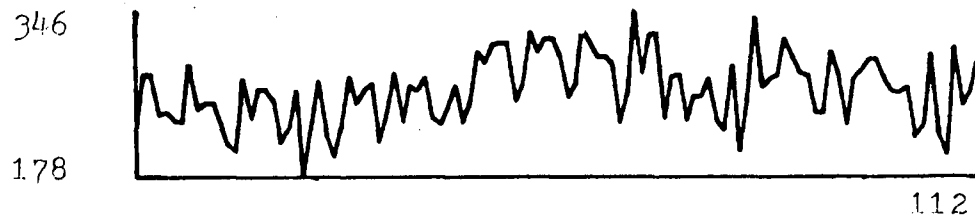
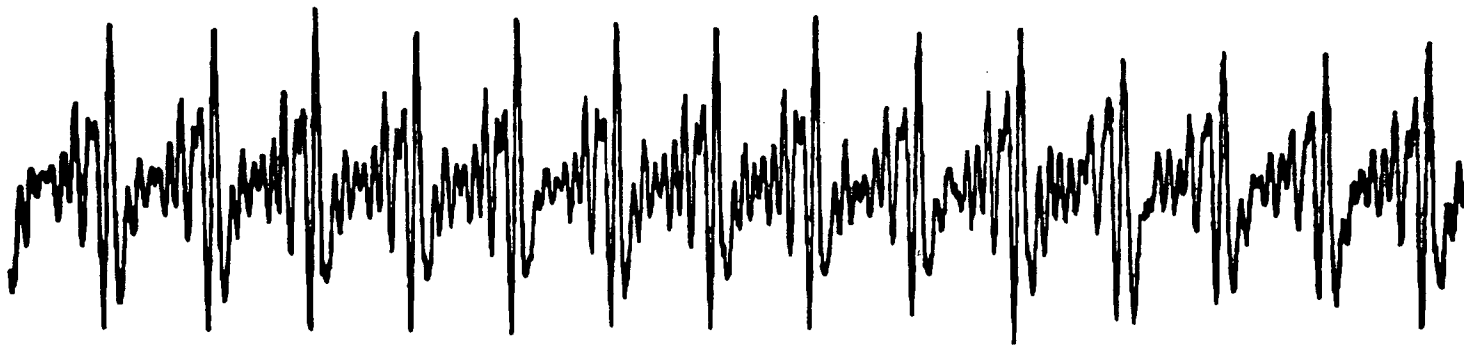
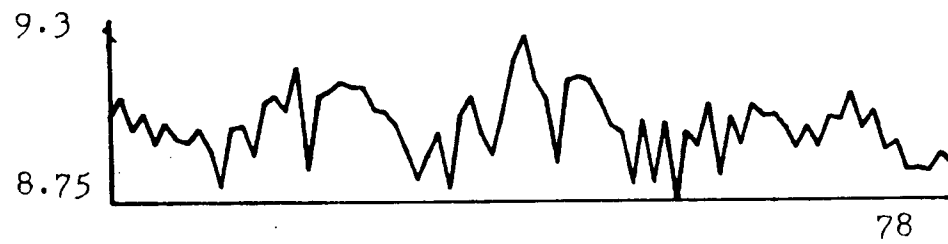


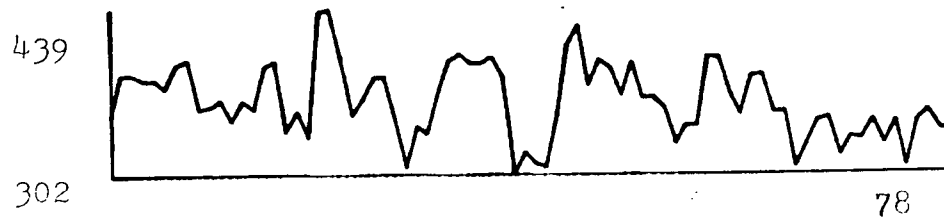
Fig4.41 Speech, ppd and lpa plots, a 62 year old male with T1 Glottic cancer



Pitch period duration (ms) vs period No.



Largest peak amplitude (dynes/cm²) vs period No.



can clearly be seen that for dry hoarse voice there is a $1/2$ subharmonic present, while for harsh hoarse voice there is a random component modulating the subharmonic.

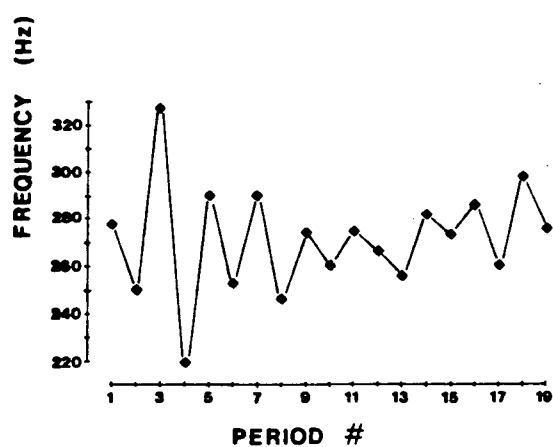
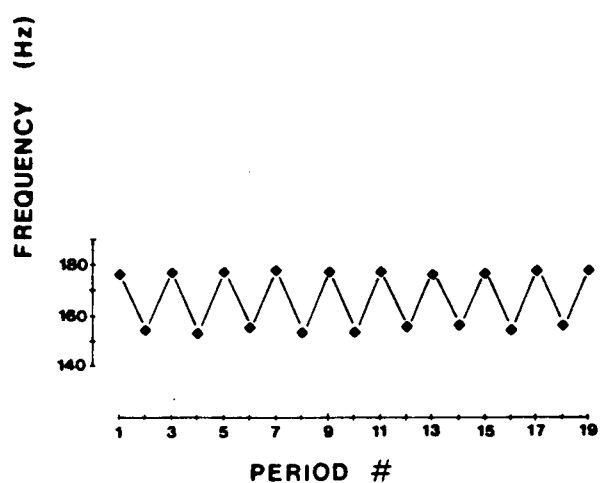


Fig4.42, Frequency vs period number for dry hoarse and harsh hoarse voice. After Monsen, (1979).

4.5. Summary

In this chapter we have presented the most significant results among the many simulations carried out. The results illustrate the effects of changing stiffness and mass separately so that the effects may be classified as due to either stiffness or mass.

In doing so, the actual simulation of particular pathologies has not truly been accomplished. This has been done for two reasons, the first of which is that very little is known about the biomechanical parameters to be used in simulating pathologies. The second reason is that it is probable that much more can be gained by separating the effects of mass and stiffness than by combining them into a potpourri of changes which are poorly understood.

The significant features shown in these simulations are

- a low value of T_{act} is necessary for any perturbations to occur.
- subharmonics of order 1/2 occur for this value of T_{act} regardless of the presence of any other changes in the model, indicating that asymmetry is not required.
- unilateral (uniform) stiffness increases enhance the subharmonic.
- unilateral mass increases inhibit the subharmonic.
- local stiffness increases introduce a small random component but inhibit the subharmonic.
- local mass increases enhance the subharmonic and produce a large random component.
- bilateral localised mass increases also results in a large random component.
- the HNR values are higher than those obtained by Kojima et al (1980) and Yumoto, Gould and Baer (1982) but since turbulent flow and random muscle excitation is not modelled, this is not surprising.
- HNR values are much lower for local mass changes than for all other cases

due to the random modulation.

- in comparison with the data base from VGH, the ppd and lpa plots for the simulated cases are very similar.
- data from Monsen also shows regular subharmonic perturbations and irregular random perturbations which appear to be a combination of subharmonic and random effects.

The following table lists the HNR values (in dB) as T_{act} is decreased from 150 to 25 ('s' indicates the presence of a subharmonic and 'r' the presence of random effects).

Table I HNR values for all simulation cases

<u>Simulation</u>	<u>HNR at 150</u>	<u>HNR at 25</u>
normal	37.7	17.5 (s)
unilateral stiffness	35.9	16.8 (s)
local stiffness	36.1	18.6 (s,r)
unilateral mass	40.26	40.8
local mass	35.7	7.9 (r)
bilateral local mass		0.38 (r)
<u>Real data</u>		
normal		22.3
chronic laryngitis		4.5 (s,r)
T1 Glottic Cancer (1)		5.5 (s,r)
T1 Glottic Cancer (2)		12.0 (r)

CHAPTER 5

ANALYSIS

In this chapter, an approximate analysis of vocal fold oscillation and perturbation is derived. The model is based on work by Titze (1983), using a single mass-spring oscillator vibrating normally to the flow. In his model vertical phase lag was represented by a travelling wave moving axial to the flow on the surface of the mass, and collision was not considered.

The analysis carried out in this chapter extends Titze's work, and shows that some of the perturbation phenomena are deterministic and implicit in the equations describing the system. Certain terms in the developed equations are, according to Hayashi (1964) and Minorsky (1962), indicative of subharmonic oscillations often found in nonlinear oscillatory systems.

5.1. The One Mass Model

The final equation derived by Titze (1983) defined the differential equation for the tissue excited by the pressure created due to the aerodynamic flow. This equation is

$$m\ddot{x} + r\dot{x} + kx = 2y^2T(B_T v + L_T \dot{v})(x_0 + x) + 2y^2TL_T v \dot{x} = P y T \quad \dots(5.1)$$

where m = tissue mass

r = tissue damping

k = tissue stiffness (kx is replaced by $kx + k\eta x^3$ for a cubic hardening function)

y = the longitudinal length of the glottis

T = thickness (from the bottom of the lower margin to the top of the upper margin)

R_T = the downstream resistance as seen from the midpoint of the glottis. It is the summation of the vocal tract load R_L and the post-midpoint glottal resistances R_t , R_m , and R_e . Thus

$$R_T = R_t + R_m + R_c + R_L$$

L_T = the downstream inductance due to the inertia of the air mass within the glottis and the vocal tract. Thus

$$L_T = L_L + L_m$$

v = the velocity of the flow

x = lateral tissue displacement as measured from the midpoint of the thickness dimension

x_0 = the initial displacement for the tissue.

P = the midpoint driving pressure

If we now substitute for v in equation 5.1, we can find terms representing the effective component aerodynamic forces. The flow velocity v in Fig 4.4 is of the same frequency and approximately in phase with the tissue displacement, with a saw tooth shape in which half the period is zero. A Fourier series for v is

$$\begin{aligned} v(t) = & V/4 - (2V/\pi^2)\cos w_0 t - (2V/3^2\pi^2)\cos 3w_0 t - (2V/5^2\pi^2)\cos 5w_0 t \dots \\ & + (V/\pi)\sin w_0 t - (V/2\pi)\sin 2w_0 t + (V/3\pi)\sin 3w_0 t - \dots \end{aligned} \quad \dots(5.2)$$

where w_0 is the fundamental frequency and V is the maximum velocity.

A reasonable approximation to v is the function incorporating the four lowest frequency terms of the Fourier expansion i.e.

$$v(t) = V/4 - (2V/\pi^2)\cos w_0 t + (V/\pi)\sin w_0 t - (V/2\pi)\sin 2w_0 t \quad \dots(5.3)$$

Fig5.1 is a plot of equation 5.3. The similarity between Fig5.1 and Fig4.4 is reasonable if we note that the discontinuities producing the saw tooth result from collision. If collision does not occur, the velocity will more closely resemble Fig5.1.

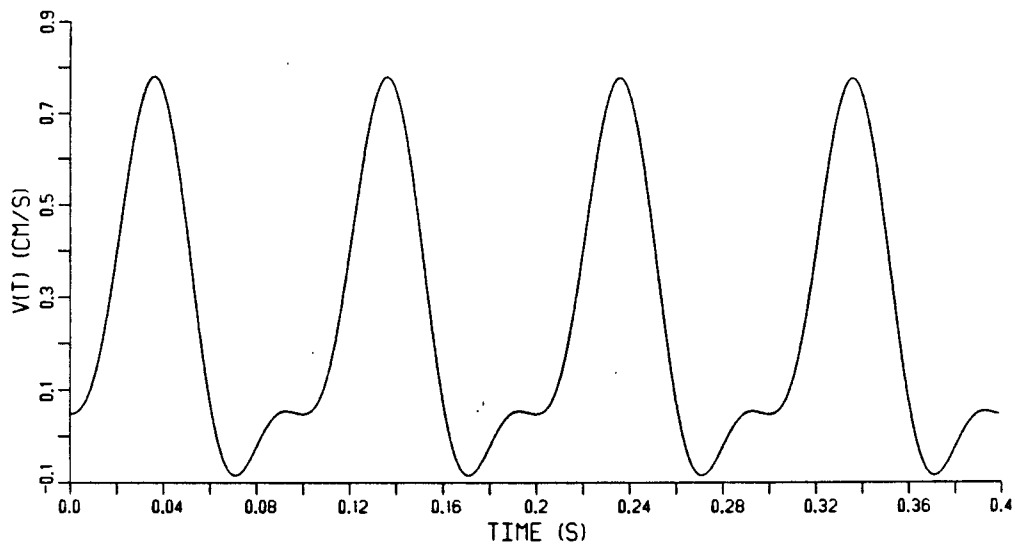


Fig5.1 Four term approximation to $v(t)$

Substitution of this time driven function for v is not informative since there are still two variables, x and t involved. However, since the velocity function is the

same frequency as the displacement, let

$$x = \sin w_0 t ; \dot{x} = w_0 \cos w_0 t ; \text{ and } \ddot{x} = -w_0^2 \sin w_0 t = -w_0^2 x$$

v is now a function of x.

Thus

$$\begin{aligned} v &= V/4 - (2V/\pi^2 w_0) \dot{x} + Vx/\pi - (V/\pi w_0) x \dot{x} \\ \dot{v} &= (2Vw_0/\pi^2) x + V\dot{x}/\pi + (Vw_0/\pi) x^2 - (V/\pi w_0) \dot{x}^2 \end{aligned} \quad \dots(5.4)$$

The right hand side of equation 5.1 becomes

$$\begin{aligned} &= 2y^2 T [R_T (V/4 - (2V/\pi^2 w_0) \dot{x} + Vx/\pi - (V/\pi w_0) x \dot{x}) (x_0 + x) \\ &+ L_T ((2Vw_0/\pi^2) x + (V/\pi) \dot{x} + (Vw_0/\pi) x^2 - (V/\pi w_0) \dot{x}^2) (x_0 + x) \\ &+ L_T (V/4 - (2V/\pi^2 w_0) \dot{x} + (V/\pi) x - (V/\pi w_0) x \dot{x}) \dot{x}] \end{aligned} \quad \dots(5.5)$$

Expanding 5.5, and setting $a = 2y^2 T$

$$\begin{aligned} &= a [R_T V x_0 / 4 - (R_T x_0 2V/\pi^2 w_0) \dot{x} + (R_T x_0 V/\pi) x - (R_T x_0 V/\pi w_0) x \dot{x} + (R_T V/4) x \\ &- (R_T 2V/\pi^2 w_0) x \dot{x} + (R_T V/\pi) x^2 - (R_T V/\pi w_0) x^2 \dot{x} + (L_T x_0 2Vw_0/\pi^2) x + (L_T x_0 V/\pi) \dot{x} \\ &+ (L_T x_0 Vw_0/\pi) x^2 - (L_T x_0 V/\pi w_0) \dot{x}^2 + (L_T 2Vw_0/\pi^2) x^2 + (L_T V/\pi) x \dot{x} + (L_T Vw_0/\pi) x^3 \\ &- (L_T V/\pi w_0) x \dot{x}^2 + (L_T V/4) \dot{x} - (L_T 2V/\pi^2 w_0) \dot{x}^2 + (L_T V/\pi) x \dot{x} - (L_T V/\pi w_0) x \dot{x}^2] \end{aligned} \quad \dots(5.6)$$

The impedances R_T and L_T consist of a constant vocal tract load and a variable glottal load that varies inversely with the size of the glottal constriction. Consequently during the periods just after the glottis opens and just before it closes, the impedances become very large and the pressures are most significant in terms of magnitude. Fig 5.2 is an illustration of the types of forces that can occur near the closure points. The terms in equation 5.6 can be associated with each of the diagrams.

Points to note are

1. Bernoulli forces

Fig 5.2a illustrates those forces which are negative regardless of the direction of motion. Collating these terms we have

$$a[-(L_T V / \pi w_0) x \dot{x}^2 - (L_T 2V / \pi^2 w_0) \dot{x}^2 - (L_T V / \pi w_0) x \dot{x}^2 - (L_T x_0 V / \pi w_0) \dot{x}^2]$$

These forces are the Bernoulli forces, producing a negative 'sucking' force regardless of the direction of motion of the mass. It should be noted that these terms are divided by w_0 , which is very large when the fundamental frequency is 100 Hz. Consequently, these forces while significant, are not dominant.

2. Velocity dependent forces opposing motion

The forces in Fig 5.2b are velocity dependent, but oppose the direction of motion. These terms are

$$a[-(R_T x_0 2V / \pi^2 w_0) \dot{x} - (R_T x_0 V / \pi w_0) x \dot{x} - (R_T 2V / \pi^2 w_0) x \dot{x} - (R_T V / \pi w_0) x^2 \dot{x}]$$

Similar to those terms producing the Bernoulli effect, a w_0 divisor is present, and thus the forces are not dominant.

3. Forces opposing the Bernoulli forces

The forces in Fig 5.2c,

$$a[R_T V x_0 / 4 + (R_T x_0 V / \pi) x + (R_T V / \pi) x^2 + (R_T V / 4) x + (L_T x_0 2V w_0 / \pi^2) x + (L_T 2V w_0 / \pi^2) x^2 + (L_T x_0 V w_0 / \pi) x^2 + (L_T V w_0 / \pi) x^3]$$

oppose the Bernoulli forces. Although there is no w_0 divisor they are relatively small near closure as they are multiplied by x (the displacement determining the constriction), which is very small near closure. $R_T V x_0 / 4$ is of

constant magnitude, aiding the opening cycle and opposing the closing cycle. It is dominated, however, by the velocity dependent forces aiding motion.

4. Velocity dependent forces aiding motion

The forces represented by Fig 5.2d are the largest pressure dependent forces driving the system as they are not all functions of x or $1/w_0$. They are, however, functions of the impedance L_T , indicating that inertial loads due to the vocal tract or the glottis are important in sustaining oscillation. The velocity dependent terms are:

$$a[(L_T x_0 V/\pi)\dot{x} + (L_T V/\pi)x\dot{x} + (L_T V/4)\dot{x} + (L_T V/\pi)x\dot{x}]$$

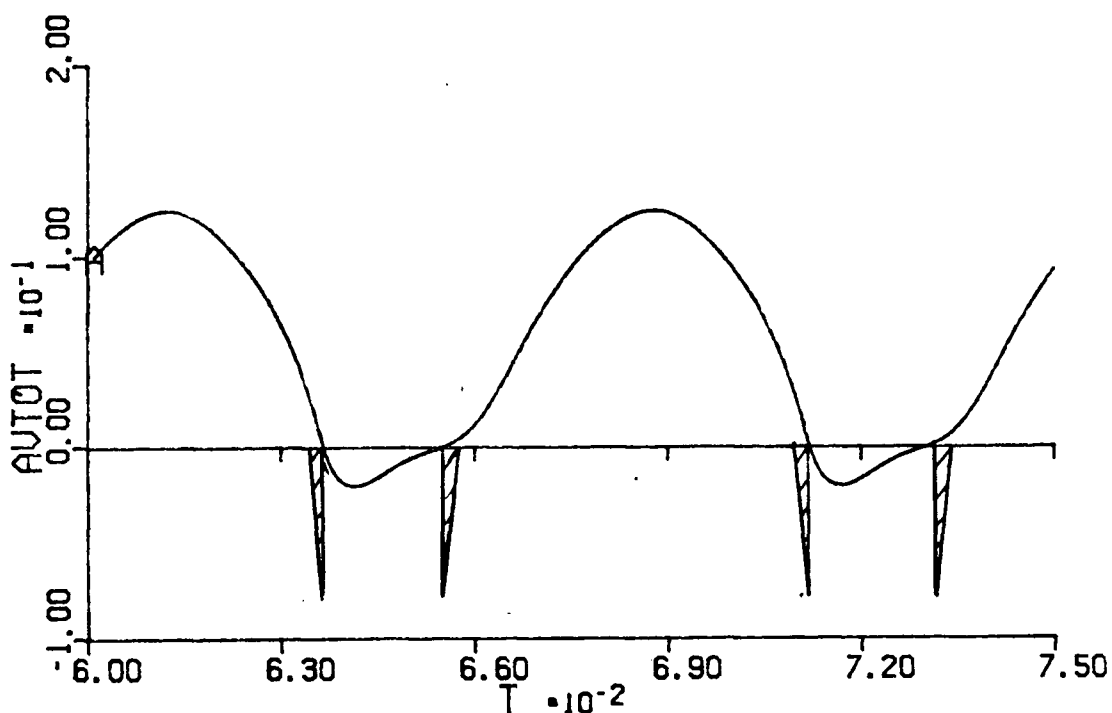


Fig 5.2a Bernoulli forces

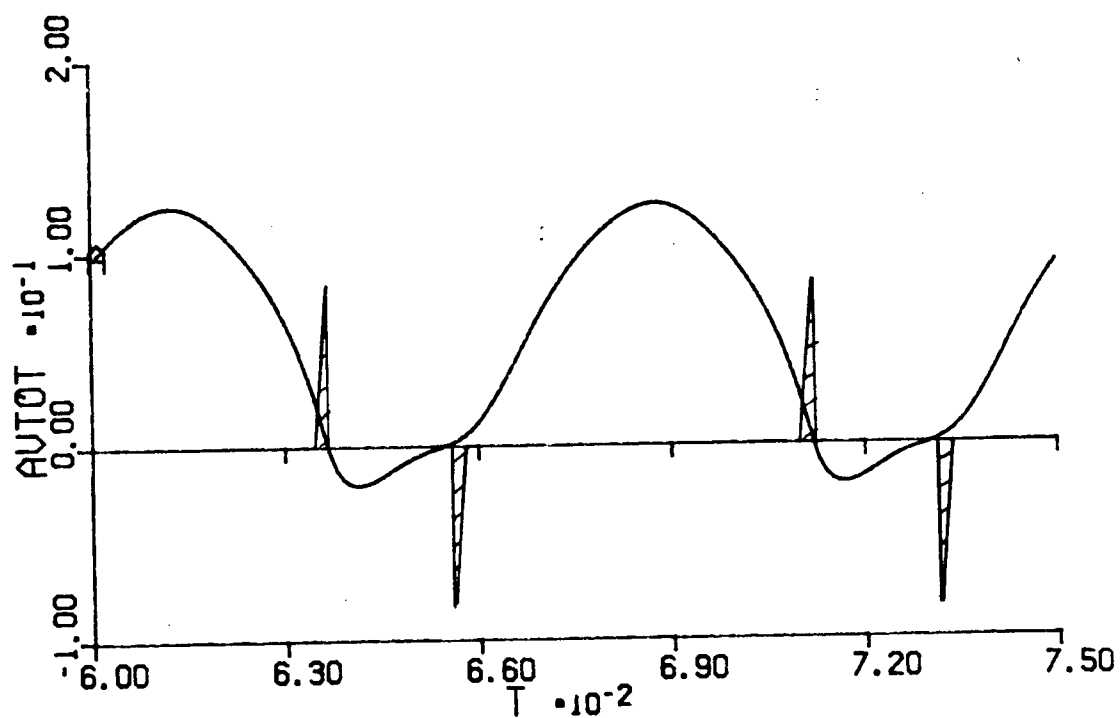


Fig5.2b Velocity dependent forces opposing motion

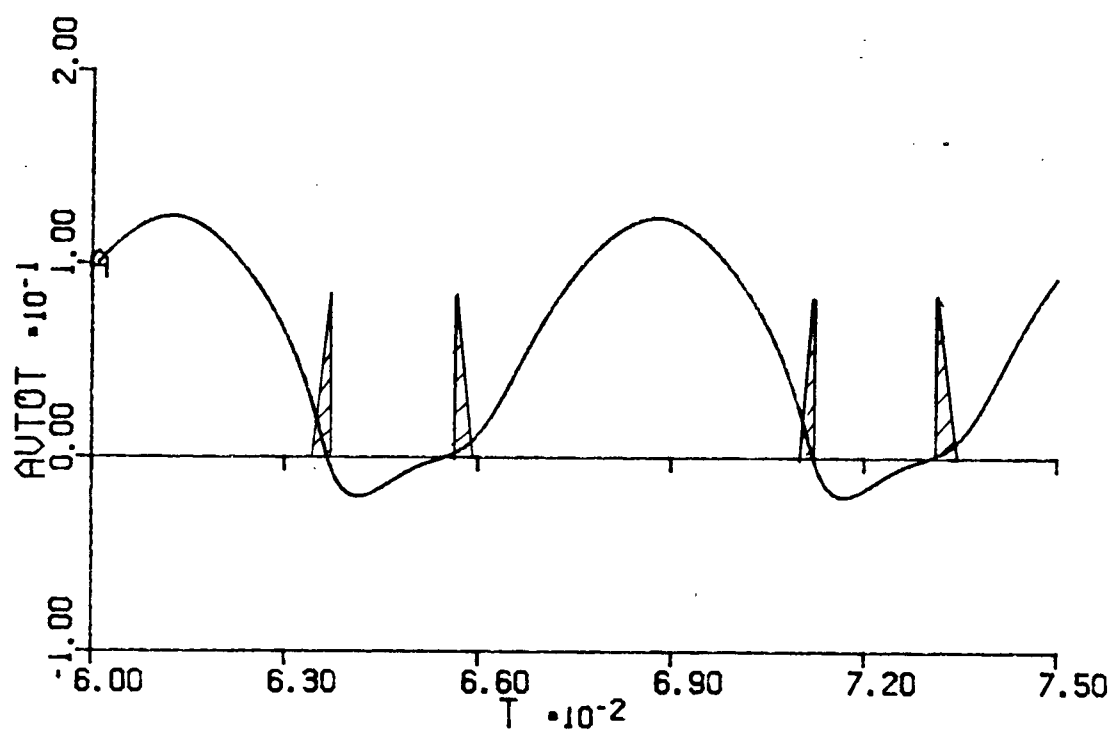


Fig5.2c Forces opposing the Bernoulli forces

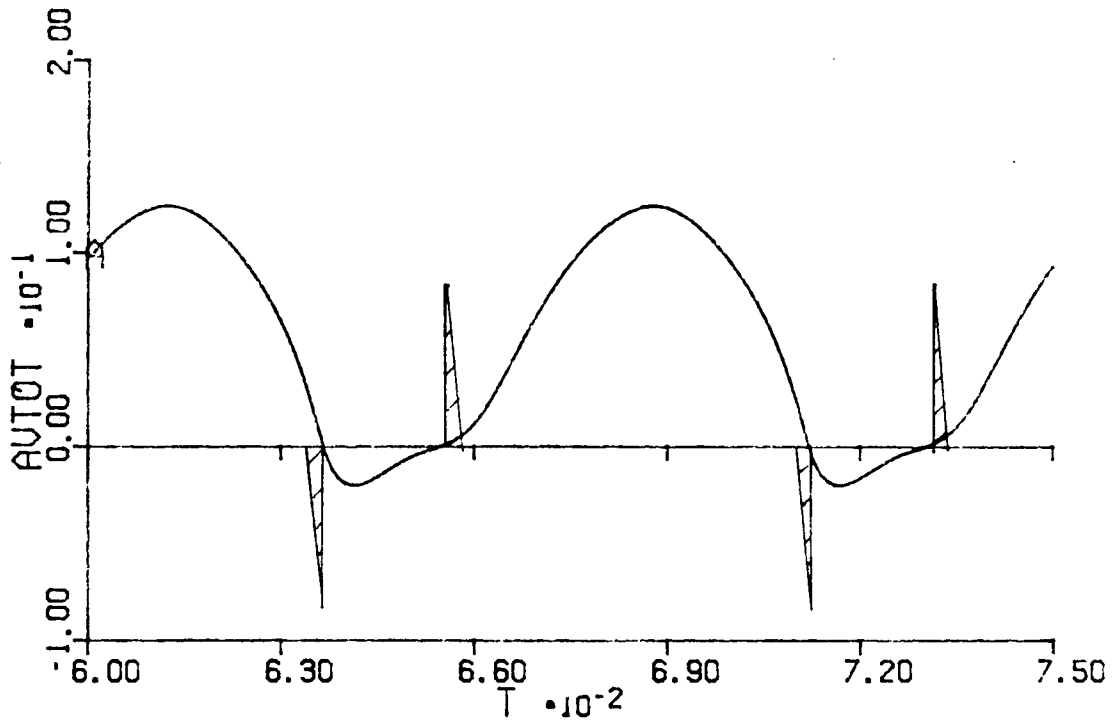


Fig 5.2d Velocity dependent forces aiding motion

The dynamic equation for the system can be found by substituting 5.6 into 5.1 and rearranging, giving

$$\begin{aligned}
 m\ddot{x} + \dot{x} [& (r - a(L_T V/4 + L_T x_0 V/\pi - R_T x_0 2V/\pi^2 w_0)) \\
 & - a(2L_T V/\pi - R_T 2V/\pi^2 w_0 - R_T x_0 V/\pi w_0)x + (aR_T V/\pi w_0)x^2 + (aL_T x_0 V/\pi w_0)\dot{x} \\
 & + (a2L_T V/\pi w_0)x\dot{x}] \\
 & + x(k - a(R_T x_0 V/\pi + R_T V/4 + L_T x_0 2Vw_0/\pi^2)) \\
 & - ax^2(R_T V/\pi + L_T x_0 Vw_0/\pi + L_T x_0 2Vw_0/\pi^2) \\
 & + x^3(k\eta - aL_T Vw_0/\pi) + (aL_T 2V/\pi^2 w_0)\dot{x}^2 - aR_T Vx_0/4 = 0 \quad \dots(5.7)
 \end{aligned}$$

The terms with a w_0 divisor may be neglected when considering the most significant forces since for a typical frequency of 100 Hz, $w_0 = 200\pi$. Thus, for example, consider the relationship between

$$L_T x_0 V/\pi : R_T x_0 2V/\pi^2 w_0 \quad \dots(5.8)$$

which equates to

$$L_T : 2R_T/200\pi^2$$

From Titze (1983) an appropriate $L_T : R_T$ ratio is 1 : 50. Thus the ratio in 5.8 is approximately 20 : 1, or a 5% contribution by the R_T term. A comparison of all other $1/w_0$ terms reveals a similar or smaller contribution, and consequently in this study these terms are neglected.

Equation 5.7 can thus be simplified to

$$\begin{aligned} m\ddot{x} + \dot{x} & \left((r - a(L_TV/4 + L_{Tx_0}V/\pi)) - (a2L_TV/\pi)x \right) \\ + x & \left(k - a(R_{Tx_0}V/\pi + R_TV/4 + L_{Tx_0}2Vw_0/\pi^2) \right) \\ - x^2 & a(R_TV/\pi + L_{Tx_0}Vw_0/\pi + L_T2Vw_0/\pi^2) \\ + x^3 & (k\eta - aL_TVw_0/\pi) - aR_TVx_0/4 = 0 \end{aligned} \quad \dots(5.9)$$

A number of terms important to sustained oscillation can be observed in equation 5.9.

1. Damping

The damping or velocity coefficient is no longer constant, and may be negative when the sinusoidally varying $(a2L_TV/\pi)x$ is greater than $(r - a(L_TV/4 + L_{Tx_0}V/\pi))$. The $(r - a(L_TV/4 + L_{Tx_0}V/\pi))$ term is a threshold preventing the damping coefficient from becoming negative, the positive inertial impedance L_T countering the dissipation from r and reducing the threshold. When the overall damping is negative, energy is absorbed from the air flow by the tissue to compensate for tissue losses.

The sinusoidal variation in $-(a2L_TV/\pi)x$ produces an alternately negative and positive damping which results in the absorption and dissipation

of energy. Thus the inductive load L_T is one of the dominant factors in maintaining the energy balance in the system, as it helps to overcome the dissipative effect of the tissue damping.

As L_T varies inversely with the size of the constriction, when the displacement is very large, L_T is very small, resulting in a higher threshold, and possibly preventing a negative coefficient value. The system thus dissipates energy momentarily until the high threshold decreases with increasing L_T and decreasing x . Another possible explanation for the momentary dissipation is that the system becomes a Van der Pol oscillator due to the $(aR_TV/\pi w_0)x^2\dot{x}$ term in 5.7. Equation 5.7 resembles the Van der Pol equation

$$\ddot{x} + \mu(x^2 - 1)\dot{x} + x = 0$$

when x becomes very large, as the x^2 becomes significant in spite of the decreasing R_T , and the w_0 divisor.

Near closure, however, L_T is large, so the threshold is easily overcome, and for positive displacement energy is absorbed.

Observation of the energy diagram of Fig 4.4 (the normal fold) illustrates the described phenomena, with dissipation occurring for both negative displacement and large positive displacement, and absorption for all positive displacements less than peak.

2. Vertical phasing

Until now the effects of vertical phasing have not been discussed and in fact equation 5.1 was developed by Titze (1983) without introducing this phenomenon (although later in his paper vertical phasing is considered). It is clear from equation 5.9 that oscillation is possible without it, but we will now

consider the effect it has on oscillation, based on Titze's study (1983).

When vertical phasing is present, the glottis is divergent when the fold is approaching closure, as the upper margin orifice A_{mtot} is larger than the lower A_{vtot} and midpoint margins. If we consider a fold in which no vertical phasing occurs, ($A_{vtot} = A_{mtot}$), the viscous resistance R_m will be greater than for the divergent glottis since the area A_{mtot} is less. In fact, it is possible for the divergent glottis shape to exhibit a small pressure recovery, causing R_m to be very small. As a result R_T is reduced in magnitude.

The opposite phenomenon occurs when the vocal fold is in its opening cycle, as the lower margin leads the upper, resulting in a convergent profile. A_{mtot} is less than for a fold not exhibiting vertical phasing, and consequently R_m and R_T are greater. The impedances L_m and L_T respond in a similar way.

The effect of the vertical phase lag then, is to provide glottal impedances which are asymmetrical in magnitude (direction, or velocity dependent). Thus, during glottal opening, the impedances for a convergent profile increase causing an increased midpoint pressure. As a result the masses are aided in the direction of motion. During closure, on the other hand, the impedances decrease and consequently the forces opposing closure decrease allowing the masses to be 'sucked' in.

3. Stiffness

The terms proportional to x , x^2 , and x^3 determine the 'spine' of both the energy diagram and the frequency response. The presence of terms involving L_T and R_T modify the spine by decreasing the gradient. Note that the presence of the x^2 term is solely due to the aerodynamic flow and not

the tissue.

$k\eta x^3$ is commonly known as the cubic hardening function in Duffings equation (cit p378, Meirovitch, 1975), and its presence results in a frequency response which does not possess a fixed resonance frequency. For a damped system, Duffings equation is of the form

$$\ddot{x} + w^2x = \epsilon(-2\rho w\dot{x} - w^2(\alpha x + \beta x^3) + F\cos \Omega t) \quad \dots(5.10)$$

with a frequency response as in Fig5.3. $F\cos \Omega t$ is an external harmonic forcing function.

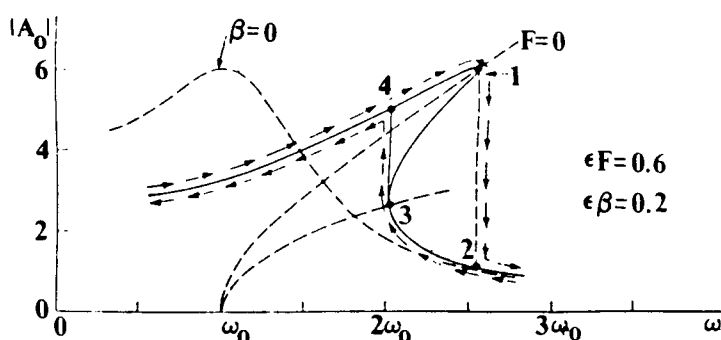


Fig5.3 Frequency response for Duffings equation. After Meirovitch (1975)

Ishizaka and Flanagan (1972) examined the effect of increasing subglottal pressure P_L on the fundamental frequency (Fig5.4) for various η (η_v , η_c , and η_m were all varied by the same amount). The η controlled the 'spine' of the stiffness function.

η is the equivalent of β in Duffings equation. The trend of increasing fundamental frequency with greater η (or β) is apparent in both Fig5.3 and Fig5.4.

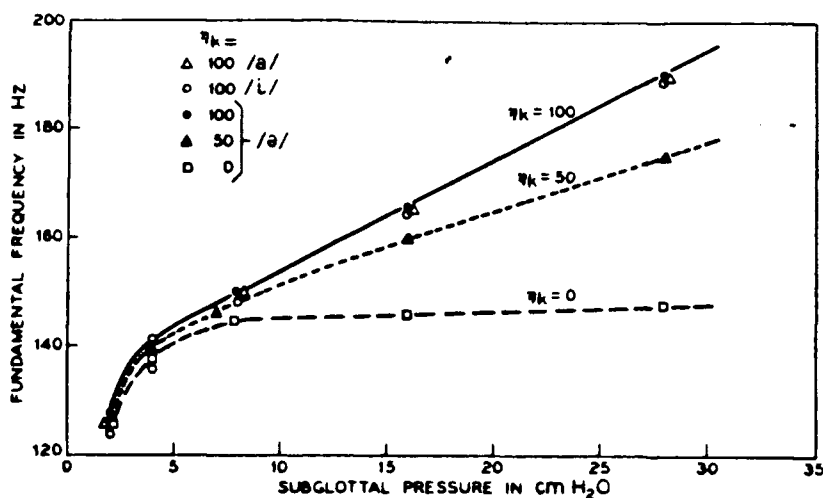


Fig5.4 P_l vs f_0 for the Ishizaka-Flanagan model. After Ishizaka and Flanagan, 1972

4. Subharmonic Perturbations

According to Hayashi (1964), it is possible for submultiple oscillations of the fundamental frequency to appear with the fundamental. Two requirements are necessary in the differential equation, and these are

(1) stiffness functions of the form

$$bx^2 + cx^3 + \dots$$

which will give subharmonics of $1/2$ and $1/3$ (order 2 and 3).

and (2) The equation must be driven by an external harmonic function.

Equation 5.9 fulfills the first requirement, but not the second. The second requirement is necessary as it allows the system to traverse a non unique phase trajectory (this is prohibited in a second order autonomous system). A second order equation such as equation 5.9 is not an adequate representation of the proposed model, as data from real pathological speech (Monsen, 1979), and the simulations carried out in this study have shown non unique trajectories and a $1/2$ subharmonic. The equations must therefore be reconsidered by expanding to the two mass system.

5.2. Fourth Order Systems

Titze's model (1983), assuming a single mass system, results in the autonomous, second order differential of equation 5.9. A phase plot of the motion of such a system would be expected to display a trajectory which moves out to a single limit cycle following a unique trajectory (without crossing over itself). Fig 4.3 shows, for the normal cord, a trajectory which is not unique. This implies that the proposed simulation model does not conform exactly to equation 5.9. For a trajectory to deviate from its unique path either

- (1) an external harmonic forcing function is present,
- or (2) the system is of higher than second order.

A single mass system such as the Flanagan-Landgraf model cannot exhibit a non-unique phase trajectory, and therefore cannot undergo the subharmonic oscillations seen in chapter four.

To explain the presence of the 1/2 subharmonic, the proposed simulation model must be thought of as a system of two second order systems coupled together (representing the upper and lower masses respectively) to form a fourth order autonomous system in which each second order system acts as a driving function for the other.

5.2.1 Expansion to a Fourth Order System

If, instead of a one mass system, two masses are assumed, then the pressures at some distance above and below the midpoint can be used to drive each of the masses. As in the Ishizaka-Flanagan model, a coupling spring between the two masses is added, producing an additional force $k_c(x_v - x_m)$. Thus, let P_v be the driving pressure at the midpoint of the vocalis mass, and P_m be the equivalent pressure for

the mucosa. Assume that the relationships between P_v , P_m and P are calculated by moving upstream and downstream from the midpoint respectively. Then

$$P_v = P + (L_v/2)(\dot{u} + \dot{u}_d) + (R_v/2)(u + u_d)$$

and

$$P_m = P - (L_m/2)\dot{u} - (R_m/2)u \quad \dots(5.11)$$

where u , \dot{u} are the flow and flow differential, with

$$u = 2y(x_0 + x)v$$

u_d and \dot{u}_d as the displacement flow and differential due to the incremental area change caused by lateral tissue displacement ie. $u_d = 2yT\dot{x}$

R_v , R_m , L_v , and L_m are the glottal impedances defined in the proposed model.

Then,

$$P_v = 2y(R_T v + L_T \dot{v})(x_0 + x) + (L_T v/T)2yT\dot{x} + L_v \dot{u}/2 + L_v \dot{u}_d/2 + R_v u/2 + R_v u_d/2 \quad \dots(5.12)$$

If $u_d = 2yT\dot{x}$, $\dot{u}_d = 2yT\ddot{x}$, $u = 2y(x_0 + x)v$,
and $\dot{u} = 2y(x_0 + x)\dot{v} + 2y\dot{x}v$ then

$$\begin{aligned} P_v &= 2y(x_0 + x)(R_T v + L_T \dot{v} + R_v v/2 + L_v \dot{v}/2) \\ &+ 2y(L_T v + L_v v/2)\dot{x} + 2yT(R_v/2)\ddot{x} + 2yTL_v \ddot{x}/2 \\ &= 2y(x_0 + x)(R_T + R_v/2)v + (L_T + L_v/2)\dot{v} \\ &+ 2y\dot{x}((L_T + L_v/2)v + TR_v/2) + 2yTL_v \ddot{x}/2 \end{aligned} \quad \dots(5.13)$$

Substituting equation 5.4 for v and \dot{v} and letting

$$R^\circ = R_T + R_v/2 \text{ and}$$

$$L^\circ = L_T + L_v/2$$

an expression purely in terms of x can be derived. A further substitution of

$$x = (x_v + x_m)/2$$

$$\dot{x} = (\dot{x}_v + \dot{x}_m)/2$$

$$\ddot{x} = (\ddot{x}_v + \ddot{x}_m)/2$$

gives an expression in terms of x_v and x_m .

Thus $P_v y_T =$

$$\begin{aligned} & (\ddot{x}_v + \ddot{x}_m)aL_v/4 + (a/2)(\dot{x}_v + \dot{x}_m)[L^\circ V/4 + L^\circ x_0 V/\pi + R_v/2] \\ & + (a/4)(\dot{x}_v + \dot{x}_m)(x_v + x_m)[2L^\circ V/\pi] \\ & + (a/2)(x_v + x_m)[R^\circ x_0 V/\pi + R^\circ V/4 + L^\circ x_0 2Vw_0/\pi^2] \\ & + (a/4)(x_v + x_m)(x_v + x_m)[R^\circ V/\pi + L^\circ x_0 Vw_0/\pi + L^\circ 2Vw_0/\pi^2] \\ & + (a/8)L^\circ Vw_0/\pi(x_v + x_m)^3 + aR^\circ Vx_0/4 \end{aligned} \quad \dots(5.14)$$

where the terms with a w_0 divisor are neglected as before. The derivation of the dynamical equation is identical to that for equation 5.9 except for the terms $(2yTR_v/2)\ddot{x} + 2yTL_v\ddot{x}/2$ (which are not functions of v or \dot{v}), and the use of R° and L° . Thus, if the vocalis differential equation is

$$m_v\ddot{x}_v + r\dot{x}_v + k_v x_v + \eta_v k_v x_v^3 + k_c(x_v - x_m) = yTP_v \quad \dots(5.15)$$

then the following equation is derived:

$$\begin{aligned} & \ddot{x}_v(m_v - aL_v/4) + \dot{x}_v((r_v - a(L^\circ V/4 + L^\circ x_0 V/\pi + R_v/2)/2 - (a/2)(L^\circ V/\pi)x_v) \\ & + x_v(k_v + k_c - (a/2)(R^\circ x_0 V/\pi + R^\circ V/4 + L^\circ x_0 2Vw_0/\pi^2)) \\ & - x_v^2 a(R^\circ V/\pi + L^\circ x_0 Vw_0/\pi + L^\circ 2Vw_0/\pi^2) \\ & + x_v^3(k_v \eta_v - (a/8)(L^\circ Vw_0/\pi)) - aR^\circ Vx_0/4 \end{aligned}$$

$$\begin{aligned}
&= \\
&\ddot{x}_m L_v T/4 + \dot{x}_m ((a/2)(L^\circ V/4 + L^\circ x_0 V/\pi + R_v/2) + (a/2)(L^\circ V/\pi)x_m) \\
&+ x_m (a/2)(R^\circ x_0 V/\pi + R^\circ V/4 + L^\circ x_0 2Vw_0/\pi^2) \\
&+ x_m^2 (a/4)(R^\circ V/\pi + L^\circ 2Vw_0/\pi^2) + x_m^3 (a/8)(L^\circ Vw_0/\pi) \\
&+ [a/4(L^\circ 2V/\pi)(x_m \dot{x}_v + x_v \dot{x}_m) + (a/2)(x_v x_m)(R^\circ V/\pi + L^\circ x_0 Vw_0/\pi \\
&+ L^\circ 2Vw_0/\pi^2) + (a/8)(L^\circ w_0 V/\pi)(3x_v^2 x_m + 3x_m^2 x_v)] \quad \dots(5.16)
\end{aligned}$$

the $[]$ terms are cross coupling terms. It can be seen that both the left and right sides of equation 5.11 have a similar form to 5.4, and assuming x_m is a sinusoidal function, we can expect it to act as a driving function for the left hand side.

In a similar fashion, the mucosal pressure is derived as:

$$P_m = P - (L_m/2)(2y(x_0 + x)\dot{v} + 2y\dot{x}v) - (R_m/2)2y(x_0 + x)v \quad \dots(5.17)$$

where u and \dot{u} have been replaced by the same expressions used previously.

Substituting equation 5.1 for P ,

$$P_m = 2y(x_0 + x)[(R_T - R_m/2)v + (L_T - L_m/2)\dot{v}] + 2y\dot{x}(L_T - L_m/2)v \quad \dots(5.18)$$

Substituting

$$\begin{aligned}
R_T - R_m/2 &= R^\dagger \\
L_T - L_m/2 &= L^\dagger,
\end{aligned}$$

and equation 5.4 for v , equation 5.17 can be expressed in terms of x only. The resulting equation 5.19 neglects the terms involving the divisor w_0 , for reasons described previously.

$$yTP_m = ax\dot{x}(L^\dagger)V/4 + L^\dagger x_0 V/\pi + ax\dot{x}(L^\dagger 2V/\pi)$$

$$\begin{aligned}
& + ax(R\dot{V}x_0V/\pi + R\dot{V}V/4 + L\dot{V}2Vw_0/\pi^2) \\
& + ax^2(R\dot{V}V/\pi + L\dot{V}x_0Vw_0/\pi + L\dot{V}2Vw_0/\pi^2) + ax^3L\dot{V}Vw_0/\pi + aR\dot{V}Vx_0/4 \\
& \dots(5.19)
\end{aligned}$$

Equation 5.19 can be equated to a mucosal tissue differential equation similar to 5.16 and x can be replaced by $(x_v + x_m)/2$ to eventually yield

$$\begin{aligned}
& \ddot{x}_m m_m \\
& + \dot{x}_m(r_m - (a/2)L\dot{V}V/4 + L\dot{V}x_0V/\pi) - (a/4)x_m L\dot{V}V/\pi) \\
& + x_m[k_m + k_c - (a/2)(R\dot{V}x_0V/\pi + R\dot{V}V/4 + L\dot{V}x_02Vw_0/\pi^2)) \\
& - (a/4)x_m^2(R\dot{V}V/\pi + L\dot{V}x_0Vw_0/\pi + L\dot{V}2Vw_0/\pi^2)) \\
& + x_m^3(k_m \eta_m - (a/8)(L\dot{V}Vw_0/\pi)) - aR\dot{V}x_0V/4 \\
& = \\
& \dot{x}_v((a/2)(L\dot{V}V/4 + L\dot{V}x_0V/\pi) + (a/4)x_v L\dot{V}2V/\pi \\
& + x_v[k_c + (a/2)(R\dot{V}x_0V/\pi + R\dot{V}V/4 + L\dot{V}x_02Vw_0/\pi^2)) \\
& + (a/4)x_v^2(R\dot{V}V/\pi + L\dot{V}x_0Vw_0/\pi + L\dot{V}2Vw_0/\pi^2) + (a/8)x_v^3(L\dot{V}Vw_0/\pi) \\
& [+ (a/4)((\ddot{x}_m x_m + x_m \ddot{x}_v)L\dot{V}2V/\pi \\
& + (ax_v x_m/2)(R\dot{V}V/\pi + L\dot{V}x_0Vw_0/\pi + L\dot{V}2Vw_0/\pi^2) \\
& + (3a/8)(L\dot{V}Vw_0/\pi)(x_v x_m^2 + x_m x_v^2)] \dots(5.20)
\end{aligned}$$

Equation 5.20 has a very similar form to 5.16, but lacks inertially coupled terms. By expanding the equations to fourth order, the driving forces now become obvious as each of the masses provides the driving force for the other. The second requirement for subharmonic oscillation has thus been fulfilled.

5.3. Perturbations

5.3.1 Further Examination of Subharmonics

According to Hayashi (1964), a $1/2$ subharmonic is apt to occur only when the stiffness nonlinearity is asymmetrical

$$\text{ie. } ax + bx^2 + cx^3$$

whereas $1/3$ subharmonics can occur when $ax + cx^3$ terms are present, regardless of the presence of bx^2 . Thus, since $ax + bx^2 + cx^3$ is present, it is possible for subharmonics of both $1/2$ and $1/3$ to occur.

Thus far, equations 5.16 and 5.20 have proven too difficult to solve, so the appearance of subharmonics, while likely, has not been analytically proven. However, the simulations presented in chapter four demonstrate phenomena that appear to be subharmonics, and it is our contention that this is the mechanism for regular perturbation in the model, and that it is also manifested in real vocal fold vibration.

As evidence we recall the real data obtained by Monsen (Fig 4.42) and we also present in Fig 5.5 the pitch vs period number plot attained by Isshiki and Ishizaka (1976) for various tension imbalances using the two mass model. It is evident from Fig 4.42 and Fig 5.5 that small subharmonic components of $1/2$, $1/3$ and higher order modulate the fundamental.

In our simulations, the subharmonic only appears for low T_{act} . There are two possible reasons for this, and they are rooted in both linear and nonlinear oscillation.

In a linear oscillatory system, the amount of damping determines the amplitude and decay of the vibration. As an analogue, one might expect equations 5.16 and 5.20 to respond similarly. The F_{tot} vs time plot in Fig 4.6 shows an impulsive force indicating a rich harmonic spectrum. By decreasing T_{act} the tissue damping and stiffness decreases (equations 3.7 and 3.10). Thus the superharmonics and subharmonics can appear without being attenuated by the damping.

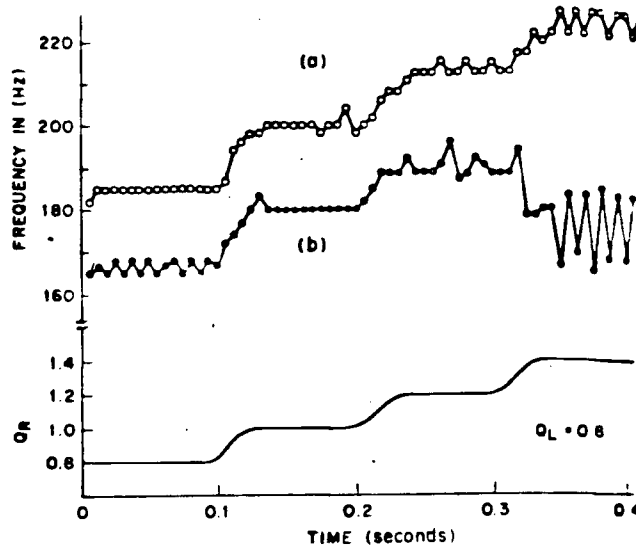


Fig5.5 Frequency Fluctuation Plot. After Isshiki and Ishizaka (1976)

An analysis by Minorsky (1962) for a nonlinear system with a $1/2$ subharmonic showed the subharmonic amplitude to be optimally dependent on the stiffness. Thus for greater stiffness the conditions may have been less than optimal for the subharmonic, and by decreasing the stiffness conditions became more favourable.

5.3.2 Irregular Perturbation

At present, it is unknown how irregular vibrations appear. It is likely that they are a combination of subharmonics and harmonics from the nonlinear stiffnesses, and from the cross coupled terms appearing in 5.16 and 5.19. Also, since the driving forces tend to be impulsive, it is possible that the higher harmonics of vibration are excited, confusing the situation even more.

Another alternative is that the system is undergoing chaotic oscillation, though this difficult to show.

5.4. Summary

A mathematical approximation has been introduced extending Titze's treatise on the physics of vocal fold oscillation (Titze, 1983). Subharmonics have been proposed as a possible cause of regular jitter and shimmer perturbations that have appeared in simulations using the proposed model, Isshiki and Ishizaka's model, and in speech data gathered by Monsen (1979).

The equations developed in chapter five indicate that the regular perturbation phenomena is deterministic and implicit in the model. They also illustrate the components of the driving forces discussed in section 3.9. The effects of these terms can be seen in the energy exchange diagrams of chapter four. It is interesting to note that the energy input to overcome damping is provided by the tract loading, and is not implicit in the differential equations for the tissue. This implies that sustained oscillation requires aerodynamic flow.

Finally, we observe that the appearance of perturbations requires a third or higher order system. The proposed model, and that of Ishizaka and Flanagan (1972), are fourth order coupled systems in a form appropriate for sustaining oscillation and producing subharmonics.

CHAPTER 6

SUMMARY, CONCLUSIONS AND RECOMMENDATIONS

In the course of this investigation we have established a hybrid simulation model by building upon existing models. This has been used to successfully investigate asymmetric conditions of vibration. A set of flow coupled differential equations has also been proposed that appears to explain much of the phenomena observed in real speech. The following sections outline the major findings of this study and propose further areas for investigation.

6.1. Summary and Conclusions

The simulation model is a hybrid of the discrete two mass Ishizaka-Flanagan model (1972) and the discrete multiple mass model proposed by Titze (1973, 1974). Proposing such a hybrid enables us to asymmetrically vary and localise biomechanical parameters. The resulting simulations have revealed interesting phenomena such as subharmonics and random perturbations.

The results are presented in terms of phase plots and energy exchange plots which are ideal for examining the oscillatory stability and the displacement dependent energy dissipation-absorption mechanism. Speech, flow, cross-sectional area and displacement plots are also presented so that time dependent relationships can be observed.

Perturbations have been analysed using pitch period duration (ppd) and largest peak amplitude (lpa) plots which attempt to show the changes in period and amplitude of the speech wave over time. These changes may be regular (periodic) or random.

The plots show end point phenomena i.e. the period length and the peak amplitude, but do not adequately represent the effect jitter and shimmer have throughout the entire fundamental period. To display these effects better, the HNR value (Yumoto, Gould and Baer, 1982) has been calculated.

The flow coupled differential equations developed in chapter five extend work done by Titze (1983), unifying the differential equations into a function dependent only on the lateral displacement. By doing this, the various components of the aerodynamic forces are isolated and our understanding of their effects clarified. For example, vertical phasing and vocal tract loading have been incorporated to explain the oscillatory stability in terms of velocity dependent negative damping, Bernoulli effect forces and absorption-dissipation mechanisms. Vertical phasing and loading have previously been discussed by Titze (1980, 1983, 1985), Ishizaka (1981), and Conrad (1983) within their own models, and this analysis reinforces their conclusions and furthermore predicts phenomena that may be the cause of perturbed speech.

Oscillatory instabilities observed in the simulation model have been explained assuming subharmonics of order $1/2$ and $1/3$, which are present in the differential equations. Although the simulation model is not directly derived from the differential equations of chapter five, the energy exchange diagrams produced by the proposed hybrid simulation model (the equations and results of chapters three and four) and by Titze (1983) (which is the result that would be derived from the simulation of the equations of chapter five) are so similar that there is little doubt that the two cases represent the same situation. It should be noted that the $1/3$ subharmonic is inherent in the tissue parameters but the $1/2$ subharmonic appears through the presence of the flow. The manifestation of these subharmonics is dependent on the value of T_{act} , the longitudinal tension, which determines the damping and stiffness in the tissue. Other instabilities which are more random in nature have been observed for localised

parameter changes. The equations do not show how this irregular perturbation arises, but it is likely that it is due to the interaction of the many sub and superharmonics of the fundamental.

The results of the simulations agree well with data from the Isshiki and Ishizaka model, the Voice Lab, and Mosen (1979), which all display both regular subharmonic perturbations and random perturbations. The HNR ratios are higher in the simulation results than for real speech data, probably because the model does not incorporate turbulent flow, random perturbations in parameter values or three dimensional flow characteristics.

We can conclude that the model has successfully simulated the major phenomena present in vocal fold oscillation, demonstrating the flexibility and suitability for modelling pathologies.

The flow coupled differential equations, although not solved for, are in a form recognisable as likely to sustain oscillation, and under certain conditions, produce regular and irregular perturbations. Thus the investigation has succeeded in attaining all the proposed goals, and we feel that the field will benefit greatly from this study since few analysis or simulation studies have been performed in the area of speech perturbations. It is gratifying to observe the predicted phenomena in the real pathological speech data, indicating that this study is on the right track.

6.2. Recommendations

Since this study performed a limited range of simulations much work can be done before it is necessary to advance to a three dimensional finite element model. It is necessary, however, for more scientific data on biomechanical parameters for pathological cases to be made available, especially for changes in mass, stiffness and tension. (we should recognise however that the two mass form of the proposed model

is not an ideal anatomical model of the vocal cords, and consequently some of the recorded parameters are not suitable for use in the mass-spring approach).

Given this data, future work could be attempted in any of the following areas using this model to investigate:

- the effect of the posterior glottic chink
- the random variation of biomechanical parameters using a normal distribution. This would model the effect of nerve innervation and muscle contraction in the fold.
- the incorporation of turbulent flow to model the irregular medial edge when a polyp or carcinooma exists.
- the analytical solution to the proposed differential equations
- the simulation of the set of differential equations proposed in chapter five.

The last recommendation is the most logical as it would demonstrate conclusively the similarity between the hybrid simulation model (and the Ishizaka-Flanagan model) and our equations (derived from Titze (1983)). This would enable us to unify the previous discrete models with the recently proposed analytical models of Titze (1983) and Conrad (1983).

REFERENCES

- Anathapadmanabha, T.V., and Fant, G. (1982). "Calculation of true glottal flow and its components", Speech Transmission Lab. Quarterly Report 1/82: 167-184.
- Aronson, B. (1980). "Clinical Voice Disorders: an Interdisciplinary Approach", Thieme-Stratton,
- Baer, T. (1975). "Investigation of phonation using excised larynges", unpublished PhD thesis, MIT, Cambridge, Mass.
- van den Berg, J.W., Zanema, J.T., and Doornenbal, P. (1957). "On the air resistance and the Bernoulli effect of the human larynx", J. Acoust. Soc. Am. 29: 626-631.
- van den Berg, J.W. (1958). "Myoelastic-aerodynamic theory of voice production", J. Speech and Hearing Res. 1:3, 227-244.
- van den Berg, J.W. (1960). "Vocal ligaments versus registers", Current Problems in Phoniatric Logopedics 1, 19.
- Conrad, W.A. (1983). "Collapsible tube model of the larynx", Proc. in Vocal Fold Physiology, edited by I.R. Titze and R.C. Scherer, The Denver Center for the Performing Arts, Denver: 328-348.
- Davis, S.B. (1976). "Computer evaluation of laryngeal pathology based on inverse filtering speech", unpublished PhD, Univ. Cal. Santa Barbara, also SCRL Monograph 13, Speech Comm. Res. Lab., Inc. Santa Barbara, Calif.
- Fant, G. (1960). "Acoustic theory of speech production", Mouton and Co., S-Gravehenge.
- Farnsworth, D.W. (1940). "High speed motion pictures of human vocal cords", Bell Lab. Record. 18: 23-208.
- Ferrein (1741). cited in van den Berg, J.W. (1958)
- Flanagan, J.L. (1960). "Speech Analysis, Synthesis, and Perception", Springer-Verlag, New York.
- Flanagan, J., and Landgraf, L.L. (1968). "Self oscillating source for vocal-tract synthesizers", IEEE Trans. Audio and Electronics, AU-16: 57-64.
- Gerratt, B.R., and Hanson, D.G. (1985). "Glottographic measures of laryngeal function in individuals with abnormal motor control", paper presented at the Fourth International Conference on Vocal Fold Physiology, June 1985.
- Hayashi, C. (1964). "Nonlinear Oscillations in Physical Systems", McGraw Hill, New York.
- Hirano, M. (1977). "Structure and vibratory behaviour of the vocal folds", Dynamic Aspects of Speech Production, edited by M. Sawashima and F.S. Cooper, University of

Tokyo Press, Tokyo: 13-27.

of the vocal folds in normal and pathological states", Proc. of the vocal folds in normal and pathological states", Proc. Conf. on the Assessment of Vocal Pathology, ASHA Reports 11, American Speech-Language-Hearing Association, Maryland: 11-30.

Hirano, M. (1981). "Clinical Examination of Voice", Springer-Verlag, Wien.

Ishizaka, K. (1980). "Equivalent lumped mass models of vocal fold vibration", Vocal Fold Physiology, edited by K.N. Stevens and M. Hirano, University of Tokyo Press, Tokyo: 231-244.

Ishizaka, K., and Flanagan, J.L. (1972). "Synthesis of voiced sounds from a two mass model of the vocal cords", Bell System Technical Journal 51: 1233-1268.

Ishizaka, K., and Matsudaira, M. (1968). "What makes the vocal cords vibrate?", Proc. 6th Int. Cong. Acoust., Tokyo vol 2: B9-12.

Ishizaka, K., and Matsudaira, M. (1972). "Fluid Mechanical Consideration of Vocal Cord Vibration", SCRL Monograph 8, Santa Barbara, California.

Ishizaka K., Matsudaira, M., and Kaneko, T. (1976). "Input acoustic-impedance measurement of the subglottal system", J. Acoust. Soc. Am., 60: 190-197.

Isshiki, N., and Ishizaka, K. (1976). "Computer simulation of pathological vocal cord vibration", J. Acoust. Soc. Am., 60: 1193-1198.

Kaneko, T., and Ishizaka, K. (1968). "On equivalent mechanical constants of the vocal cords", J. Acoust. Soc. Japan, 24: 312-313.

Kojima, H., Gould, W.J., Lambiase, A., and Isshiki, N. (1980). "Computer analysis of hoarseness", Acta Otolaryngologica 89: 547-554.

Koike, Y. (1980). "Sub and supraglottal pressure variation during phonation", Vocal Fold Physiology, edited by K.N. Stevens and M. Hirano, University of Tokyo Press, Tokyo: 181-188.

Ludlow, C. (1981). "Research needs for the assessment of phonatory function", Proc. Conf. on the Assessment of Vocal Pathology, ASHA Reports 11, American Speech-Language-Hearing Association, Maryland: 3-8.

Meirovitch, L. (1975). "Elements of vibration analysis", McGraw Hill, New York.

Minorsky, N. (1962). "Nonlinear Oscillations", Van Nostrand, Princeton, N.J.

Monsen, R.B. (1981). "The use of a reflectionless tube to assess vocal function", Proc. Conf. on the Assessment of Vocal Pathology, ASHA Reports 11, American Speech-Language-Hearing Association, Maryland: 141-150.

Moore, G.P. (1976). "Observations on laryngeal disease, laryngeal behaviour and voice", Annals of Otolaryngology 85: 553-564.

Morrison, M., Rammage, L., Belisle, G., Pullan, C.B., and Nicol, H. (1983). "Muscle tension dysphonia", Journ. Otolaryngology, V 12, 5:

- Rothenberg, M. (1983). "Source-tract acoustic interaction in breathy voice", in *Vocal Fold Physiology*, edited by I.R. Titze and R.C. Scherer, Denver Center for the Performing Arts.
- Schiller, L. (1922). cited in Ishizaka K., and Matsudaira, M. (1972).
- Stevens, K. (1977). "Physics of laryngeal behaviour and larynx modes", *Phonetica* 34: 264-279.
- Tanabe, M. (1972). "Vibratory pattern of the vocal cord in unilateral paralysis of the cricothyroid muscle", *Acta Otolaryngologica* (Stockholm) 74:339-345.
- Titze, I.R. (1973). "The human vocal cords: a mathematical model, Part 1", *Phonetica* 28: 129-170.
- Titze, I.R. (1974). "The human vocal cords: a mathematical model, Part 2", *Phonetica* 29: 1-21.
- Titze, I.R., and Strong, W. (1976). "On the mechanics of vocal fold vibration", *J. Acoust. Soc. Am.* 60: 1366-1380.
- Titze, I.R., and Talkin, D.T. (1979). "A theoretical study of the effects of various laryngeal configurations on the acoustics of phonation", *J. Acoust. Soc. Am.* 66: 60-74.
- Titze, I.R. (1980). "Comments on the myoelastic-aerodynamic theory of phonation", *J. Speech and Hearing Res.* 23: 495-510.
- Titze, I.R. (1983). "Mechanisms of sustained oscillation of the vocal folds", in *Vocal Fold Physiology*, edited by I.R. Titze and R.C. Scherer, Denver Center for the Performing Arts, Denver: 349-357.
- Titze, I.R., and Haghighi, F.A. (1983). "Simulation of particle trajectories of vocal fold tissue during phonation", in *Vocal Fold Physiology*, edited by I.R. Titze and R.C. Scherer, Denver Center for the Performing Arts, Denver.
- Titze, I.R. (1985). "The physics of flow induced oscillation of the vocal folds", *Annual Report of the Denver Center for the Performing Arts*, No 1.
- Yanagihara, N. (1967). "Significance of harmonic changes and noise components in hoarseness", *J. Speech and Hearing Res.* 10: 531-541.
- Yumoto, H., Gould, W.J., and Baer, T. (1982). "Harmonics to noise ratio as an index of the degree of hoarseness", *J. Acoust. Soc. Am.* 71: 1544-1550.

APPENDICES

APPENDIX A: Vocal Tract Measurement Data

The cross sectional area and length data for the vocal tract in the vowel configuration /a/ are derived from Fant (1960). The sections are variable in length with section 1 beginning just above the glottis and section 10 ending with the lips. The areas and lengths to each section are listed below:

Section	Area (cm ²)	Length (cm)
1	2.6	0.5
2	1.3	1.5
3	4.0	0.5
4	2.1	1.0
5	0.7	2.0
6	1.3	1.5
7	2.2	2.0
8	4.7	2.0
9	8.0	4.5
10	5.0	2.0

APPENDIX B: Harmonics to Noise Ratio

The harmonics to noise ratio (HNR) used in this study is derived from Yumoto, Gould and Baer (1982), in which a measure was proposed for quantifying the hoarseness severity in sustained vowels (hoarseness describes a perceived abnormality of the voice).

According to Yanagihara (1967) hoarse vowels have high frequency noise content, noise in the main formants, and a loss of high frequency harmonics. The degree of hoarseness can thus be assessed by the extent to which noise replaces the harmonic structure. The algorithm used by Yumoto, Gould and Baer (1982) is based on the assumption that sustained vowels consist of periodic components (the harmonics) and additive noise (of zero mean amplitude).

The effects of both jitter and shimmer are included in this algorithm, the deviations from the expected amplitude are assumed to be caused by the noise, so HNR is a direct measure of the shimmer in the signal. The effect of jitter is more subtle, and arises because of the formulation of the HNR.

Consider a speech signal $f(t)$. It can be considered as a concatenation of many waves $f_i(\tau)$ of period τ , where τ is a constant length for all pitch periods. If $f_i(\tau)$ is averaged over n periods, then

$$f_{av}(\tau) = \frac{1}{n} \sum_{i=1}^n f_i(\tau)$$

The noise component, or amplitude shimmer, should then average to zero. This average signal wavelet then possesses the harmonic content of the general speech signal.

The presence of jitter means that the signal is not truly periodic and so τ must be considered very carefully. In Yumoto et al, τ is chosen to be the longest pitch period T_{max} in the signal. Consequently, for the purposes of calculating the

average signal, all shorter duration wavelets are padded with zeros. The average wavelet then does not see any contribution from the short wavelets, but has a small contribution from the longer wavelets (reduced by a factor of n).

The energy content of the average wavelet is the harmonic measurement H used for calculating HNR. Thus

$$H = n \int_0^{T_{\max}} (f_{av}(\tau))^2 d\tau$$

The noise N is measured as the deviation by each wavelet f_i from the average wavelet over T_i . Thus

$$N = \sum_1^n \int_0^{T_i} (f_i(\tau) - f_{av}(\tau))^2 d\tau$$

Integrating to T_i means that the effect of T_{\max} is only apparent for the pitch period that has a period of T_{\max} . By integrating over T_i the contribution due to varying period lengths is incorporated in the noise computation.

In our experience HNR is highly dependent on pitch period estimation and so it is necessary that a highly accurate marking scheme be employed. If the period is estimated to the nearest sampling point errors that appear to be jitter can occur due to poor sampling resolution. Consequently, for this implementation of HNR a simple visual marking scheme using the speech signal is used as an initial estimate, and then a parabolic interpolation of the peak in the autocorrelation wave of the speech signal is used as a more accurate estimate. The HNR measurements using the VGH database have shown that pathologies range from 0 to 20 dB, which is compatible with the simulation results. The difference between these results and those of Yumoto et al. may be attributable to the more accurate period estimation technique.