A STUDY OF VOICE PRODUCTION IN NORMAL AND DYSPHONIC SUBJECTS.

ANDREW W. KELMAN B.Sc.

THESIS SUBMITTED TO THE UNIVERSITY OF GLASGOW FOR THE DEGREE OF DOCTOR OF PHILOSOPHY.

DEPARTMENT OF VOICE PATHOLOGY
VICTORIA INFIRMARY,
GLASGOW

AND

DEPARTMENT OF CLINICAL PHYSICS
AND BIO-ENGINEERING,
11 WEST GRAHAM STREET,
GLASGOW.

MAY 1977

CONTENTS

ACKNOWLEDGEMENTS.

The research described in this dissertation was supported by a grant from the Scottish Home and Health Department (grant number C168). I would like to thank Mr.I.C.Simpson, Consultant, and Mrs. M.T. Gordon, Chief Speech Therapist, both from the E.N.T. Department, Victoria Infirmary, for their foresight in the instigation of the project. I would like to acknowledge the help given to me by the whole Speech Therapy Department, but especially by Mrs. M.T. Gordon and Mrs. F.M.Morton, in the collection of some of the clinical data discussed in CHAPTER 3. The study described in CHAPTER 7 was undertaken with the collaboration of my fellow research student Mr. S. Gatehouse.

I would like to thank my supervisor Professor Sir Andrew Watt Kay, Department of Surgery, University of Glasgow, and also Mr. N.M.Orr, Department of Clinical Physics and Bio-Engineering, for their advice during the investigation.

I would like to acknowledge Dr. J.S.Orr, Dr. J. Kirk and Dr. D.J.Sumner, also of the Department of Clinical Physics and Bio-Engineering, for their helpful suggestions in the content of this thesis.

Finally, I would like to thank Professor J.M.A. Lenihan for his support and encouragement throughout the course of this study.

ABSTRACT

The aims of the research described in this thesis were to investigate some of the factors involved in voice production, and to develop some of the techniques employed, to a stage at which they could be used clinically to assess phonatory function. Available techniques for assessing vocal performance were completely subjective, and, thus, there was a great need to provide objective means of measuring the degrees of vocal dysfunction of the many patients attending the Speech Therapy Department.

In CHAPTER 1, the vocal tract and musculature of the larynx and vocal folds are described. The Myoelastic-Aerodynamic theory of vocal fold vibration is briefly discussed, and a review of some of the relevant literature is given. The instrumentation used in the study is detailed in CHAPTER 2.

The maintenance and control of air flow is of prime importance to voice production, and in CHAPTER 3, aerodynamic measurements are utilized in providing two series of tests which can be used to assess phonatory function. These are now used routinely in the Speech Therapy Department.

The production of voiced sounds depends upon the interruption of the air supply from the lungs by the regular vibration of the vocal folds. Thus the frequency spectra of voiced sounds, and in particular of vowel sounds, consist of harmonic series, with regular frequency components. The frequency analysis of vowel sounds is described in CHAPTER 4, and the presence of abnormal components related to dysphonic conditions. The results obtained from an assessment based on spectral analysis (due to Yanagihara) are compared to those obtained from the aerodynamic assessment developed in CHAPTER 3, with the latter assessment proving of greater value as a routine clinical technique.

The vibratory pattern of the vocal folds is studied in CHAPTER 5 using the non-invasive laryngograph technique. The period of major excitation of the vocal tract, i.e. the closing phase of the vocal fold vibration, is related to certain properties of the spectrum of the laryngograph waveform, namely the spectral gradient and also the presence of fine structure in the spectrum. Abnormal spectral components in the voice spectra are shown to originate from the vocal fold vibratory pattern, and a simple model of the vocal folds is used to explain some of the phenomena.

In CHAPTER 6 the control of a sustained vowel phonation is discussed, and the relationships between air flow rate, sound level and the vibratory pattern of the vocal folds are studied.

In CHAPTER 7 the electromyographic activity associated with the lip musculature, during the production of a consonant - vowel - consonant syllable, is studied. The differences in the outputs obtained using different electrode sites and configurations are explained, and the measured activity attributed to different muscle groups by a process of spatial mapping.

The main results of the experiments are summarized in CHAPTER 8. This research has led to a greater understanding of the processes involved in voice production, and also to the development of objective techniques which can be used directly in the assessment of dysphonic patients.

ILLUSTRATIONS

ABBREVIATIONS

CL.T.	CLOSING TIME OF VOCAL FOLD VIBRATORY PATTERN
CVC	CONSONANT-VOWEL-CONSONANT
EMG	ELECTROMYOGRAPHIC
F_o	FUNDAMENTAL FREQUENCY
F1	FREQUENCY OF FIRST FORMANT
F2	FREQUENCY OF SECOND FORMANT
FR	FLOW RATE
I.P.A.	INTERNATIONAL PHONETIC ALPHABET
ips	INCHES PER SECOND (TAPE SPEED)
L_x	LARYNGOGRAPH OUTPUT
m.dep.ang.oris	MUSCLE DEPRESSOR ANGULI ORIS
m.dep.lab.inf.	MUSCLE DEPRESSOR LABII INFERIORIS
m.lev.lab.sup.	MUSCLE LEVATOR LABII SUPERIORIS
m.ment.	MUSCLE MENTALIS
m.orb.oris	MUSCLE ORBICULARIS ORIS
m.zyg.maj.	MUSCLE ZYGOMATICUS MAJOR
MFR	MEAN FLOW RATE
M/P & MFR/PFR	MEAN FLOW RATE / PEAK FLOW RATE
MPT	MAXIMUM PREDICTED PHONATION TIME
MVC/BVC	MEASURED VITAL CAPACITY / BALDWIN'S VITAL CAPACITY
O.T.	OPENING TIME OF VOCAL FOLD VIBRATORY PATTERN
PHON.VOL.	PHONATION VOLUME
SD/M%	STANDARD DEVIATION/MEAN %
SL	SOUND LEVEL
S.T.P.	STANDARD TEMPERATURE AND PRESSURE
VC	VITAL CAPACITY

CHAPTER 1.

THE VOICE PRODUCTION SYSTEM

CHAPTER I - THE VOICE PRODUCTION SYSTEM

1.1 INTRODUCTION

Man's most important method of communication is speech, although the importance of this function may only become apparent when the ability to produce speech is either impaired or lost completely. Speech consists of the controlled production of series of sounds whose meanings are prescribed according to the particular language, and thus the ability to produce speech is dependent on the capability of producing the wide variety of voiced sounds.

Voice production is dependent upon the action of the respiratory system, the laryngeal and articulatory musculature, the neural system including various feedback loops, and the overall control and coordination provided by the respiratory and speech centres in the brain. Defects occurring in any of these aspects result in alterations in the voice output, and such defects range from congenital abnormalities, e.g. cleft palate, diseases of the respiratory system, neurological diseases, e.g. Parkinson's disease, diseases of the central nervous system, e.g. brain tumour or diseases which are more or less localized to the larynx itself.

In this chapter, a brief description is given of the voice production system, the respiratory system, and the detailed musculature of the larynx and vocal folds. The Myoelastic-Aerodynamic theory of vocal fold vibration is described, and a review of some of the relevant literature is given.

1.2 THE UPPER VOCAL TRACT

The main parts of the upper vocal tract are shown diagramatically in Figure 1.1. The driving force for all voice production (phonation) is provided by the lungs. Air is expelled from the lungs via the bronchi and trachea, and is then forced

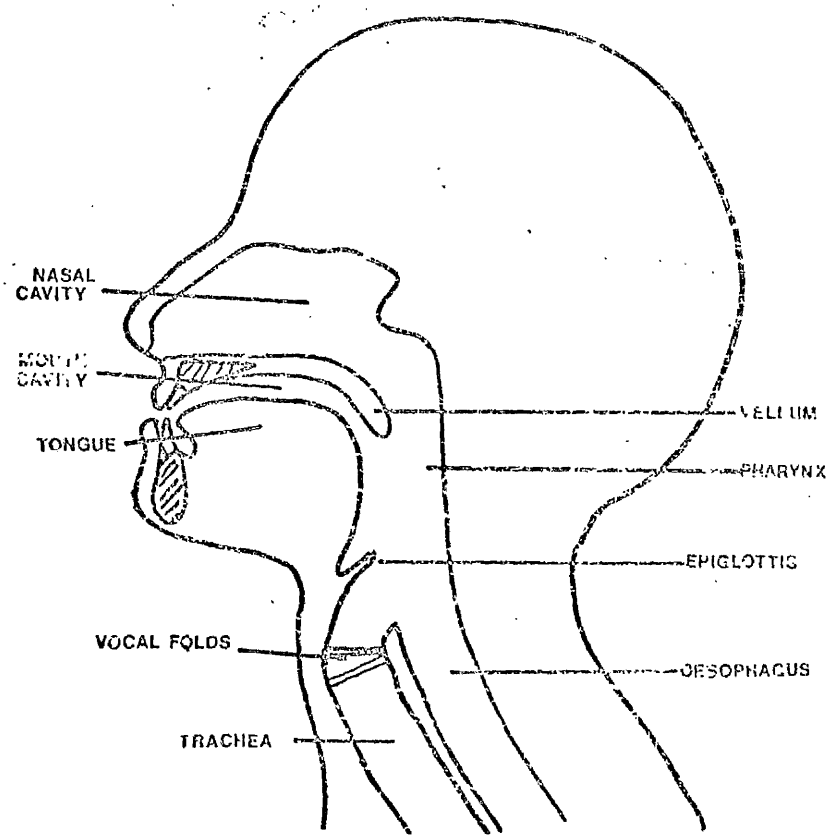


Fig. 1.1. Diagram showing the main regions of the upper vocal tract.

through a constriction at some point in the upper vocal tract, resulting in the production of a sound source. In the case of the bi-labial plosive /p/, (TABLE 1.1) the sound source is produced by the build up of air pressure behind the lips, followed by its sudden release. The sound source for the labio-dental fricative /f/ results from the fricative noise produced when the air is forced through a constriction formed between the upper front teeth and the lower lip. Similarly, other sounds can be categorized according to the site of the constriction and the mechanism involved in the production of the sound source. However, the sounds of interest to the present investigation are the "voiced" sounds, in which the constriction is produced in the larynx by the vibration of the vocal folds. The area of the constriction varies at a frequency commonly in the range 100 to about 300Hz in normal voice production, and the sound source results from the rapid periodic interruption of the air supply from the lungs. This gives rise to a more regular sound in terms of frequency content, than those arising from purely fricative noise. The basic sound source is modified by the resonant properties of the cavities of the upper vocal tract, i.e. the pharyngeal cavity, mouth cavity, nasal cavity. The transmission characteristics of the upper vocal tract can be controlled by altering the size and shape of the various cavities according to the adjustment of the relevant musculature, thus giving rise to the wide spectrum of voiced sounds. The vellum is used to close off the nasal cavity for non-nasal sounds, and other important features include the position of the tongue and the size and shape of the lip opening.

Voice sounds have been categorized in the International Phonetic Alphabet (I.P.A.) according to sound source, tongue position, lip opening size and shape etc., and TABLE 1.1 gives the phonetic symbols and examples of words containing the sounds which are

TABLE 1.1

PHONETIC SYMBOL	KEY WORD
/i/	heed
/I/	hid
/ɛ/	head
/ɑ/	father
/ɔ/	ball
/ʊ/	hood
/u/	who'd
/ʌ/	hud
/ɜ/	heard
/p/	pup
/b/	pub

EXAMPLES OF THE RELEVANT PHONETIC SYMBOLS (I.P.A.) AND WORDS CONTAINING THESE SOUNDS.

of interest to this investigation.

1.3 THE RESPIRATORY SYSTEM

The main function of the respiratory system is to allow gas exchange to take place, i.e. the uptake of oxygen and the removal of carbon dioxide. However, the musculature involved in respiration also takes part in other functions, one of which is phonation. The muscles which control lung volume include the external and internal intercostal muscles, which are attached between adjacent ribs, and also the diaphragm, which is attached around the base of the thoracic cage and separates the thoracic cavity from the abdominal cavity. During quiet respiration, the lung volume is increased for inspiration due to the contraction of the diaphragm, which thus descends, and also the active contraction of the external intercostal muscles, which results in the ribs moving upwards and outwards. During expiration, the diaphragm and external intercostal muscles relax, and the volume of the thoracic cage is decreased by the recoil forces produced by the elastic tissue of the lungs. The diaphragm accounts for about 70% of the tidal volume during quiet respiration (Ganong,1967; Campbell,1974). During forced ventilation the internal intercostal muscles may also contract to move the ribs downwards during expiration. However, it has been shown (Draper et al,1959, Proctor,1974) that during phonation the diaphragm is completely relaxed, apart perhaps for the first few seconds of a phonation initiated at full lung volume, and that the fine control required during phonation is provided by the intercostal muscles.

The respiratory function is under autonomic control from the respiratory centres in the medulla oblongata in the brain. These centres receive information about the carbon dioxide and oxygen tensions in the blood via chemoreceptors situated in various parts of the body (Ganong,1967). There are

also stretch receptors in the lung and the thoracic musculature which relay information to the respiratory centres via the vagus nerve. The respiratory pattern can be modified to a certain extent by the cerebral cortex, and is thus under limited voluntary control, e.g. during phonation, so long as the respiratory requirements are still met.

1.4 MUSCULATURE OF THE VOCAL FOLDS

The larynx consists of a series of cartilaginous rings, the main cartilages being the thyroid cartilage (TC), the cricoid cartilage (CC), the arytenoid cartilages, the corniculate cartilages and the cuneiform cartilages. The cartilages are articulated so as to allow movement between them, and the position of the larynx in the neck is maintained by the external laryngeal muscles. The larynx is attached superiorly to the hyoid bone (HB) which is in turn suspended from the back of the tongue. (Figure 1.2).

The vocal folds are whitish structures situated in the larynx at the level of the thyroid cartilage. The folds are not of uniform construction, but consist of the vocalis muscle covered by a layer called the elastic conus. The elastic conus is thicker near the edge of the vocal fold, and this portion is called the vocal ligament. The ligament is covered by a thin epithelial layer called the mucosa (mucous membrane) (Hirano,1973).

The vocal folds are connected to the cartilages of the larynx at the angle of the thyroid cartilage anteriorly, and to the vocal processes of the arytenoid cartilages posteriorly. The lateral part of each fold and the lateral wall bounding the fold is formed by the thyroarytenoid muscle.

During respiration the vocal folds are relaxed and lie around the laryngeal wall, presenting as little air resistance as possible. In phonation, the position and mechanical properties of the folds

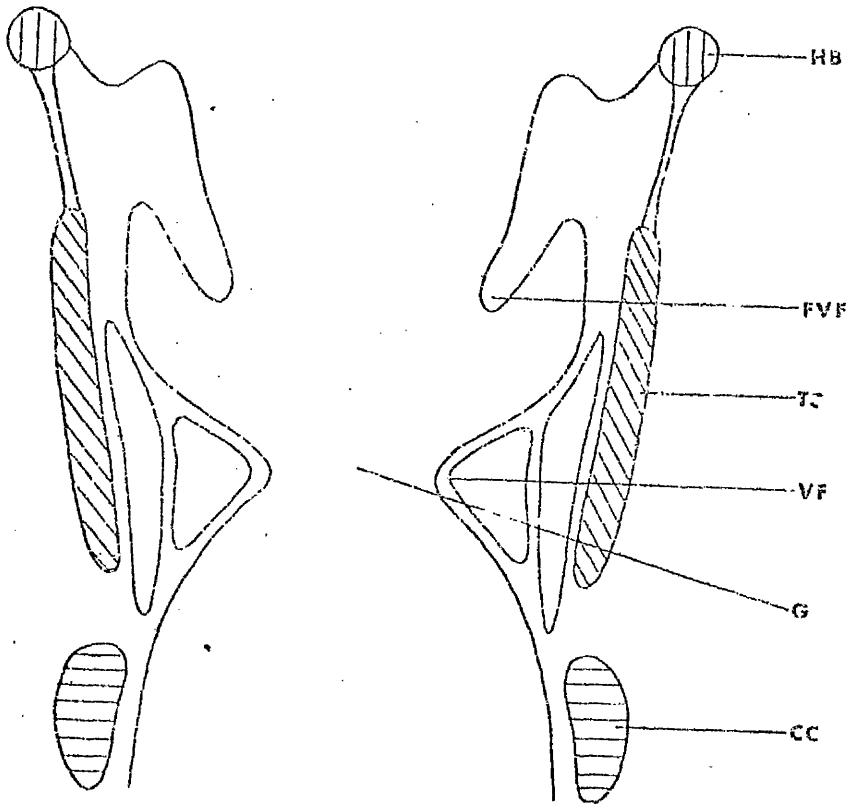


Fig.1.2.
Diagram of coronal section through larynx.

HB - HYOID BONE, FVF - FALSE VOCAL FOLD,
TC - THYROID CARTILAGE, VF - VOCAL FOLD,
G - GLOTTIS, CC - CRICOID CARTILAGE.

are governed by the action of the intrinsic laryngeal muscles, which are named according to the cartilages to which they are attached. The cricoarytenoid and interarytenoid muscles serve to adduct the folds,(i.e. draw the folds into a more central position) just prior to the onset of phonation, this being accomplished by a rotation of the arytenoid cartilages. The properties of mass, tension and elasticity are principally determined by the cricothyroid and vocalis muscles. (Green,1957; Van den Berg,1958). Higher up in the larynx lie the false vocal folds (ventricular folds), but these play little part in normal phonation.

Figure 1.3 shows diagrammatically the vocal folds as seen from above during a normal laryngeal examination, and Figure 1.4 shows a photograph taken during such an examination. The space between the open vocal folds is called the glottis.

The phonatory centres in the brain are situated in the cerebral cortex and control the adjustment of the vocal folds and respiratory pattern during phonation, although the ultimate control of respiration still remains with the respiratory centres in the medulla. The phonatory centres are, in turn, under the control of the motor speech area (Broca's Area) situated in the posterior portion of the sub-frontal convolutions of the cerebrum.

The larynx is innervated by two branches of the vagus nerve, (tenth cranial nerve) the superior recurrent laryngeal nerve and the inferior recurrent laryngeal nerve. The superior branch innervates only the cricothyroid muscle, all the other intrinsic laryngeal muscles being supplied by the inferior branch.

The final voice output is also under the control of several feedback systems. The importance of the auditory system as a feedback control system is demonstrated by the monotonous voice of a person

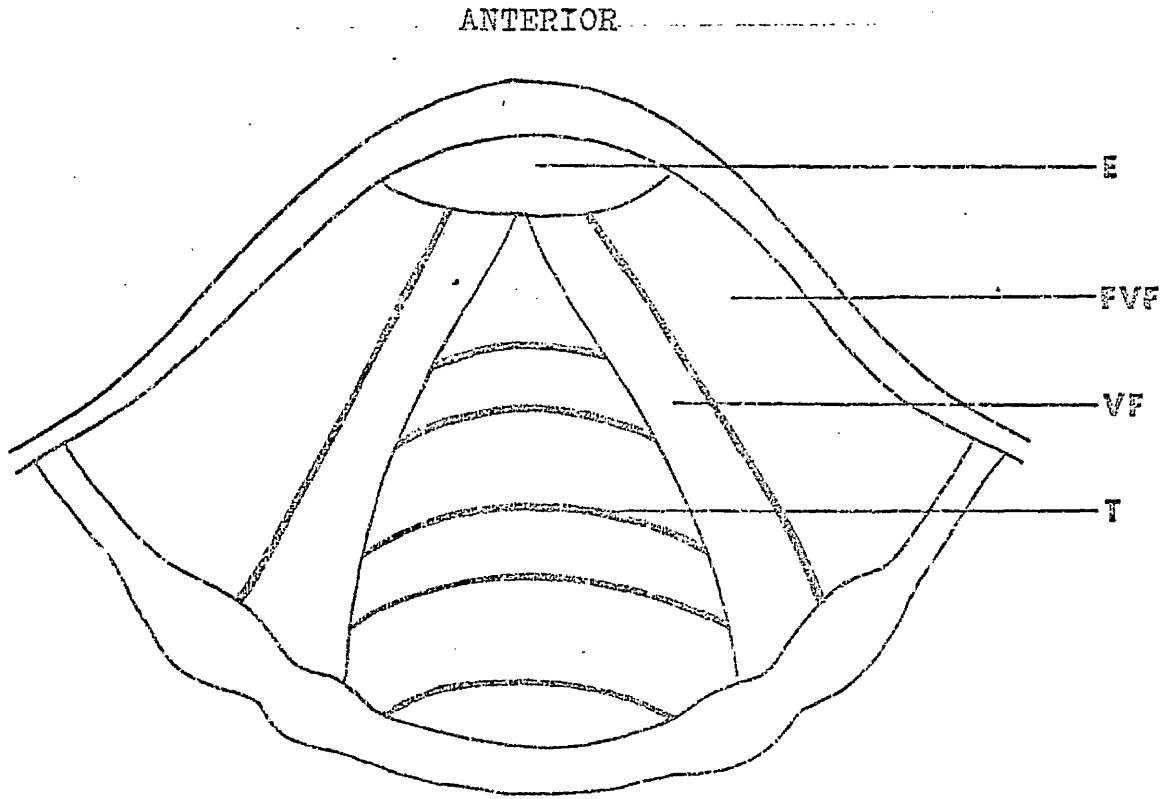
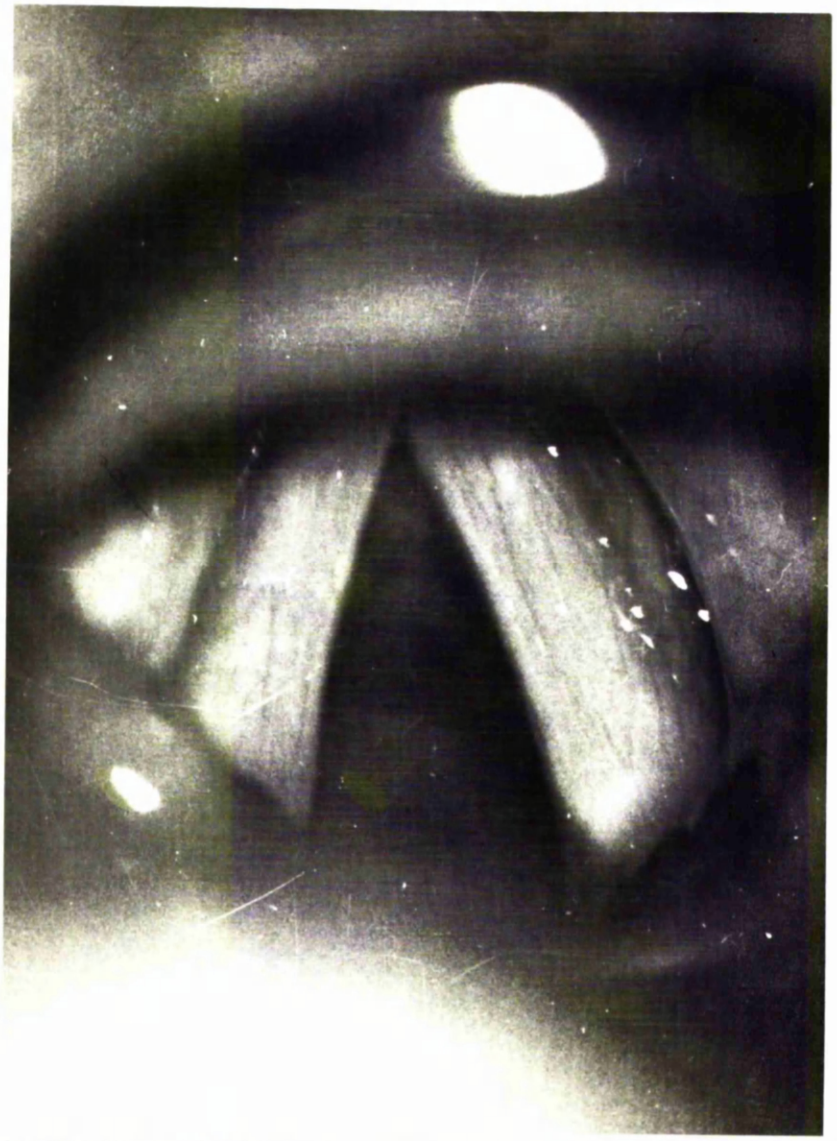


Fig.1.3.

Diagram of vocal folds as seen during laryngeal examination.
E - EPIGLOTTIS, FVF - FALSE VOCAL FOLDS, VF - VOCAL FOLDS,
T - TRACHEA

ANTERIOR



POSTERIOR

Fig.1.4. Vocal folds photographed via a laryngeal mirror.

who is completely deaf, this being due to lack of fundamental frequency or intensity variations, and even a subject who suddenly develops a total deafness will gradually acquire this monotonous voice quality. A possible cause of some cases of stammering is the partial breakdown of the auditory feedback loop, and such cases can be treated using delayed auditory feedback, in which the subject's speech is replayed to him, but with a small time delay introduced. However, the fact that a trained singer can immediately produce a particular note before any auditory feedback can take place, indicates that there exist other monitoring systems which allow for the precise adjustment of the vocal folds and other laryngeal structures,(Kirchner and Suzuki,1968). Kirchner and Wyke (1964,1965) reported finding positional mechanoreceptors in the laryngeal joints in the cat, which when stimulated resulted in reflex contractions and relaxations of the intrinsic laryngeal muscles. Wyke (1974) also reported on stretch receptors within the intrinsic laryngeal muscles, and their affect on the adductor and abductor muscles of the vocal folds. Thus the adjustment of the vocal folds is subject to the feedback control from several feedback loops as well as that of the auditory pathway.

1.5 THE VIBRATION OF THE VOCAL FOLDS

There have been numerous theories suggested as to the cause of the vibration of the vocal folds, ranging from those drawing comparisons with musical instruments (either reed instruments or stringed instruments, hence the name vocal cords instead of vocal folds), to those relying entirely on aerodynamic or elastic forces, and also a theory involving direct neural control over each cycle of vocal fold vibration. This theory, the neuromuscular theory, was proposed by Husson (1953), and according to it, each cycle of vocal fold vibration resulted from the response of the thyroarytenoid muscle to a nerve

impulse from the central nervous system, transmitted via the recurrent laryngeal nerve. Thus the fold vibration was under the direct control of the central nervous system, and the frequency of vibration depended upon the frequency of nerve stimulation. In order to account for the motor nerve fibre absolute refractory period of about 4 msec, and the muscle absolute refractory period of about 1 msec., a polyphasic theory of conduction was postulated, monophasic up to 500Hz, biphasic from 500-1000Hz and triphasic from 1000-1500Hz etc.,(i.e. alternate nerve and muscle groups would only be stimulated at every second or third cycle). It should also be possible for the vocal folds to vibrate in the total absence of air flow through the glottis.

However, the results of Husson could not be repeated by other investigators, (Van den Berg, 1958; Dedo and Dunker,1967). It was shown that electromyographic activity consistent with the frequency of phonation was caused by electrode microphonics,(i.e. the physical vibration of the electrodes) and that thyroarytenoid muscle action potentials did not exceed a frequency of about 50Hz. It was also known that the vocal folds could still vibrate at phonation frequencies even after the recurrent laryngeal nerves had been severed, and it was also shown that the vocal folds could not vibrate in the absence of air flow, (von Leden,1961). Von Leden also reported that his histological and anatomical findings were inconsistent with the neuromuscular theory, and this theory is no longer considered to be tenable.

The current theory describing the vibration of the vocal folds is the myoelastic-aerodynamic theory. According to this theory, the mechanical properties of the folds, e.g. mass, tension, elasticity etc., are governed by the action of the laryngeal muscles (SECTION 1.4) and are continuously monitored by the several feedback loops, (Kotby and Haugen,1970).

Prior to phonation, the vocal folds are drawn into a central position by the laryngeal adductor muscles. At the onset of air flow through the now narrow glottis, the folds are either drawn together or forced apart, and vibrations are maintained by the balance between the aerodynamic forces caused by the air flow, and the tissue forces resulting from the displacement of the folds from their equilibrium position. Figure 1.5 illustrates a schematic section through the larynx. The vocal folds are not thin membranes, but have a thickness of the order of 0.2 - 0.5 cms. Typical values for the dimensions of the glottis are length 1.0 cm and width 0.25 cm, although there is considerable variation among individuals and also within each individual under different vocal adjustments. The force F_1 shown in Figure 1.5 is due to the subglottic pressure generated by the lungs and always acts to separate the folds. F_T represents the force due to the displacement of the vocal folds from their equilibrium position, and can act in either direction, i.e. when the folds are together, F_T acts to separate them, but when they are at their furthest extension F_T acts to pull them inwards. Thus F_T always tends to restore the folds to their equilibrium position. F_b is the aerodynamic force resulting from the air flow through the glottal constriction and always acts to pull the folds together. This force, the Bernoulli force, rises to a maximum at maximum air flow rates and abruptly decreases to zero when the glottis closes and air flow ceases. (Lieberman, 1968).

The dependence of the aerodynamic, Bernoulli, force on air flow can be demonstrated by considering the frictionless flow of an incompressible gas through the system shown in Figure 1.5. Let the subscripts 1, 2 and 3 refer to the subglottic, glottic and supraglottic parts of the system respectively. For any system containing gas flow under

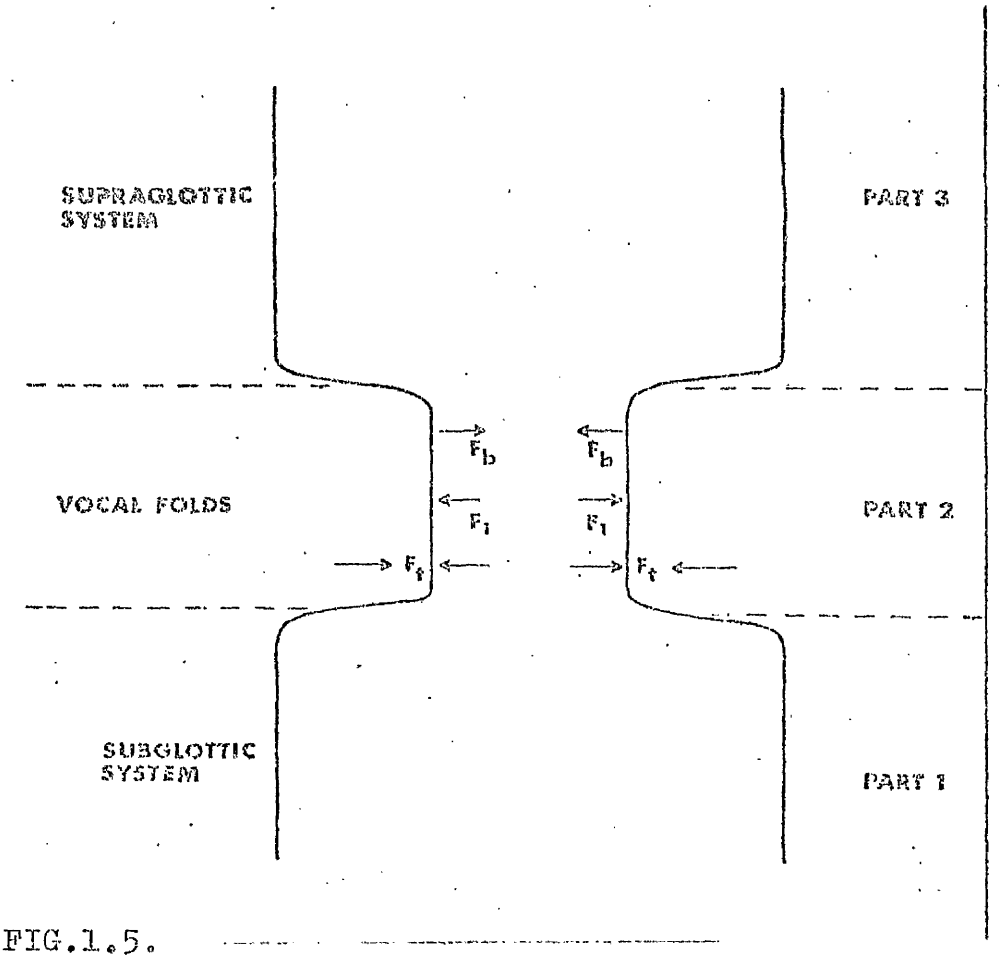


FIG.1.5.

Schematic section through the larynx.

F_T - TISSUE FORCE, F_l - FORCE DUE TO SUB-GLOTTIC PRESSURE,

F_b - BERNOULLI FORCE DUE TO AIR FLOW THROUGH GLOTTIS.

pressure P, with particle velocity v, volume V, and a mass of gas m, the principle of conservation of energy gives

$$\frac{1}{2} mv^2 + PV = \text{constant}$$

If the gas has density ρ, this becomes

$$\frac{1}{2} \rho v^2 + P = \text{constant} = k$$

In particular :

$$\frac{1}{2} \rho v_1^2 + P_1 = k \text{ (subglottic)}$$

$$\frac{1}{2} \rho v_2^2 + P_2 = k \text{ (glottic)}$$

and

$$\frac{1}{2} \rho v_3^2 + P_3 = k \text{ (supraglottic)}$$

where P_3 = atmospheric pressure, and v_1, $v_3 \ll v_2$, because of the small glottal area compared to the area of the trachea. Thus the pressure drop across the constriction, ΔP, is approximated by :

$$\Delta P \doteq \frac{1}{2} \rho . v_2^2 = \frac{1}{2} . \rho . \frac{U^2}{A^2}$$

where U = air flow rate, and A = area of the constriction (glottis). Thus the vocal folds tend to be pulled together by the differential pressure caused by the air flow through the glottis. When non frictionless flow is considered, the pressure drop is modified by a term dependent on the dimensions of the glottis and the viscosity of air, so that

$$\Delta P = k_1 \left(\frac{1}{2} . \frac{U^2}{A^2} \right) + k_2 U ,$$

where $k_1 \doteq 0.9$ is due to turbulent flow and depends on the geometry of the entrance and exit of the constriction, and

$$k_2 = 12 \mu d / l . w^3 ,$$

where the glottis has length l, width w, depth or thickness d, and μ is the viscosity of air. (van den Berg et al, 1957).

The glottal impedance also contains an inductive element, L, due to the mass of air trapped in the

constriction (Flanagan, 1958; Fant, 1970), where:

$$L = \frac{\rho d}{A}$$

Thus the total pressure drop across the glottis is given by :

$$\Delta P = k_1 \left(\frac{1}{2} \cdot \frac{U^2}{A^2} \right) + k_2 U + \frac{\rho d}{A} \cdot \left(\frac{dU}{dt} \right)$$

After making suitable approximations, this equation can be solved to give the air flow velocity function, U, and the results obtained show similar frequency characteristics to functions obtained by a process of inverse filtering of voice output functions.

The myoelastic--aerodynamic theory predicts the dependence of vocal fold vibration on glottal air flow rate and thus subglottic pressure, and glottal impedance (Isshiki, 1964,1965). The importance of the physical properties of the vocal folds results from the dependence on A, which is also a function of time. During phonation, the folds act as vibrators having a double structure, i.e. the main body of the vocalis muscle and vocal ligament, and also the mucous membrane, resulting in a very complex vibratory pattern.

1.6 DYSPHONIA

Dysphonia is a general term encompassing the whole range of vocal disorders. The complexity of the voice production mechanism gives rise to a very large number of different types of dysphonia. Diseases of the respiratory system which produce either airway obstruction or changes in the mechanical properties of the lungs will produce abnormalities in the phonatory pattern, e.g. bronchitis may result in a decrease in the volume of air available for phonation with a corresponding decrease in the time required between breaths or decrease in air flow rate. Neurological diseases affecting either specific nerves and muscle groups or else the brain stem itself will also cause impaired phonatory behaviour, e.g. a lesion of the recurrent laryngeal

nerve can result in the paralysis of one or both vocal folds, with a considerable change in vocal performance. Any conditions disturbing the normal vibratory pattern of the vocal folds will also give rise to noticeable changes in voice function, e.g. carcinoma of the vocal fold, vocal nodule, or any inflammatory condition. Many vocal problems also have psychogenic origins or stem from the general misuse of the voice.

It is therefore important to have effective methods of assessing vocal function. Currently in most clinics, voice assessment relies entirely upon the subjective views of an observer, either clinician or speech therapist, and the terminology used to describe abnormal vocal function is not universally standardised. The vocal folds are still examined using a steady light source, although it is impossible to observe all but the most gross abnormalities in the vibratory pattern, since the vibration frequency is in the approximate range 100-300Hz. Few clinics routinely use a voice synchronised stroboscopic light source to view the vocal folds in motion. Thus there is a need to develop objective means of assessing vocal function, which can be applied rapidly and efficiently in a clinical department and still be non-invasive to the patient.

1.7 AIMS OF THE PRESENT INVESTIGATION

Much of the interest in voice production originated from the development of communications systems, where the fewer the number of parameters which had to be transmitted the greater the number of transmissions which could be dealt with simultaneously. The frequency analyser had shown that many sounds, particularly vowel sounds, had frequency spectra consisting of harmonic series, with superimposed envelope curves, each consisting of a series of peaks, or

formants. The formant frequencies of the vowel sounds were found to be dependent on the particular vowel (Potter et al, 1947; Peterson and Barney, 1952), and the formants were attributed to the resonant properties of the upper vocal tract. (CHAPTER 4). Electrical analogue models were developed to duplicate these resonant characteristics, and this led to an interest in the form of the source function or driving function to be used, i.e. in the vibratory pattern of the vocal folds. Techniques were thus devised which could also be used to study abnormal voice production. The motion of the vocal folds in normal and pathological conditions was described by Timcke (Timcke et al, 1958, 1959; Von Leden et al,1960) after viewing a high speed film of the folds made by Bell Telephone Laboratories. However, this technique is not practicable in a clinical department, and it was only recently that a satisfactory method of routinely studying vocal fold vibration became available (Fourcin and Abberton, 1971), and this technique, using the laryngograph, was used throughout this investigation. This method involves passing an A.C. current between two electrodes placed externally on the neck, and measuring the changes in current as the vocal folds vibrate. The apparatus and technique will be discussed in CHAPTERS 2 and 5.

The respiratory pattern during voice production had also been studied with a view to clinical applications. Isshiki and Von Leden (1964) measured the range of air flow rates for easiest phonation of the sustained vowel /a/ to be 76-182 ml/sec. for a group of normal subjects, whereas the ranges obtained from subjects suffering from pathological conditions exceeded these values. These findings were not unexpected since any condition which resulted in either incomplete glottal closure or any other abnormality in the vibratory pattern of the vocal folds would decrease the efficiency of the system, and it was

likely that this would be demonstrated by a less efficient usage of the air available for phonation. Yanagihara reported on respiratory studies carried out firstly on a group of normal subjects and later extended to include a group of dysphonic patients (Yanagihara et al,1966; Yanagihara and Von Leden, 1967; Von Leden,1968). A linear relationship was found between the volume of air expired during a maximally sustained phonation (phonation volume) and the subject's vital capacity (VC). This, together with the mean flow rate, and the ratio (phonation time / maximum predicted phonation time) was used as a test of vocal performance. The maximum predicted phonation time (MPT) was calculated from the formulae :

$$MPT = \frac{VC}{110} \times 0.67 \qquad \text{(males)}$$

$$MPT = \frac{VC}{100} \times 0.59 \qquad \text{(females)}$$

where the values 110 and 100 ml/sec were the mean flow rates obtained from the normal group, and it was found that the mean values for (phonation volume / vital capacity) were 0.67 and 0.59 for males and females respectively. However, the use of mean values gives rise to the possibility of large errors in the calculated parameters.

The aims of the investigation described in this thesis were several. The standard methods of assessing vocal performance were totally subjective, and the objective tests which were available did not appear to be wholly satisfactory. However, the importance of the control of air flow during phonation was evident, and it was intended to devise objective methods of assessing vocal function based on aerodynamic measurements. (CHAPTER 3).

A classification of hoarseness based on the presence of abnormal components in the frequency spectra of vowel sounds had been suggested by Yanagihara (1967). The principal abnormality ob-

served in such spectra was the presence of random noise components. In the present investigation, the frequency spectra resulting from many dysphonic conditions were studied, and the abnormalities in the spectra related to the vocal function. (CHAPTER 4). The laryngograph technique was used to study the vibratory pattern of the vocal folds both in normal voice production and also in dysphonic patients. (CHAPTER 5). The relationships involved in the control of sound level with respect to air flow rate and the vocal fold vibratory pattern were also to be investigated. (CHAPTER 6). The use of surface electrodes in the study of the electromyographic activity of the lip musculature was examined, and the results compared with those obtained by Leanderson and other experimenters (Leanderson et al 1967,1971; Leanderson and Lindblom, 1972) who used the more invasive technique of using needle electrodes. (CHAPTER 7).

1.8 SUMMARY

In the present chapter, the anatomy of the vocal tract was described. The respiratory system and relevant musculature was described, as was the structure of the vocal folds and their relationships with the tracheal cartilages. The mechanism involved in the vibration of the vocal folds was discussed and the importance of the role of air flow stressed. The aims of the present investigations were stated.

CHAPTER 2.

INSTRUMENTATION USED IN THE STUDY.

CHAPTER 2. - INSTRUMENTATION USED IN THE STUDY.

2.1 INTRODUCTION

The voice production mechanism was briefly described in CHAPTER 1, and from this, the importance of the various aspects of the system became evident. The investigation can be conveniently divided into the following sections.

(a) An assessment of the function of the respiratory system in general, and, in particular, during phonation,

(b) an assessment of the voice output in terms of frequency analysis,

(c) a study of the vibratory pattern of the vocal folds, and

(d) a study of the electromyographic (EMG) activity associated with particular muscle groups during certain speech gestures.

In this chapter, the instrumentation used to measure breath flow patterns, to record and analyse the voice in terms of frequency, and to obtain the vibratory pattern of the vocal folds, will be described. The details of a method of obtaining a continuous record of the fundamental frequency of phonation is also given. The system used to record and analyse the EMG activity will, more appropriately, be described in CHAPTER 7.

2.2 BREATH FLOW MEASUREMENT

The control of air flow and optimal usage of the available air is of crucial importance to satisfactory voice production. Fluctuations in air flow rate can occur very rapidly during phonation, so that the instrument used to measure the breath flow pattern must have a correspondingly short reaction time. Thus a pneumotachograph respirometry system (Mercury Electronics) was used throughout the investigation to measure the air flow pattern. This type of system is based on the principle that a flow of

air passing through a flow resistance causes a pressure differential to be developed across the resistance. The flow resistance consisted of a fine mesh housed in a plastic flowhead. The flowhead was designed to maintain laminar air flow and under these conditions, the pressure drop generated across the resistance was directly proportional to the air flow rate. The pressure drop was sensed using a differential micromanometer and an electrical output obtained which was also proportional to the air flow rate (in either direction) through the flowhead. The electrical output was integrated with respect to time to give the volume of air passing through the flowhead in a given time.

The flowhead was mounted in a rubber face mask which the subject wore during the recording of air flow rate. The face mask provided an airtight seal but did not inhibit the movement of the facial musculature, and the resistance to air flow of the flowhead was negligibly small, allowing the subject to breathe and phonate in a natural manner. The face mask also incorporated a small microphone (Amplivox) which permitted the recording of the phonations produced during the recording session.

The respirometry outputs corresponding to air flow rate and the related volumes were recorded on a four channel ink jet recorder (Mingograf, 34T) which had a flat frequency response up to 700Hz, and was run at a paper speed of 1 cm/sec. Examples of the types of outputs available are shown in Figure 2.1. The recordings illustrated in Figure 2.1 A, B were obtained during a period of quiet respiration, with the flow rate recordings showing expiration upwards and inspiration downwards. In A, the respirometer integrators were reset automatically after each stage of the respiratory cycle, and this type of display allowed the measurement of the tidal volumes, i.e. the volumes of air imspired and expired per cycle.

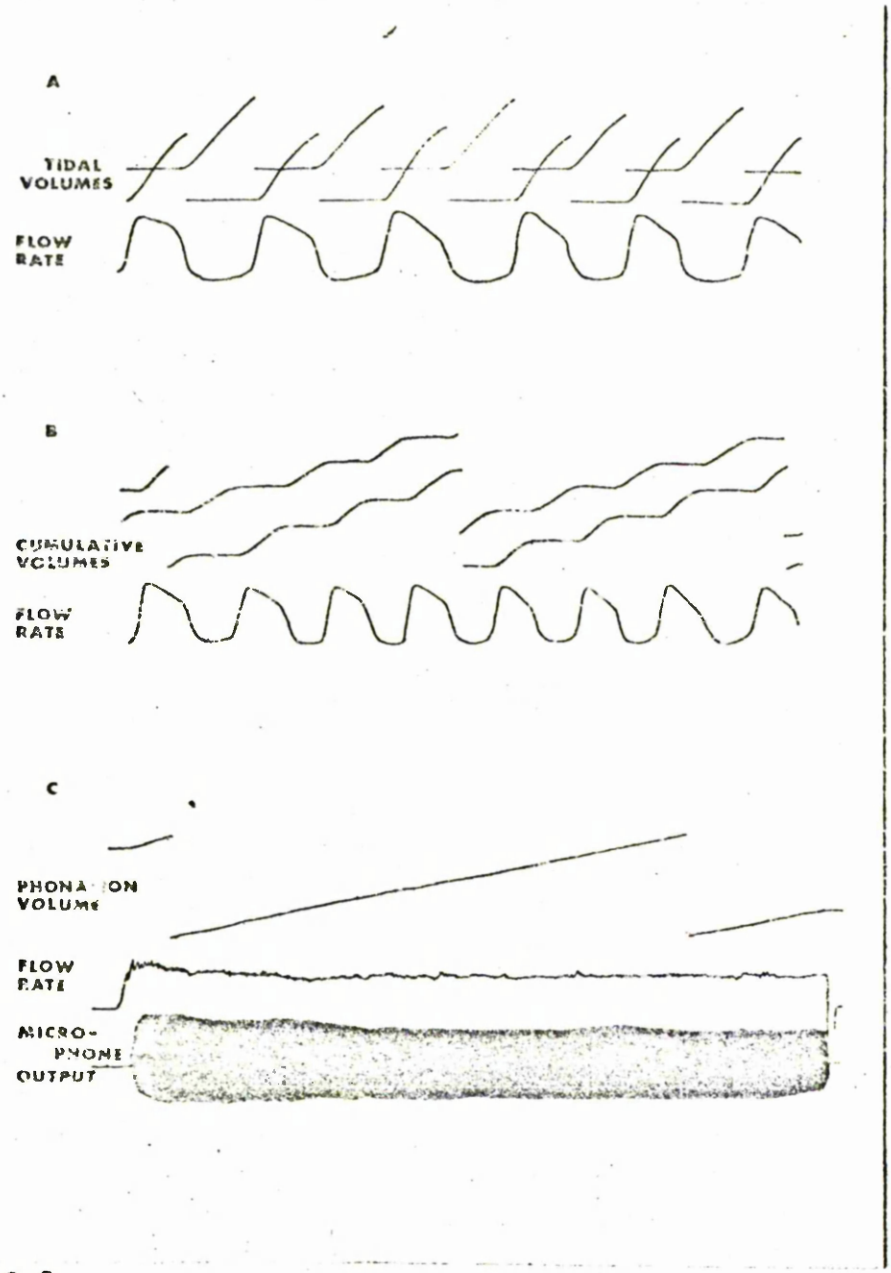


Fig.2.1.
Examples of different types of respirometer outputs.

A - TIDAL VOLUME OUTPUT OBTAINED DURING QUIET RESPIRATION.

B - CUMULATIVE VOLUME OUTPUT OBTAINED DURING QUIET RESPIRATION.

C - OUTPUT OBTAINED DURING THE SUSTAINED PHONATION OF A VOWEL SOUND.

In B, the respirometer was used in the "cumulative" mode, with the integrators storing the results of successive cycles and only resetting after a preset value had been reached, producing the typical "step" output. This facility was used in the measurement of the volume of air expired during the sustained phonation of a vowel sound (phonation volume), Figure 2.1 C, with the air flow rate and microphone outputs being recorded simultaneously.

Each air volume and flow measurement was reduced to dry air at standard temperature and pressure (S.T.P.) using the combined gas law : $\frac{PV}{T}$ = constant. The inspired air was assumed to be 50% saturated with water vapour, so that

$$V_{STP} = \frac{273}{(T + 273)} \cdot \frac{\left(P_o - \frac{1}{2}P_{H_2O}\right)}{P_o} \cdot V_M = k_T \cdot V_M$$

where V_M is the volume of air measured by the respirometer, T^oC is the ambient room temperature, P_o is atmospheric pressure, and P_{H_2O} is the saturated water vapour pressure at that temperature. The expired air was assumed to be fully saturated with water vapour and to be at body temperature of 37^oC, so that

$$V_{STP} = \frac{273}{310} \cdot \frac{\left(P_o - P_{H_2O}\right)}{P_o} \cdot V_M = 0.826\ V_M.$$

The correction factors, k_T, are tabulated in TABLE 2.1. Each flowhead was regularly calibrated using a Parkinson-Cowan gas meter.

2.3 VOICE RECORDINGS

All recording sessions took place inside a sound reduced booth to eliminate extraneous background noise. The recordings made using the microphone incorporated in the face mask were unsuitable for

TABLE 2.1

AMBIENT ROOM TEMPERATURE. T(°C)	INSPIRATION CORRECTION FACTOR k_T,	EXPIRATION CORRECTION FACTOR
14	0.944	0.826
16	0.935	0.826
18	0.929	0.826
20	0.921	0.826
22	0.913	0.826
24	0.905	0.826
26	0.897	0.826
28	0.890	0.826
30	0.882	0.826

Correction factors used to reduce measured air volumes to dry air at S.T.P.

quantitative frequency analysis because of the frequency characteristics of the microphone, and also because of the effect of the face mask on the produced sounds. Thus a recording system having a flat frequency response in the range of interest was required.

The microphone used was a Bruel and Kjaer (B. and K.) ½ inch condenser microphone (type 4131) coupled to a B. and K. microphone amplifier (type 2615), each having a flat frequency response up to 20KHz. The microphone was maintained at a constant distance from the subject's lips (20 cm.) (except for the experiments described in CHAPTER 6), and the output was recorded using a Ferrograph stereo tape recorder (type 722), which had a flat frequency response in the range 50Hz to 15KHz when run at a tape speed of 19 cm/sec. (7½ i.p.s.).

The recording system was calibrated with respect to sound level by using a B. and K. pistonphone (type 4220) which produced a tone at 124 dB. The recorded calibration tone could be altered in either 1 or 10 dB steps using the attenuator network of the microphone amplifier. A record of the sound level could be obtained by replaying the tape recording through a B. and K. level recorder (type 2305). In experiments described in CHAPTER 6, an electrical output proportional to the sound level was obtained by using the B. and K. analogue read-out device (type ZR0021), which was coupled to the pen assembly of the level recorder. This permitted the sound level to be displayed simultaneously with air flow rate via the Mingograf ink-jet recorder. The recording system had a dynamic range of 50 dB.

2.4 VIBRATORY PATTERN OF THE VOCAL FOLDS.

It is important to be able to obtain routinely information about the nature of the vocal fold

vibration, since abnormalities in this pattern usually result in abnormalities in the eventual voice sound.

There are several methods of obtaining details of the vibratory pattern of the folds. One such method involves the technique of high speed cine photography, where the motion of the folds is filmed via a laryngeal mirror at a speed of the order of 5,000-10,000 frames/sec. Such a film is then analysed frame by frame to give measurements of vocal fold length, or separation, or glottal area. However, since the subject has to endure a laryngeal mirror, phonation produced during this procedure can not be considered to be typical, and the quantity of data generated makes this method unsuitable in a clinical situation. Another technique involves the trans-illumination of the glottis and the measurement of the amount of light transmitted by using a photo-electric cell, but variations in the thickness of the vocal folds throughout a cycle and, therefore, in their light transmission properties, makes calibration extremely difficult. If the frequency transmission characteristics of the vocal tract are known, then by a process of inverse filtering of the voice output, the waveform of the original source function can be obtained. However, the transmission properties of the vocal tract are seldom available, and certainly not on a real-time basis. Thus most methods are subject to the disadvantages of either gross uncertainties (e.g. inverse filtering technique) or involving the subject in having to endure a laryngeal mirror or fibre optic bundle, and so probably producing an atypical phonation.

The method used in this investigation to obtain the vibratory pattern of the vocal folds was the laryngograph technique (Fourcin and Abberton,1971). This technique involved connecting a constant A.C. voltage source across the neck, and measuring the

change in current passed as the amount of contact between the vocal folds varied throughout the vibratory cycle.

Two planar electrodes were placed externally on the neck, one on each side of the larynx at the level of the vocal folds, and a high frequency current (1MHz) was passed between them. As the vocal folds opened and closed, the current passed between the electrodes varied, and this variation was measured. During the closing phase of the vocal fold vibration, the area of contact between the folds was increasing, with a corresponding decrease in the electrical impedance between them, and an increase in the current. During the opening phase, the change in area of contact was reversed and corresponded to a decrease in the current passed between the electrodes. Figure 2.2 illustrates diagrammatically the form of the waveform obtained from the laryngograph, with closure of the vocal folds being measured upwards. The different phases of the laryngograph waveform (Lx) and its interpretation will be discussed in CHAPTER 5.

The electrodes used were circular disc electrodes of 1.5 cm. diameter, surrounded by a guard ring which was maintained at earth potential. Several sizes and shapes of electrodes were tried, but the geometry was found not to be critical. Various materials were also tried for the electrode surface. Copper and silver plating were initially used, but in many cases, spurious electrical pick-up (mainly at mains frequency of 50Hz) dominated the signal obtained. In order to improve electrode contact, electrode jelly was used, but this resulted in a rapid corrosion of the electrode surfaces. The final choice of electrode surface was gold plating, and only rarely was it found necessary to revert to the use of electrode jelly.

The position of the electrodes was found to be important, and care was taken to place them on the larynx at the point where the Lx signal was maximal. The Lx signal was recorded on the second channel of

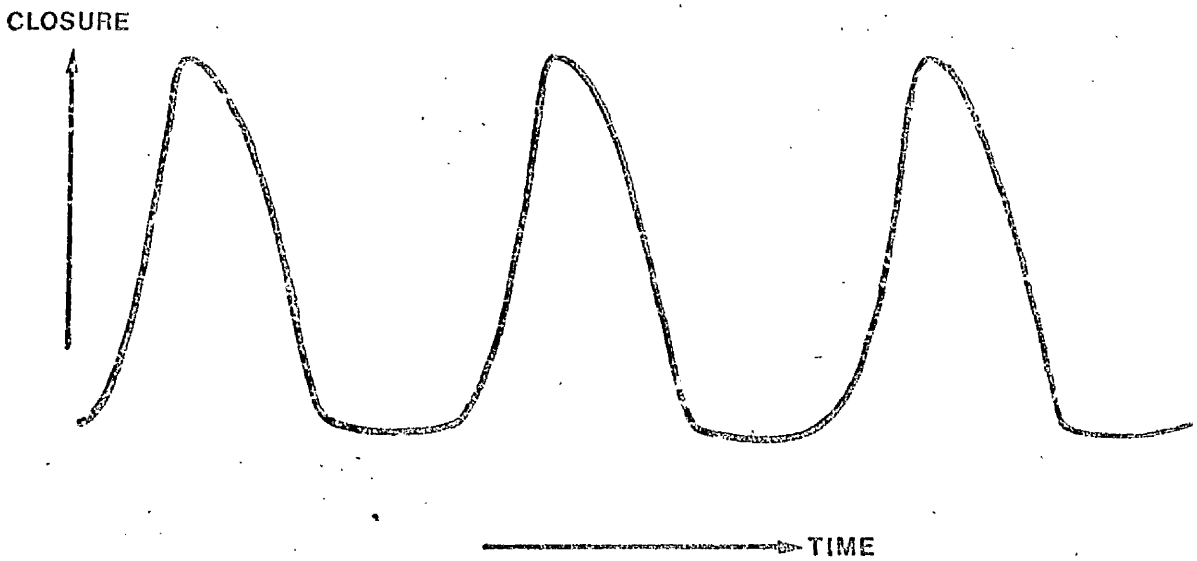


Fig. 2.2. Idealised example of type of waveform obtained from laryngograph (L_x).

the Ferrograph stereo tape recorder.

This technique was completely non-invasive to the subject, and since it did not involve simultaneous visual examination of the larynx, could be used to investigate vocal fold vibration under fairly natural conditions, even during continuous speech.

2.5 MEASUREMENT OF FUNDAMENTAL FREQUENCY.

One of the factors which makes a normal voice more interesting to listen to, (compared to the monotonous quality of the voice of a typical deaf subject) is the rapid changes in the fundamental frequency which occur. Many subjects suffering from pathological conditions of the larynx or vocal folds show a more limited range in the fundamental frequencies which they can reach. Thus the fundamental frequency of vocal fold vibration is an important parameter in voice production, and the Lx waveform readily lends itself to its measurement. During the recording of phonation sustained at a constant fundamental frequency, the Lx signal was fed into a digital frequency counter (Racal, 950) and the fundamental frequency was read directly. However, for the later experiments (CHAPTER 6) it was required to be able to record the fundamental frequency continuously and to display it simultaneously with air flow rate and sound level using the Mingograf recorder.

The circuit devised to obtain a D.C. electrical signal directly proportional to the fundamental frequency of phonation is shown in Figure 2.3. The Lx input was buffered by amplifier A1 to prevent loading or interference with the other recording instruments. The large D.C. component which was always present on the Lx signal was removed in the next stage, and the A.C. component amplified (A2). Since any spurious noise which was present occurred on the flat base of the Lx signal, this was eliminated by removing the

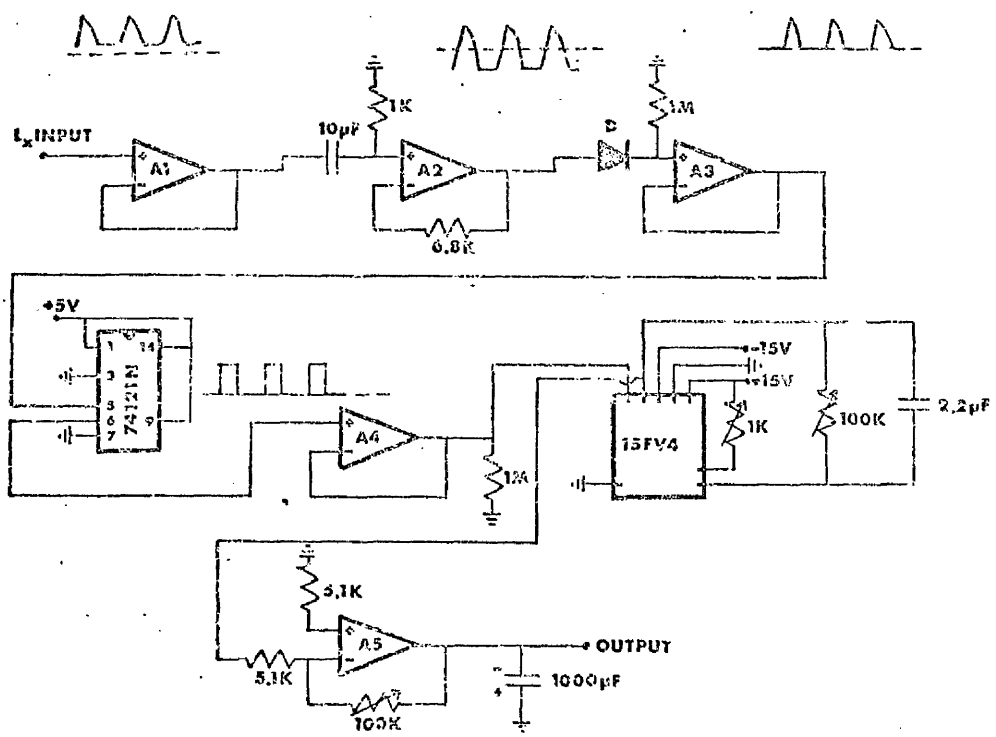


Fig.2.3. Diagram of frequency - voltage converter circuit (A1 - A5 741 Operational amplifiers).

negative part of the cycle using diode, D. The positive peak waveform was used as the Schmitt trigger input of the integrated circuit 74121N, which gave a series of positive square pulses as output. This series of pulses acted as the input to the frequency - voltage converter module (Ancom, 15 FV4), which gave as output a voltage proportional to the frequency of the incoming pulses. This D.C. output was amplified to a suitable level, A4, and any high frequency ripple smoothed out by the 1000μF electrolytic capacitor. The output was linear up to at least 500Hz, and could be recorded on the Mingograf recorder simultaneously with flow rate and sound level. The system was regularly calibrated using a signal generator.

2.6 FREQUENCY ANALYSIS

Since the initial voice source is produced by the regular vibration of the vocal folds, the frequency spectrum of a voiced sound consists of a fundamental and a series of harmonically related components. Abnormalities in the frequency spectra of the voice output can be used in an assessment of vocal function (Yanagihara,1967), and the technique can also be used to investigate abnormalities in the vibratory pattern of the vocal folds (CHAPTER 5).

The instrument used in this investigation to carry out frequency analysis was a Sonagraph (Kay Elemetrics, type 729A) with additional modules, Scale Magnifier Module (type 6076C) and Contour Display Module (type 6070A).

The signal to be analysed was re-recorded from a tape recorder onto a magnetic drum within the Sonagraph, which acted like a continuous tapeloop system and allowed any part of the recording to be analysed. In the "replay" mode, the magnetic drum rotated at 300 r.p.m., and the output obtained at the replay head was passed through on electronic filter whose centre frequency scanned the range of

interest, due to a mechanical screw assembly. The output of the tuned filter was used to control the marking of electrically sensitive paper by a stylus. The filter had two bandwidth settings, a "broad" bandwidth of 300Hz and a "narrow" bandwidth of 45Hz. The standard frequency ranges available were 0-4KHz. and 0-8KHz., although the Scale Magnifier Module permitted any frequency band to be expanded for closer examination, and also gave an output of overall signal intensity. Examples of some of the displays available are shown in Figure 2.4, the sound analysed being the vowel /a/. The frequency range 0-4KHz was used, although in A and B only the range 0-2.5KHz is shown. In Figure 2.4A, the 300Hz bandwidth filter was used, and the harmonic components were just evident as horizontal dark bands. The degree of darkness of the bands was indicative of the signal intensity in the particular frequency range. In this example, the signal intensity was greater in the range below about 1.25KHz. The lower portion of Figure 2.4B shows the same output display as A, but was obtained using the 45Hz bandwidth filter. The harmonic structure was more obvious than in A. The upper portion of B illustrates the intensity of the overall signal level, which was seen to remain fairly constant throughout the period of the recording. Figure 2.4C illustrates the "sectioner" display output, where the spectrum at any point in the recording could be obtained. The frequency scale was inverted to allow this display to be examined more easily, and in this case the horizontal lines represented the harmonic components, with their lengths indicating the signal intensity (in dB).

The Contour display module could be used to make the original displays more quantitative, and this is illustrated in Figure 2.5. In this output

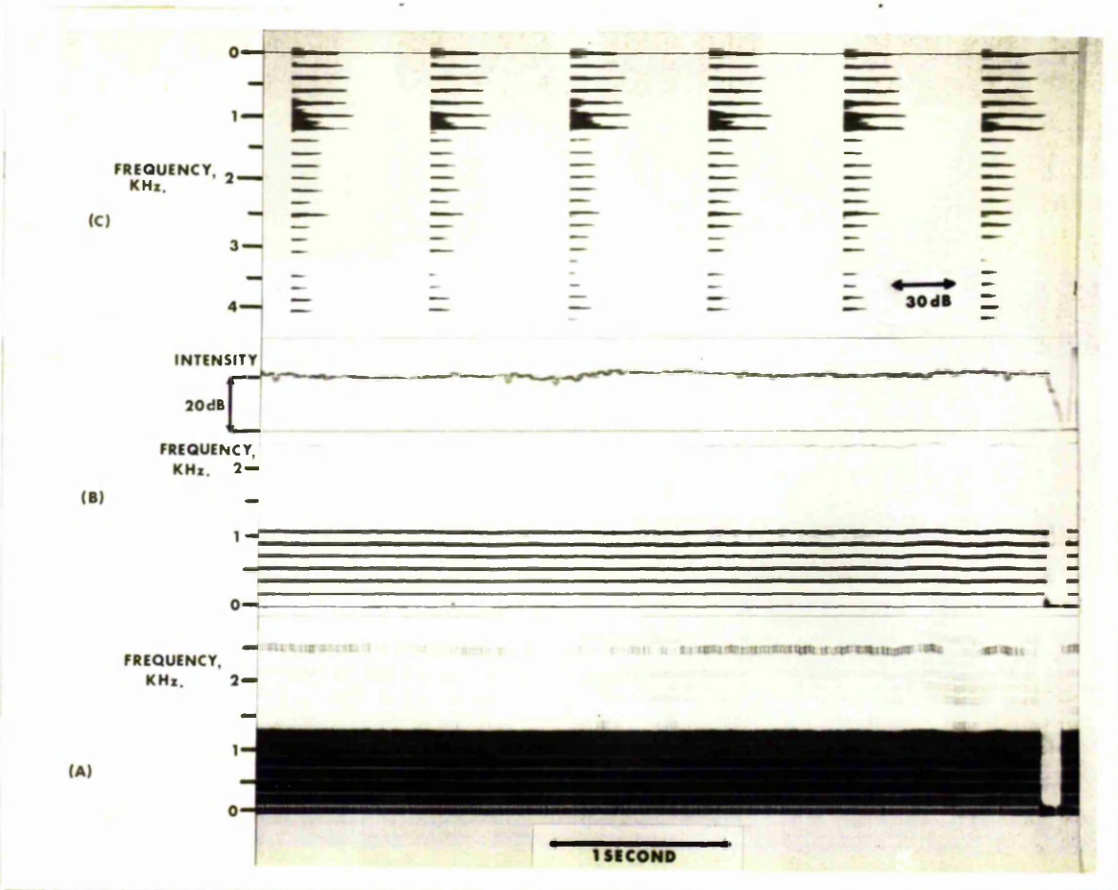


Fig. 2.4. Examples of different types of sonagraph displays obtained for vowel /a/ and demonstrating harmonic nature of the frequency spectrum of such sounds.

A - BROADBAND FILTER (300Hz BANDWIDTH)
B - NARROWBAND FILTER (45Hz BANDWIDTH) with superposition of intensity display.

C - SECTION DISPLAY USING 45Hz FILTER.

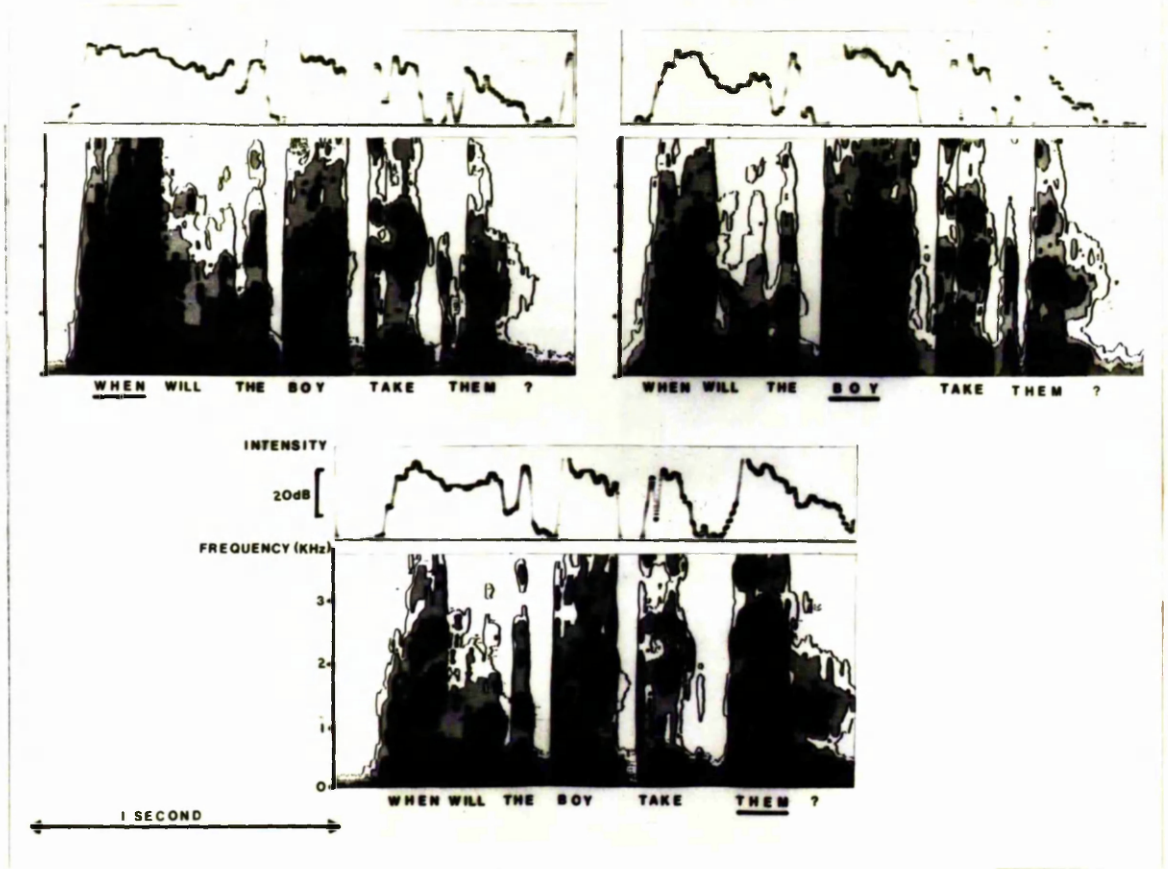


Fig. 2.5. Examples illustrating the contour display sonagraph output. In each case the stressed syllable is underlined.

display, the 300Hz bandwidth filter was used and each colour gradation represented a 6dB change in signal intensity. In the example shown, the underlined word was given greater stress, and the resultant increases in intensity and times for the stressed syllables are seen in the figure.

The calibration of the Sonagraph in terms of frequency and intensity was regularly carried out. A two channel Sonagraph was available for some of the analysis described in this thesis. Standard pieces of laboratory equipment were also available.

2.7 SUMMARY

The major pieces of specialised instrumentation used throughout the investigation were described in this chapter. These included the respirometry system for measuring breath flow patterns during quiet respiration and also during phonation, the system for recording and analysing the voice output and the vibratory pattern of the vocal folds. Examples were given to illustrate the various types of output available from the equipment. The recording and analysis of the EMG signals obtained from the lip musculature will be fully discussed in CHAPTER 7.

CHAPTER 3.

AERODYNAMIC ASSESSMENT OF PHONATORY FUNCTION.

CHAPTER 3. AERODYNAMIC ASSESSMENT OF PHONATORY FUNCTION

3.1 INTRODUCTION

The role of air flow through the glottis in the maintenance of the vibration of the vocal folds has already been discussed (CHAPTER 1). The efficiency of the phonatory process is therefore largely dependent on the efficiency of the control mechanisms of some of the muscles involved in respiration, e.g., if a subject's respiratory pattern is irregular due to lack of control of the intercostal muscles, he is also likely to produce an uncontrolled phonation, since phonation is dependent on the fine control of these muscles. In this chapter, a series of tests is developed to allow a subject's respiratory control to be assessed, according to his ability to reproduce a particular air flow pattern throughout several cycles of quiet respiration. A second series of tests combines the results of aerodynamic measurements made during the recording of a maximally sustained vowel, and also certain lung parameters, to give an assessment of the subject's capability to produce this type of phonation. The results of these tests are compared to those obtained using the standard, subjective speech therapy techniques. The results obtained by Issiki, Von Leden, and Yanagihara were discussed in CHAPTER 1.

Figure 3.1 shows diagramatically the partition of lung volumes. The vital capacity is the maximum amount of air which can be expired following a maximum inspiration, or the amount of air which can be inspired following a maximum expiration. The residual volume is the volume of air remaining in the lungs after a maximum expiration. The vital capacity gives an upper limit to the volume of air which can be used in the performance of an exercise, e.g. during

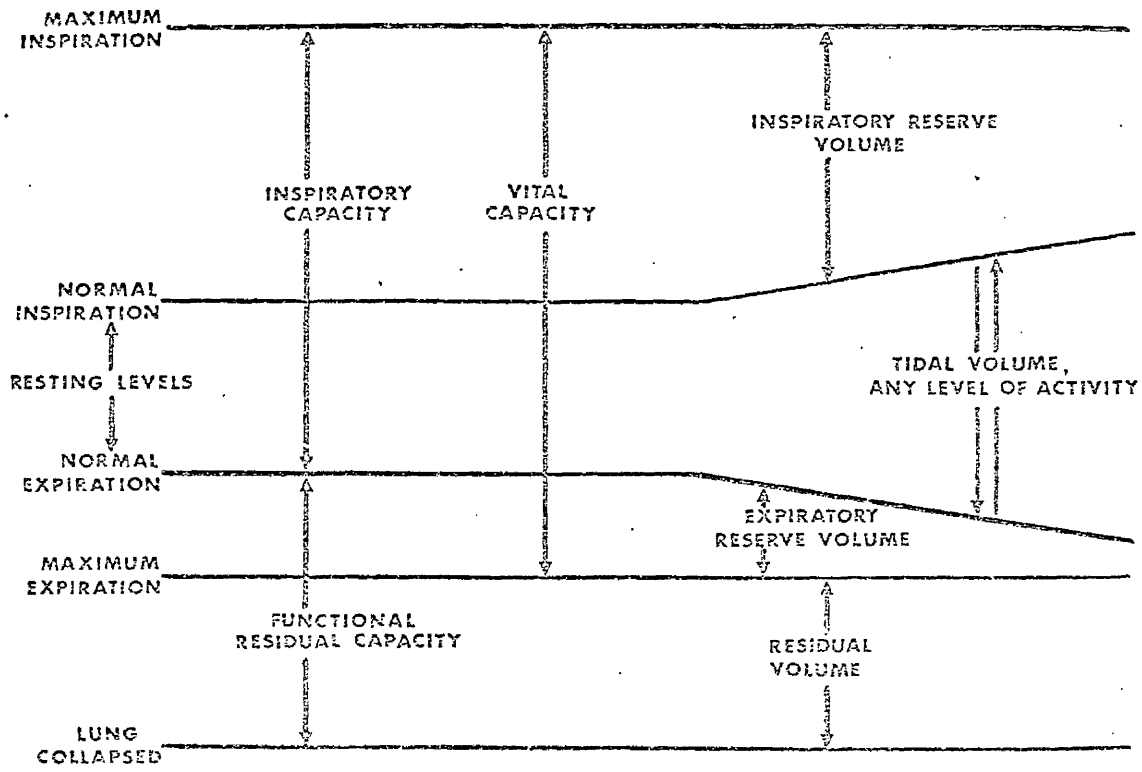


FIG.3.1 DIAGRAM ILLUSTRATING THE PARTITION OF LUNG VOLUME.

a maximally sustained phonation. The tidal volumes are the volumes of air inspired and expired per respiratory cycle, and are usually larger than the resting values. However, the absolute values of tidal volumes give no indication of the control of respiratory function.

3.2 EXPERIMENTAL PROCEDURE

The experimental system was housed inside a sound reduced booth, and is shown schematically in Figure 3.2. The instrumentation was fully discussed in CHAPTER 2.

The aim of this part of the investigation was to develop objective means of assessing respiratory control and phonatory efficiency based on aerodynamic measurements. Two groups of subjects were taken, i.e. a control group of adult subjects consisting of hospital staff and other volunteers who had no history of voice disorders, and a dysphonic group consisting of a series of patients who were suffering from various types of dysphonia, and were attending the speech therapy department. Each subject was independently assessed as to breathing function and phonatory capability according to orthodox speech therapy techniques. These consisted largely of visual inspection, and also the measurement of the maximum time for which the subject could sustain the voiceless fricative sound /s/ or the voiced fricative sound /z/, although no allowance was made for sound level or air flow rate. Each subject was thus subjectively placed in one of the four categories: normal, slightly abnormal, abnormal, and highly abnormal, for both respiratory function and phonatory capability.

During the first part of the recording session, a series of sustained vowels was recorded using the condenser microphone, while the vibratory pattern of the vocal folds was obtained by means of the laryngograph. These recordings were intended for frequency analysis and will be discussed in CHAPTER 4. The

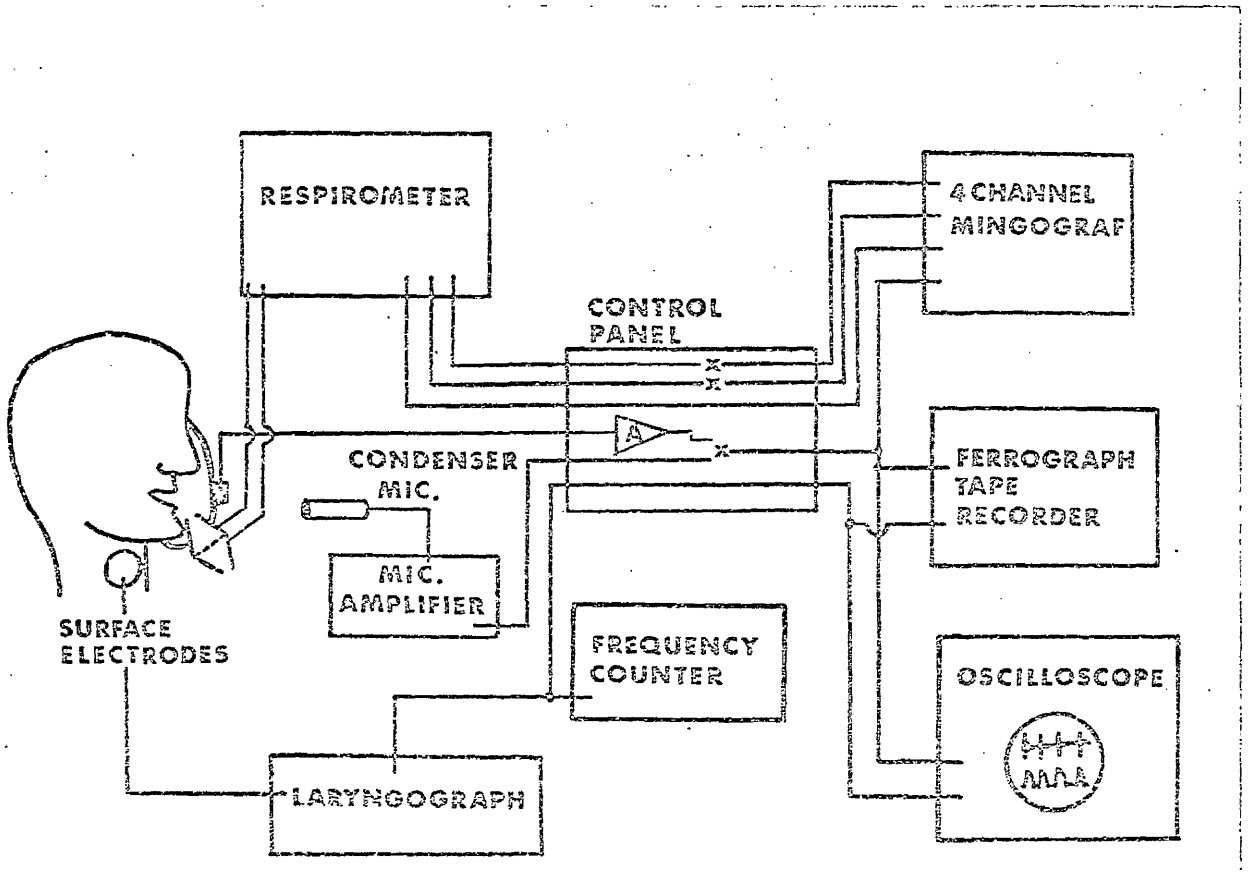


FIG.3.2 SCHEMATIC DIAGRAM OF THE EXPERIMENTAL SYSTEM.

subject was next required to wear the face mask and thus breathe through the pneumotach flowhead, and given several minutes to become used to the mask. Recordings were made of air flow rate and tidal volumes during a period of quiet respiration. The subject was then asked to sustain maximally the vowels /i/, /ɛ/ and /a/, firstly at his normal pitch, and then at the highest and lowest pitches which he could maintain while still phonating in chest register, all vowels being maintained at a "comfortable" sound level. (In this context, pitch is taken to mean the fundamental frequency of phonation, and chest register represents the range of fundamental frequencies within which the subject normally phonates). The expiratory vital capacity was then measured as the maximum volume of air expired after a maximum inspiration. Several tokens of this measure were made to ensure reproducibility. The subject was in a standing position during each exercise, and was allowed a rest period between each portion of the recording session to permit complete recovery.

3.3 ASSESSMENT OF RESPIRATORY CONTROL

The recordings made during the period of quiet respiration were firstly examined, and Figure 3.3 shows a typical example obtained from a member of the control group. The flow rate pattern was seen to be extremely regular from cycle to cycle, with the inspired and expired tidal volumes also being maintained at fairly constant levels. In contrast, Figure 3.4 shows two examples of quiet respiratory patterns recorded from members of the dysphonic group. The flow rate patterns were much less regular from cycle to cycle, and this was also reflected in the greater variation present in the tidal volumes. A quiet respiratory pattern showing great regularity from cycle to cycle was taken to mean that the subject had good respiratory control, while one showing great

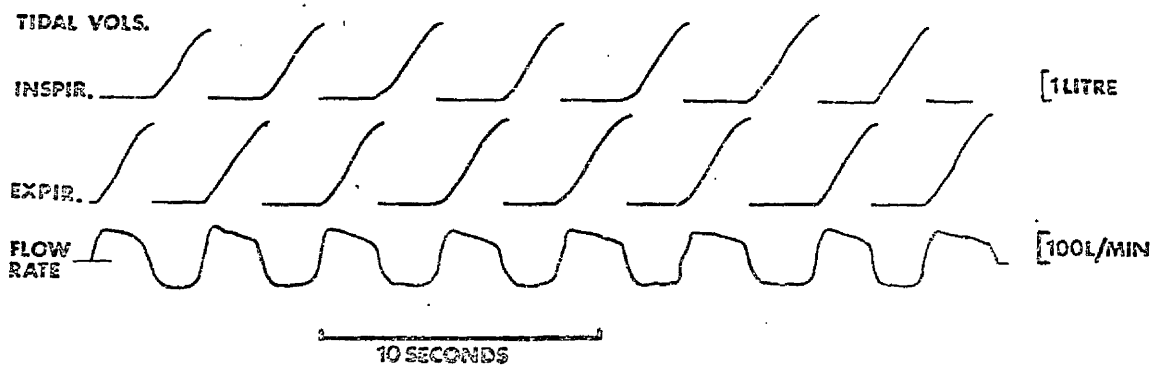


FIG. 3.3 EXAMPLE OF THE QUIET RESPIRATORY PATTERN OBTAINED FROM A MEMBER OF THE CONTROL GROUP.

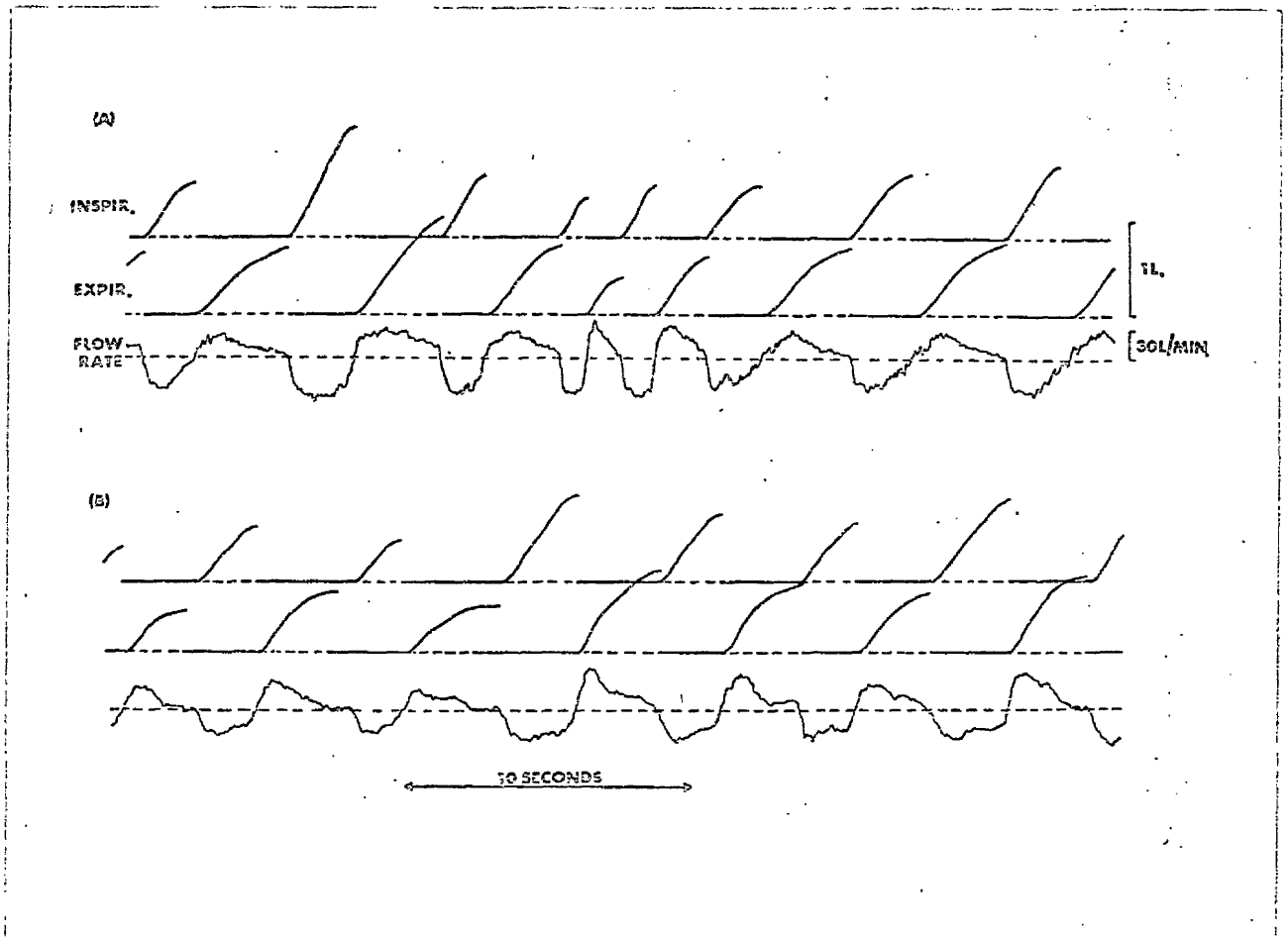


FIG.3.4 SAMPLES OF THE QUIET RESPIRATORY PATTERN OBTAINED FROM TWO MEMBERS OF THE DYSPHONIC GROUP.

variability from cycle to cycle was considered to indicate that the subject showed some lack of respiratory control. In order to quantify the amount of variation present, even for the members of the control group, the coefficient of variation was calculated for each subject for the values of inspiration and expiration tidal volumes.

Each series of quiet respiration recordings was taken and the first few cycles disregarded to eliminate possible artifacts. The mean and standard deviation were calculated for the parameters inspiration tidal volume and expiration tidal volume for a sample of between 10 and 20 consecutive cycles of each recording, and the coefficient of variation was calculated in each case, from :

$$\text{coefficient of variation} = \frac{(\text{standard deviation})}{\text{mean}} \cdot 100$$

(Garret, 1967)

This gave a measure of the variability present for each subject and also allowed inter-subject comparison. A frequency histogram of the coefficient of variation, (SD/M%), was plotted for both the control and dysphonic groups for the inspiration tidal volume data (Figure 3.5). The distribution of the control group data had a maximum in the 8-10% interval and decreased steadily at higher and lower values, the upper limit being about 18%. The members of the dysphonic group showed a greater degree of variability in the inspiration tidal volume data, as seen from the upper histogram in Figure 3.5., with (SD/M%) values extending up to about 36%.

These calculations were repeated for the expiration tidal volume data obtained from the same cycles of quiet respiration, and the frequency histograms are given in Figure 3.6. Similarly to the previous results, the control group distribution showed a maximum frequency of occurrence in the 8-10% interval,

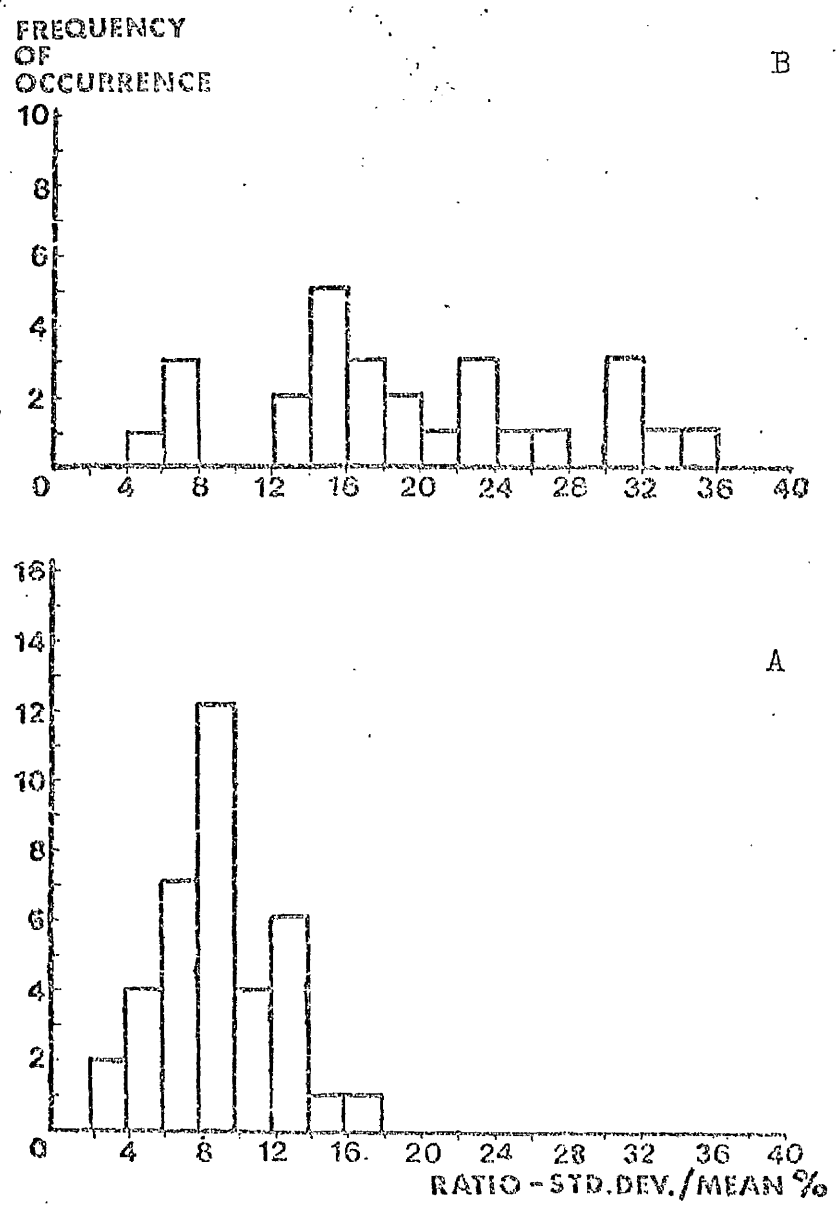


FIG.3.5 FREQUENCY HISTOGRAMS FOR COEFFICIENT OF VARIATION FOR INSPIRATORY TIDAL VOLUME DATA

A - CONTROL GROUP, B - DYSPHONIC GROUP

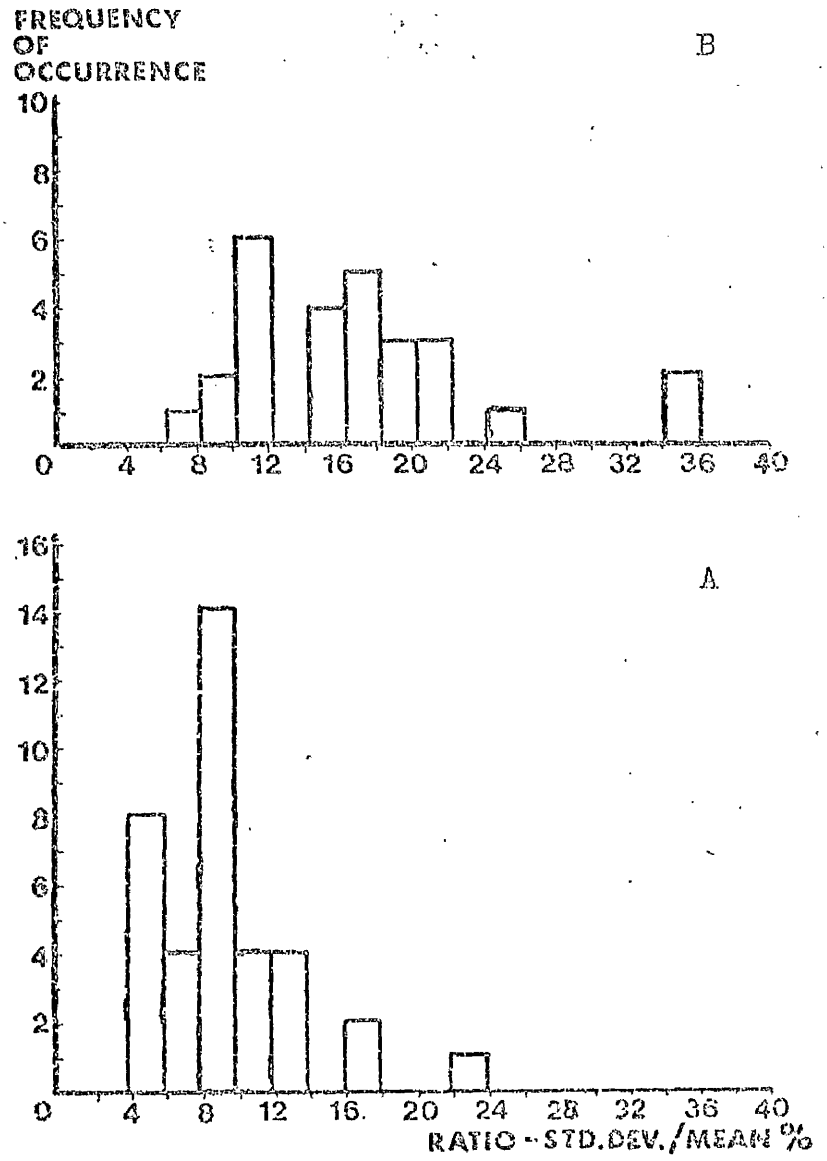


FIG.3.6 FREQUENCY HISTOGRAMS FOR COEFFICIENT OF VARIATION FOR INSPIRATORY TIDAL VOLUME DATA.

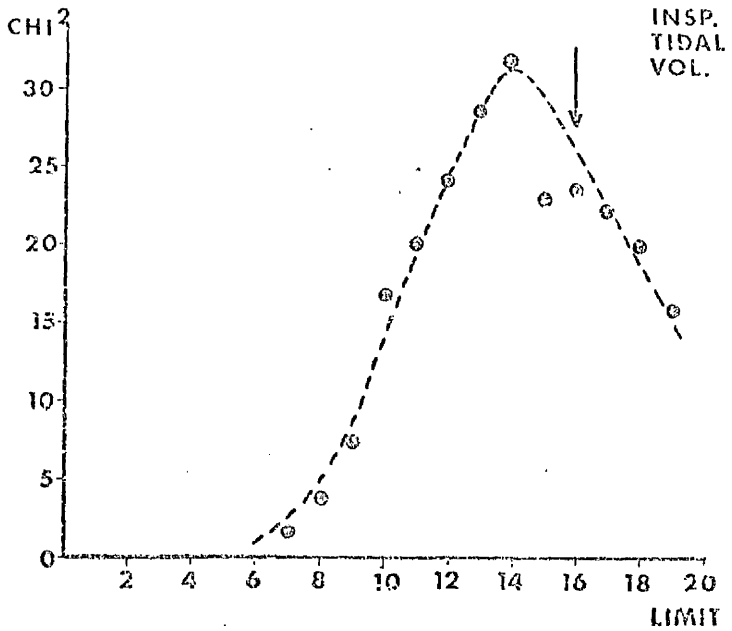
A - CONTROL GROUP, B - DYSPHONIC GROUP.

and a steady decrease towards the higher values, whereas several members of the dysphonic group had values well above those attained by the control group.

It appeared that some members of the dysphonic group demonstrated greater variability in the measured parameters than that expected from a subject with a totally normal quiet respiratory pattern. The possibility of obtaining a statistically significant limit of separation on the basis of these results was examined. The Null Hypothesis that both groups belonged to the same population was tested. If this were true, then for any value of (SD/M%), a similar proportion of each group would lie above and below this value. The significance of the differences between the proportions of each group was tested using the CHI^2 distribution, for each value of (SD/M%). (Appendix 1.1; Kirk,1976). The values of the CHI^2 coefficients were calculated and plotted against the (SD/M%) value, (limit), and these graphs are shown in Figure 3.7. In both cases, the CHI^2 coefficients reached highly significant values, e.g. for $p < 0.05$, $CHI^2 = 3.8$ and for $p < 0.0005$, $CHI^2 = 12.1$, for one degree of freedom, and so the Null Hypothesis was rejected.

A value of 16.0% was chosen as a "limit of normality" (arrowed on Figure 3.7) for both sets of tidal volume data, i.e. subjects having a variability greater than 16.0% as measured by the value of (SD/M%) would be considered to be abnormal. The values of the limits chosen were slightly to the right-hand side of the peak CHI^2 value, but still at a highly significant level. (The choice of limit values is discussed in SECTION 3.7).

Another important factor in assessing respiratory control is the variability with respect to time of the breathing pattern. As a measure of the temporal variability present for each subject, the



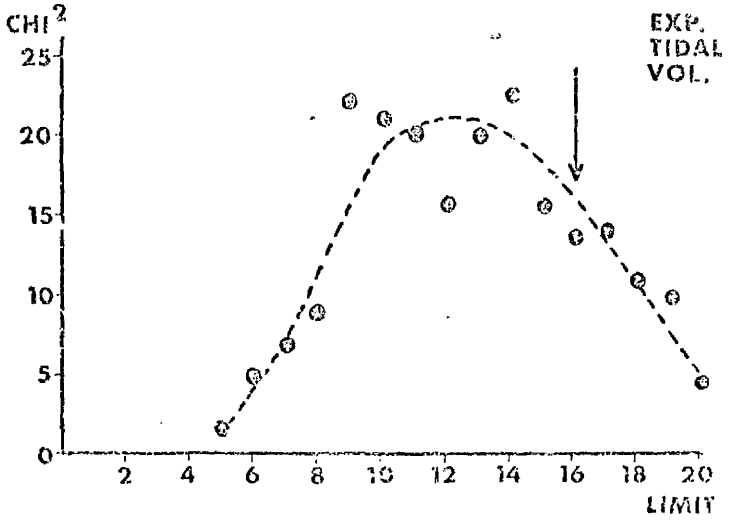


FIG.3.7 GRAPHS OF CHI^2 COEFFICIENT v LIMIT VALUE FOR COEFFICIENT OF VARIATION FOR TIDAL VOLUME DATA.

coefficients of variation were calculated for the parameters: period of the respiratory cycle, and the fraction of time per cycle spent on inspiration. (It had been found that one cycle of respiration consisted of inspiration - expiration and not expiration - inspiration (Cohen,1974)). The latter parameter was used to indicate whether or not if the respiratory period changed, the different parts of the cycle were altered in proportion. The same samples of quiet respiration were again used and the frequency histograms drawn for both groups of subjects, Figures 3.8 and 3.9. The frequency distribution obtained from the control group for the period data showed a peak in the range 2-8%, (Figure 3.8), with a few members of the group having values up to about 14%. However, there were again members of the dysphonic group having values well above those attained by the control group. The results obtained for the inspiration fraction data were very similar (Figure 3.9), with the distribution for the control group having a broad peak in the 4 - 12% range, and the variability for some members of the dysphonic group lying well beyond this interval. The CHI^2 test was again used to determine whether the two groups were significantly different with respect to these two parameters, and the CHI^2 values are plotted against the limit values of (SD/M%) in Figure 3.10. In both instances the high CHI^2 values indicated that the dysphonic and control groups were significantly different ($p < 0.0005$), and "limits of normality" were chosen at a (SD/M%) value of 14.0% (see SECTION 3.7), i.e. subjects showing variability greater than 14.0% in either of these parameters would be considered to show a lack of temporal control in their quiet respiratory patterns.

The four (SD/M%) ratios provided a means of assessing the amount of variability present in the corresponding parameters, and thus in the regularity of the quiet respiratory pattern during a representative number of cycles. The limits of normality set

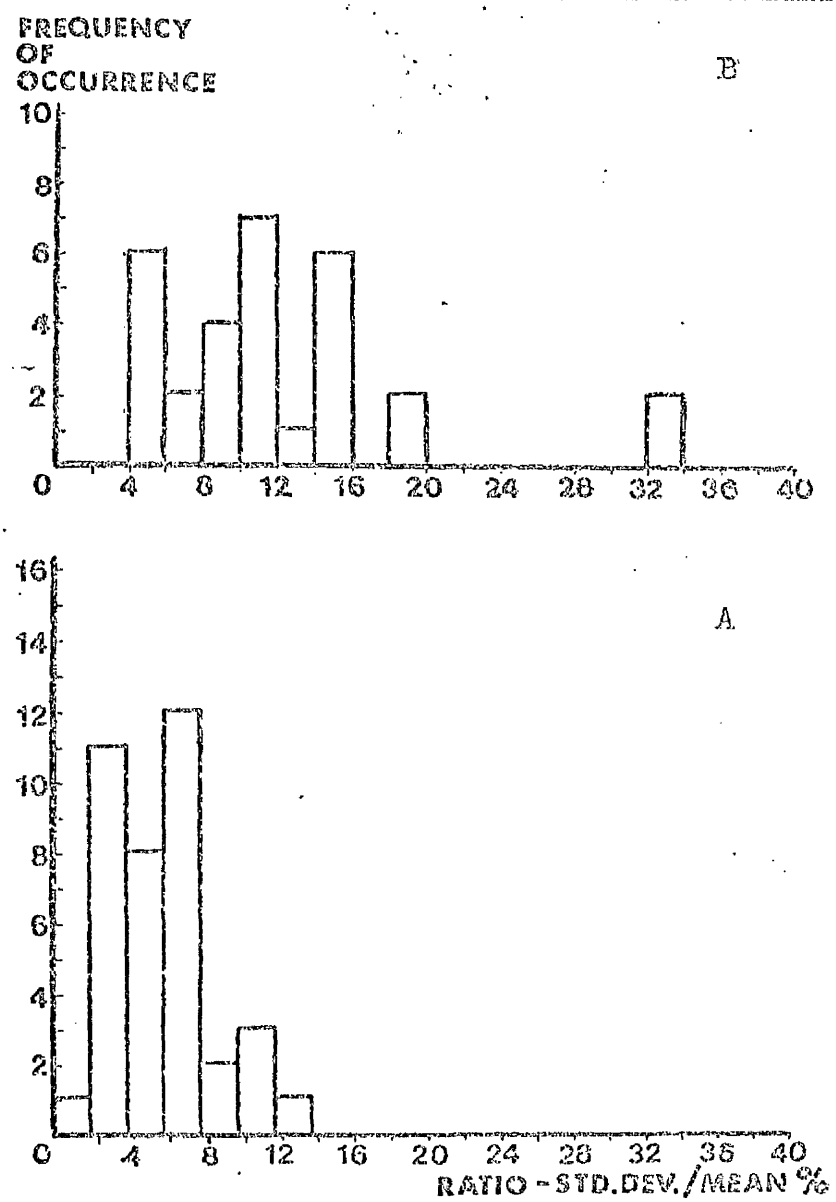


FIG. 3.8 FREQUENCY HISTOGRAMS FOR COEFFICIENT OF VARIATION FOR RESPIRATORY PERIOD DATA.

A - CONTROL GROUP, B - DYSPHONIC GROUP

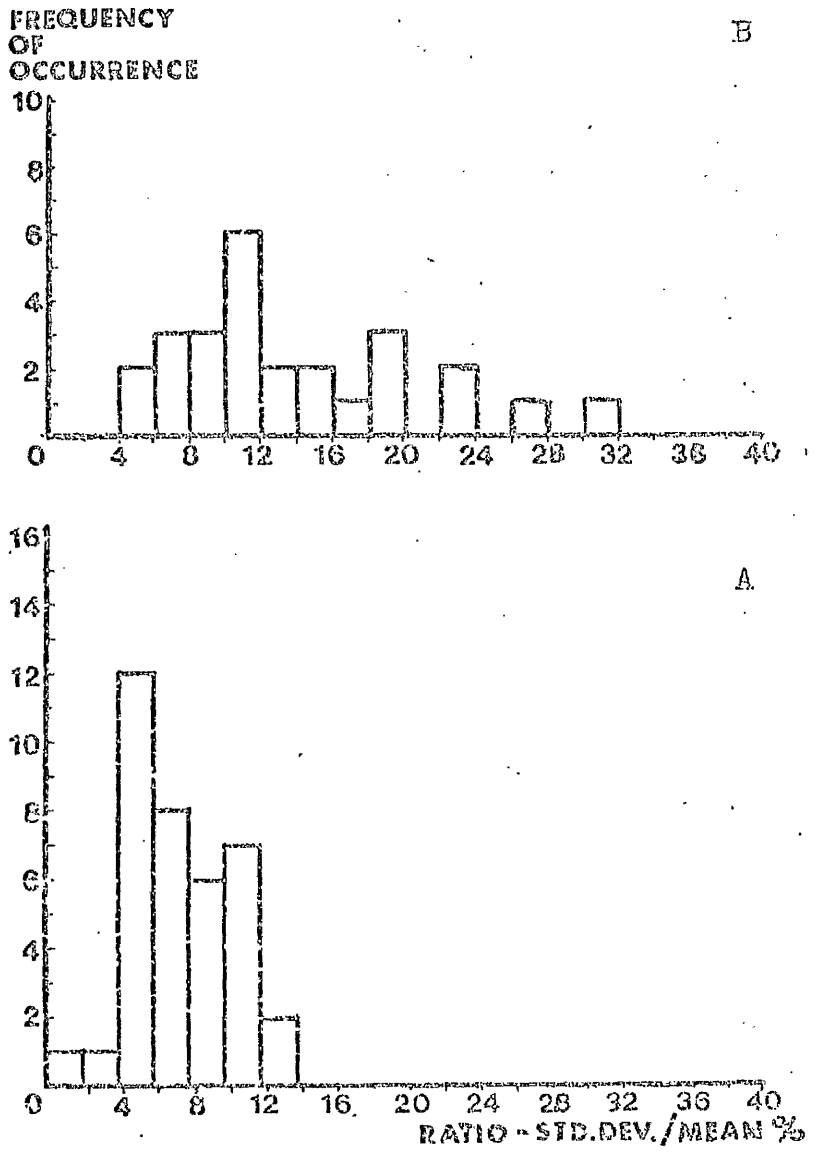


FIG. 3.9 FREQUENCY HISTOGRAMS FOR COEFFICIENT OF VARIATION FOR FRACTION OF TIME PER PERIOD SPENT ON INSPIRATION.

A - CONTROL GROUP, B - DYSPHONIC GROUP.

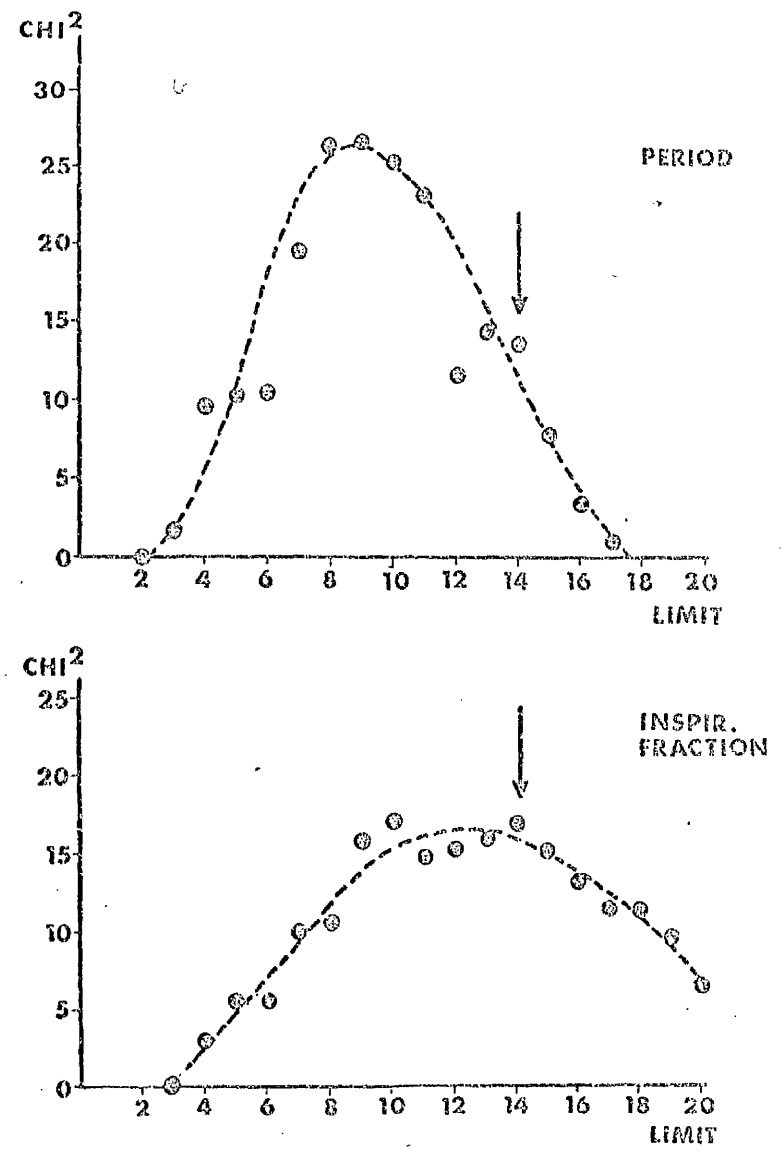


FIG.3.10 GRAPHS OF CHI^2 COEFFICIENT v LIMIT VALUE FOR COEFFICIENT OF VARIATION FOR TEMPORAL DATA.

an upper bound on the variability which could be expected from a "normal" subject on the basis of biological variation, and could be used to detect abnormality.

3.4 TESTS OF NORMALITY FOR RESPIRATORY CONTROL.

In the previous section, a method was described whereby certain upper limits could be placed on the variability of the parameters used to define the quiet respiratory pattern. Assuming that subjects having well controlled respiration show lesser degrees of variation than subjects having a lack of control, the limits could be used to provide a series of simple tests leading to a clinical assessment of respiratory control.

A scoring system was developed whereby the subject was given a score of "0" for each (SD/M%) value which was less than the corresponding "limit of normality", and a score of "1" for each result which was greater than the limit. Thus, each subject acquired a total score in the range 0 - 4. This was completed for all the subjects of the control and dysphonic groups, and the results compared with those obtained from the subjective assessment. (TABLE 3.1).

The control group consisted of 37 subjects, and of these, 34 were assigned total scores of 0, 2 of 1 and 1 of 2 (TABLE 3.1A). All the members of this group were subjectively assessed as "normal". A score of 0 was taken to indicate complete normality in the control of quiet respiration. The subject in this group who was assigned a score of 2 had results in the 16 - 18% intervals for both sets of tidal volume data, just outside the prescribed limits. The group of dysphonic subjects contained 29 members, and of these, 7 were assigned scores of 0, 6 of 1, 8 of 2, 4 of 3, and 4 had scores of 4. Of the 7 subjects who received scores of 0, 6 were subjectively assessed as "slightly abnormal" and 1 as "highly abnormal". (TABLE 3.1B).

TABLE 3.1 (A) RESULTS OBTAINED FOR CONTROL AND DYSPHONIC GROUPS USING OBJECTIVE ASSESSMENT OF QUIET RESPIRATION AND (B) COMPARISON BETWEEN SUBJECTIVE AND OBJECTIVE ASSESSMENTS.

(A)

GROUP	TOTAL NUMBER	TOTAL SCORE				
		0	1	2	3	4
CONTROL	37	34	2	1	0	0
DYSPHONIC	29	7	6	8	4	4

(B)

TOTAL SCORE	SUBJECTIVE ASSESSMENT			
	NORMAL	SLIGHTLY ABNORMAL	ABNORMAL	HIGHLY ABNORMAL
0	34	6	-	1
1	2	3	3	-
2	1	3	4	1
3	-	1	1	2
4	-	1	1	2

This subject was suffering from a mandibular tremor, and a portion of the quiet respiration recording is shown in Figure 3.11. The tremor caused a higher frequency oscillation to be superimposed on the flow rate, but did not result in a sufficient variation in the measured parameters to indicate abnormality, although in three of the four sets of data, the (SD/M%) values were just below the limits. Thus, 22 of the 39 members of this group were assessed as being "adnormal" to some degree, and a complete comparison with the subjective assessments for both groups is given in TABLE 3.1B. It appeared that the higher "score" values corresponded, in general, to the greater degrees of abnormality as subjectively assessed, and this was statistically tested using the contingency coefficient. (Appendix 1.2). This method tested the Null Hypothesis that the results obtained from the objective assessment (score results) and the subjective assessment were independent. The contingency coefficient was calculated ($C = 0.65$) and also the corresponding CHI^2 coefficient, 49.1. (12 degrees of freedom). This was highly significant ($p < 0.0005$), and the Null Hypothesis was thus rejected, leading to the conclusion that the results from both assessments showed very good agreement. However, by using the objective series of tests, a numerical result independent of the observer was obtained, and the clinical application will be discussed in a later section. (3.8).

3.5 ASSESSMENT OF PHONATORY EFFICIENCY

Some dysphonic subjects do not show irregular breathing patterns during quiet respiration and only exhibit abnormalities during phonation. It was, therefore, important to develop an objective means of assessing phonatory function on the basis of aerodynamic measurements, as these would also provide a guide to possible treatment.

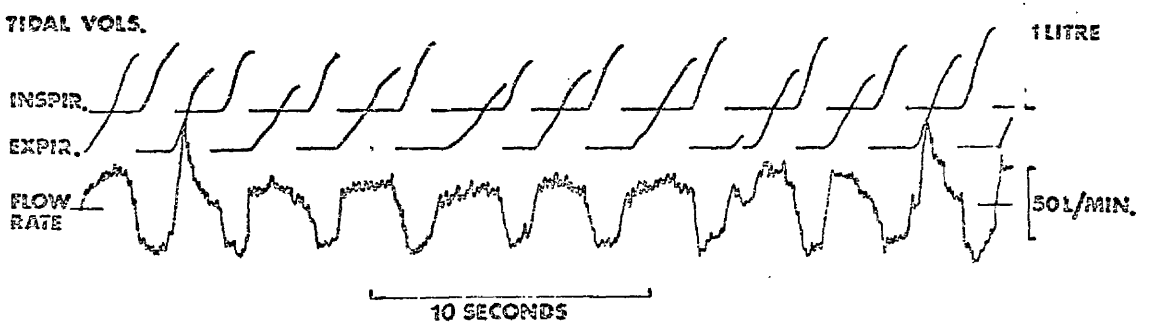


FIG.3.11 SAMPLE OF QUIET RESPIRATION RECORDING OBTAINED FROM SUBJECT WHO SUFFERED FROM A MANDIBULAR TREMOR.

Continuous speech is a very complex phenomenon from many viewpoints, not the least being the complexity of the air flow pattern. It was thus decided to investigate the air flow pattern resulting from the sustained phonation of a vowel sound, which approximated to a "steady state" situation. The vowel /a/ phonated at the subject's subjectively normal pitch was chosen as being typical of the recorded series (3.2), and all the results presented refer to this phonation. Examples of the respirometry outputs obtained during these recordings are shown in Figure 3.12. The examples in Figure 3.12A were obtained from two members of the dysphonic group, and could be compared to a "normal" phonation in Figure 3.12B. In the upper examples, the air flow rates were very irregular, with the resultant variability in the voice output, as seen on the microphone channel. It was also noted that some subjects phonated with very large air flow rates with resultant decreases in the times for which the phonations were sustained, while others used very low air flow rates and phonated with very low sound levels. The assessment of phonatory control was divided into two sections, firstly the control of the available air as measured by the air flow rate and its variability, and secondly, the quantity of air expired during the phonation as a function of the subject's lung volume, as measured by the expiratory vital capacity.

The volumes of air expired during the phonation /a/, (phonation volume), were measured from the recordings obtained from the members of the control group and the dysphonic group. The phonation times were also measured and the mean air flow rates (MFR) calculated. The frequency histograms for the MFR values for both groups were plotted (Figure 3.13), and the results obtained for the control group were found to be in the range 5 - 15 l/min. However,

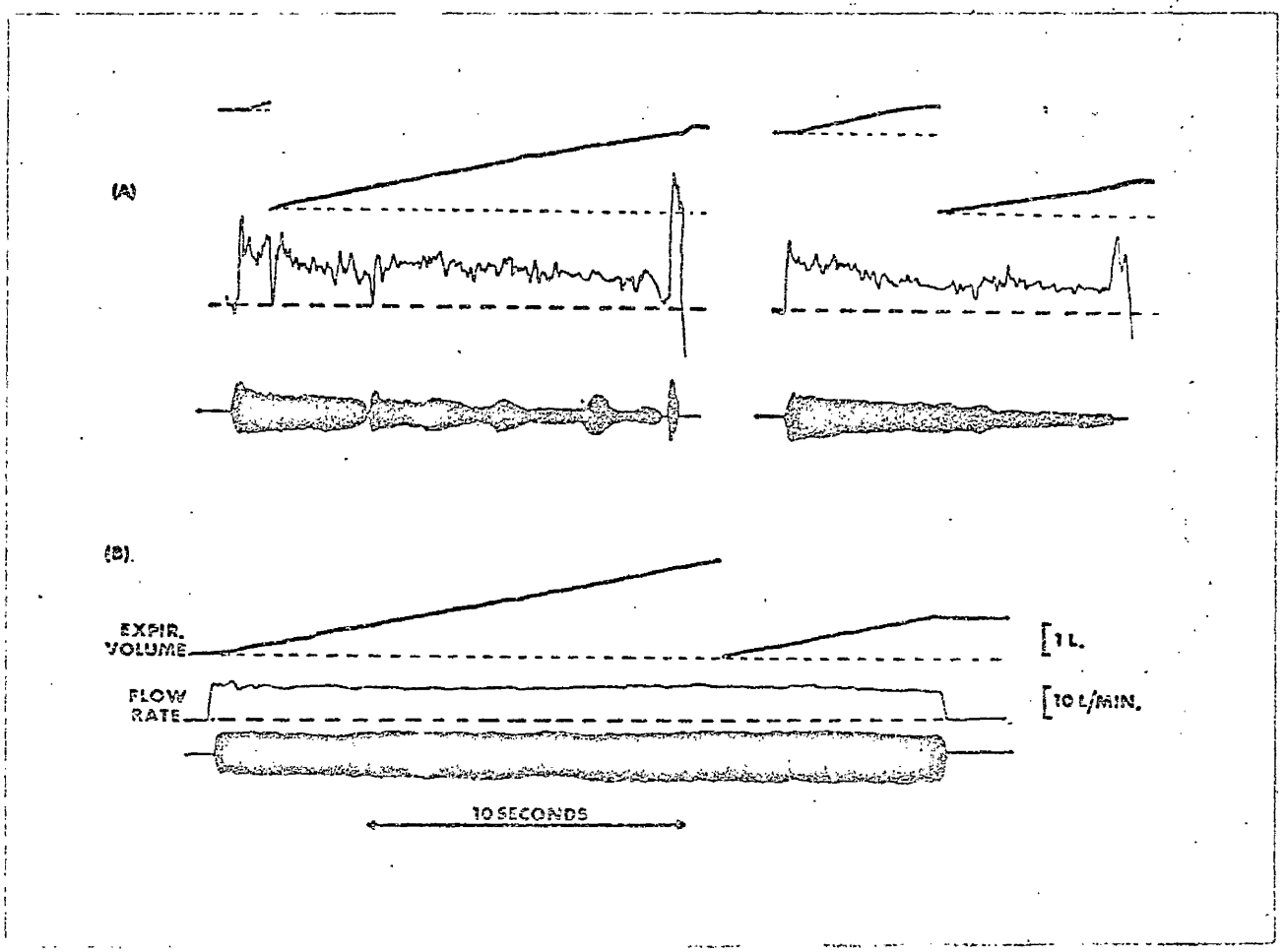


FIG.3.12 EXAMPLES OF RESPIROMETRY RECORDINGS MADE DURING THE SUSTAINED PHONATION OF THE VOWEL /a/.

A - MEMBERS OF DYSPHONIC GROUP.
B - MEMBER OF CONTROL GROUP.

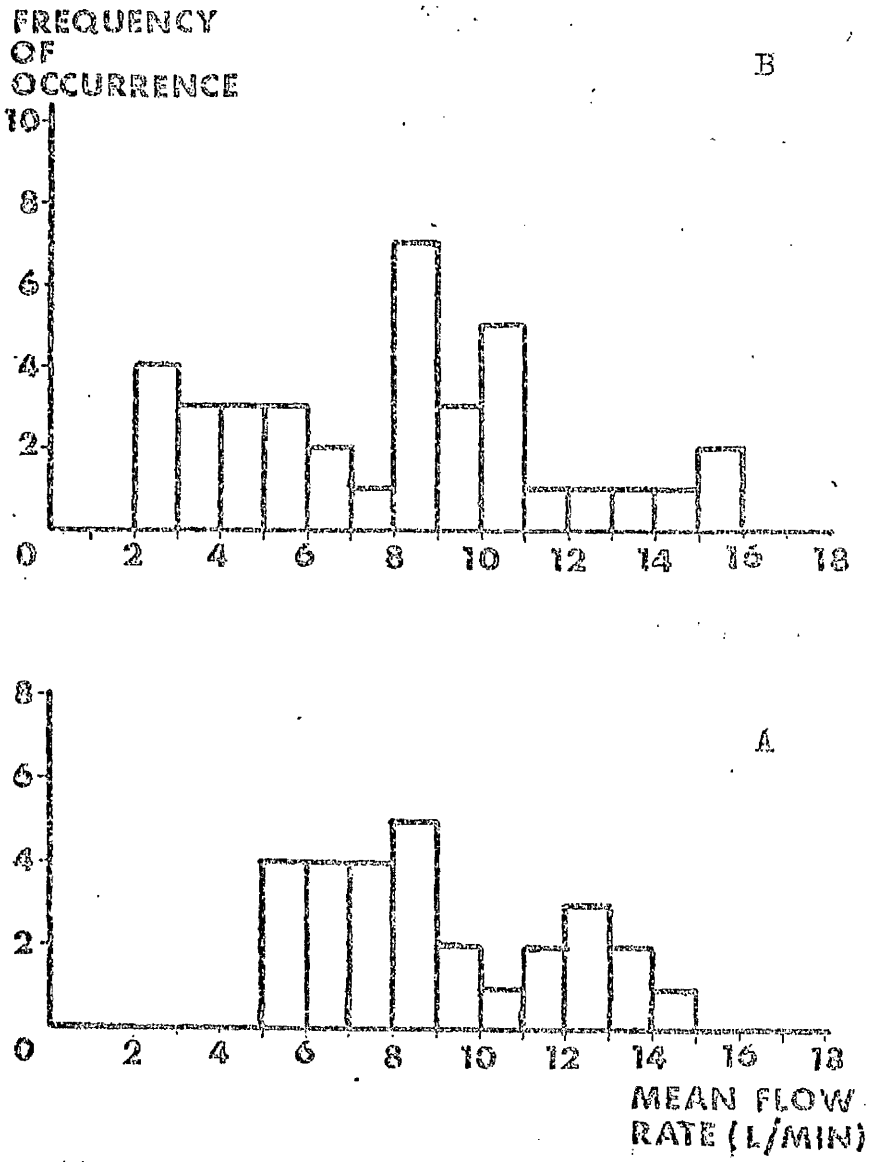


FIG.3.13 FREQUENCY HISTOGRAMS FOR MEAN FLOW RATE DATA.

A - CONTROL GROUP, B - DYSPHONIC GROUP

certain members of this group were considered to have phonated at either atypically high sound levels or fundamental frequencies (for these subjects) and thus the data from these phonations were disregarded in the assessment of flow rate limits for normal phonations. The remainder of the results were within the range 5 - 11 l/min., in agreement with the data reported by Isshiki and Von Leden (1964), and this range was taken to define the "limits of normality" for this type of phonation. Several members of the dysphonic group had MFR values outside these limits, but in these cases, the phonations did not appear to be atypical for the subjects.

It was not possible in the present investigation to obtain the standard deviation of the MFR, and so as a measure of the variation in flow rate present during the phonation, the ratio of (mean air flow rate) / (peak air flow rate), (M/P), was calculated. The peak flow rate was measured at the point of maximum flow rate in the central portion of the phonation so as to eliminate artifacts due to the initiation and cessation of voice production. The frequency histograms for the (M/P) ratio for the members of the two groups are shown in Figure 3.14. The greater degree of control achieved by the members of the control group, resulted in the members of this group attaining higher values of (M/P) than those attained by some subjects in the dysphonic group. The CHI^2 technique (Appendix 1.1) was applied to test the significance of the difference between the frequency distributions obtained for the groups, and the graph of the CHI^2 value against the limit value of (M/P) is given in Figure 3.15. The high value of the CHI^2 coefficient indicated that the two groups were very significantly different ($p \leq 0.0005$), and that the optimum value of the limit for (M/P) was 0.65, i.e. a subject with an (M/P) ratio smaller than 0.65 would be considered to be outwith the "normal" range. Thus, the values of MFR and (M/P) could be used to assess the

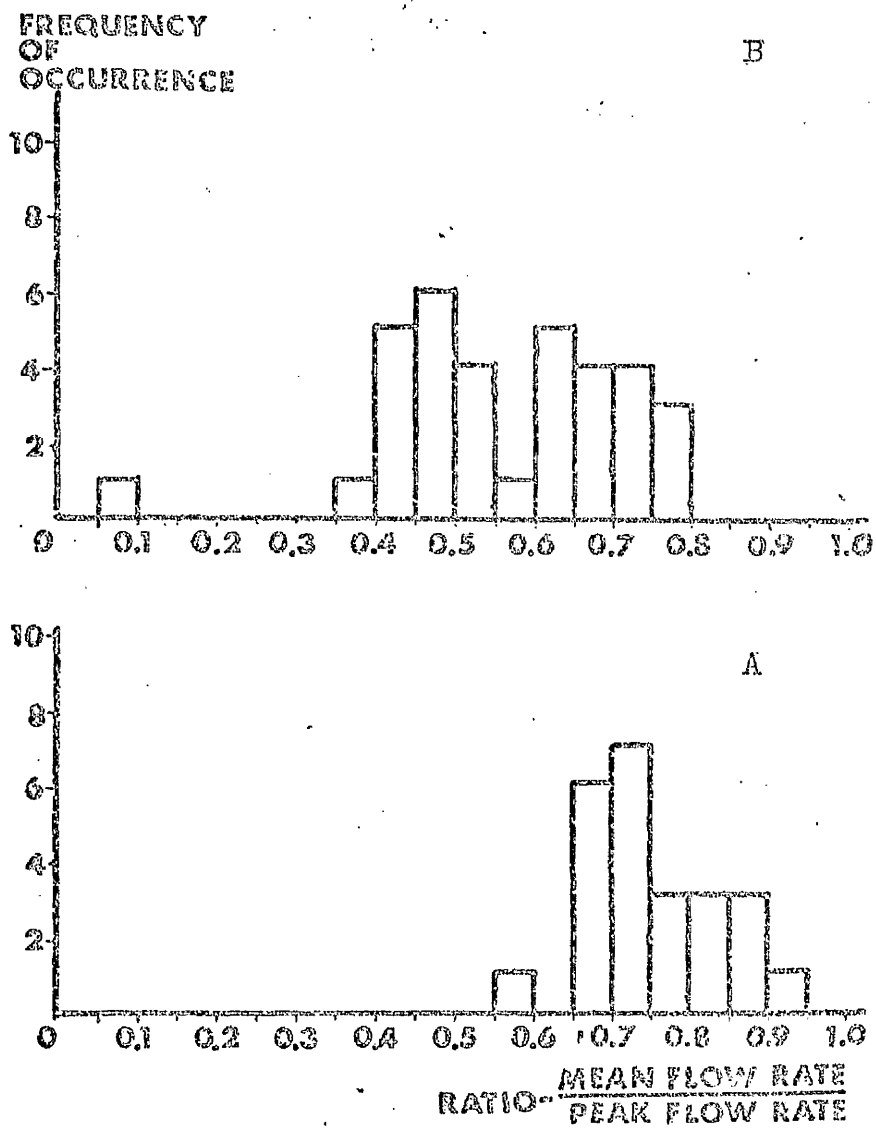


FIG. 3.14 FREQUENCY HISTOGRAMS FOR RATIO - (MEAN FLOW RATE / PEAK FLOW RATE) DATA.

A - CONTROL GROUP B - DYSPHONIC GROUP

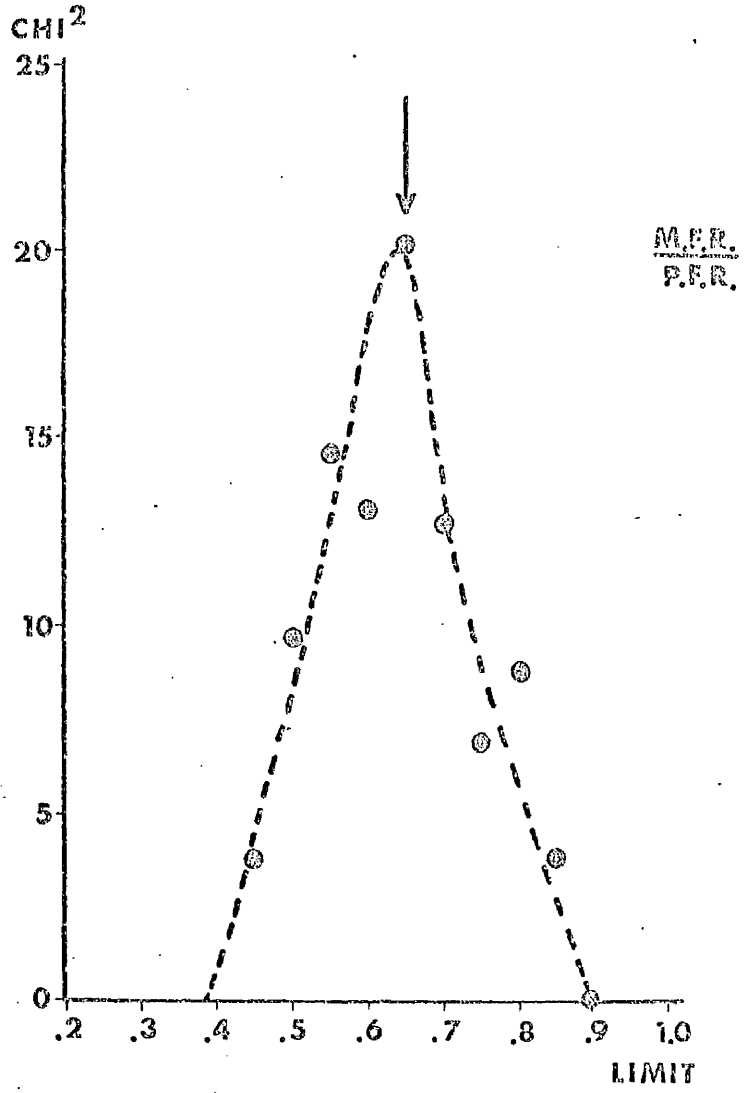


FIG.3.15 GRAPH OF CHI^2 COEFFICIENT v LIMIT VALUE FOR RATIO (MEAN FLOW RATE / PEAK FLOW RATE).

degree of control of air flow during the sustained phonation.

The phonation volume and expiratory vital capacity of each subject was obtained to give a measure of the amount of air used during the phonation, and also the amount of air available to the subject. The graph of phonation volume v vital capacity was drawn (Figure 3.16), and a linear relationship was evident for the data points of the control group, (as reported by Yanagihara and Koike (1967)). The regression line was calculated as :

PHON.VOL. = 0.64 x (VIT.CAP.) + 0.22

with a correlation coefficient, $r = 0.81$ ($p < 0.001$). This line, together with the 95% confidence limits, is shown on the graph, and it was observed that several members of the dysphonic group had data points below the lower 95% confidence limit. It was further noted that many members of this group were in the lower regions of graph, nearer the origin, and that it was possible to draw a line, transverse to the regression line, which would act as a lower bound to the data points of the control group. An arbitrary line with equation:

PHON.VOL. + VIT.CAP. = C C = constant

was chosen, and the CHI^2 technique applied to test the statistical significance of such a line for various values of C. (Figure 3.17). The optimum limit for C was found to be C = 3.5, with a CHI^2 value of 13.2 ($p < 0.005$, one degree of freedom), and the line

PHON.VOL. + VIT.CAP. = 3.5

is also shown on the graph (-.-.). A "region of normality" in the phonation volume v vital capacity space was defined as the upper region bounded by the lower 95% confidence limit of the regression line, and the transverse line with equation

PHON.VOL. + VIT.CAP. = 3.5.

The close correlation between phonation volume and vital capacity indicated the importance of vital

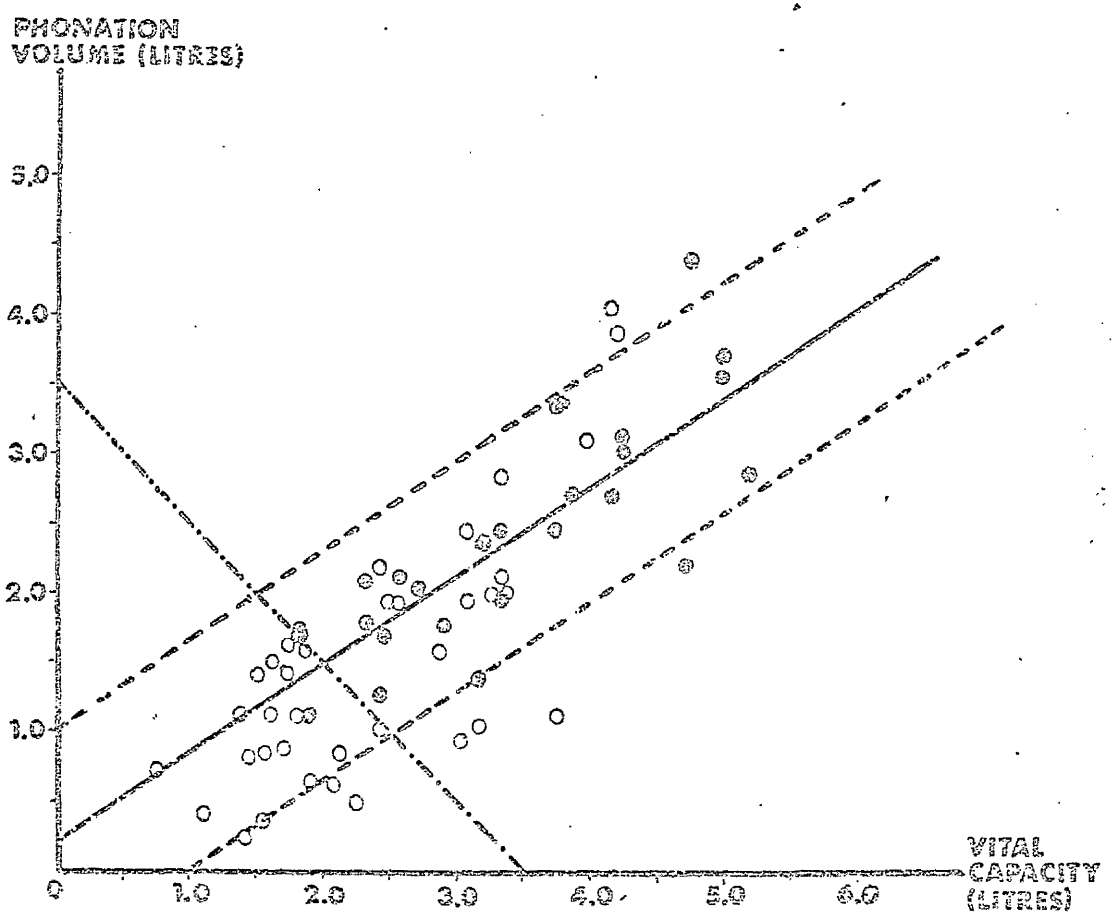


FIG.3.16 GRAPH OF PHONATION VOLUME v VITAL CAPACITY

⊛ – CONTROL GROUP
○ – DYSPHONIC GROUP

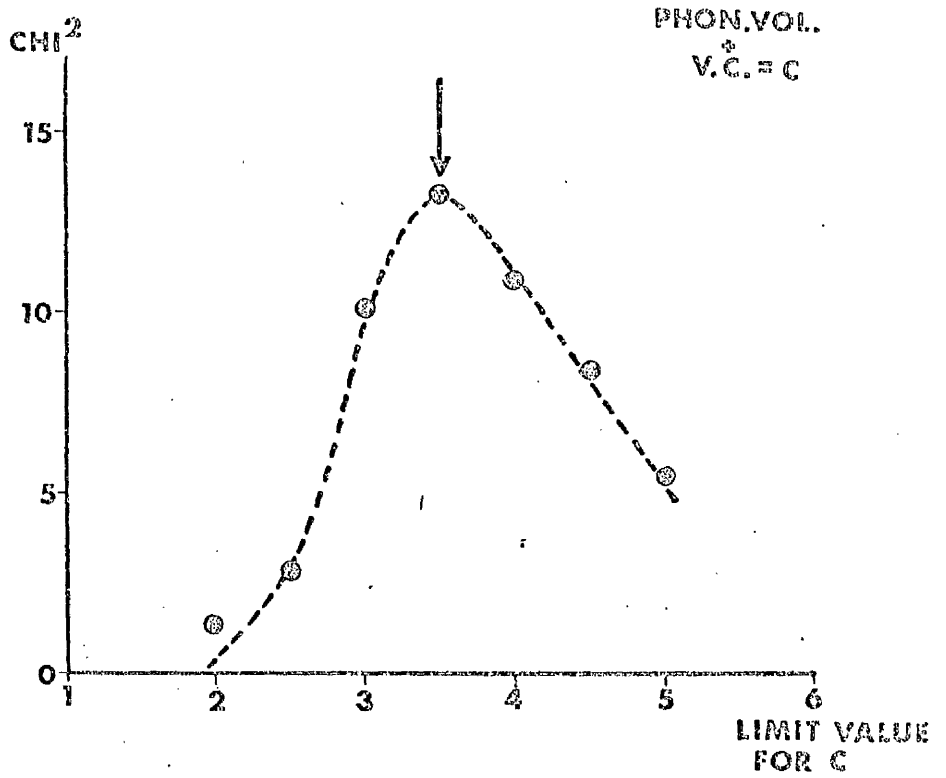


FIG.3.17 GRAPH OF CHI^2 COEFFICIENT v LIMIT VALUE FOR C.

capacity as a parameter relevant to phonation, and thus its use as a measure of lung performance. In particular, it was necessary to determine whether the measured vital capacity (MVC) was less than would be expected for the subject. The expected vital capacity (BVC) was calculated for each subject on the basis of height, age and sex using Baldwin's Formulae, (Baldwin et al 1938) :

B.V.C. = (27.63 - 0.112 x AGE) x HEIGHT—(MALE)

and

B.V.C. = (21.78 - 0.101 x AGE) x HEIGHT—(FEMALE)

with age in years, height in cm., and vital capacity in ml.. The ratio (MVC/BVC) was calculated for each subject, and the resultant frequency histograms are shown in Figure 3.18. Although Baldwin's formulae were strictly applicable to subjects in the supine position, the distribution obtained for the control group was symmetrical with a mean of 0.95 (SD = 0.20), indicating that the formulae were fairly reliable in the present case. However, many members of the dysphonic group had (MVC/BVC) ratios much less than unity, and the CHI^2 technique was applied to test for a significant limit of separation between the groups (Figure 3.19). The limit chosen was (MVC/BVC) = 0.7 ($p < 0.005$) and subjects with results less than this value would be considered to be outwith the normal range. The last two parameters indicated air usage and availability during the sustained phonation. The ranges of "normality" could now be used as the basis of an objective assessment of phonatory performance.

3.6 TESTS OF NORMALITY OF PHONATORY FUNCTION

A scoring system, similar to that described earlier (3.4), was developed to give an assessment of phonatory function, i.e. using the limits of normality for the MFR, (MFR/PFR), the area in the phonation volume v vital capacity space, and the value of (MVC/BVC), each subject was assigned a score of "0" or "1"

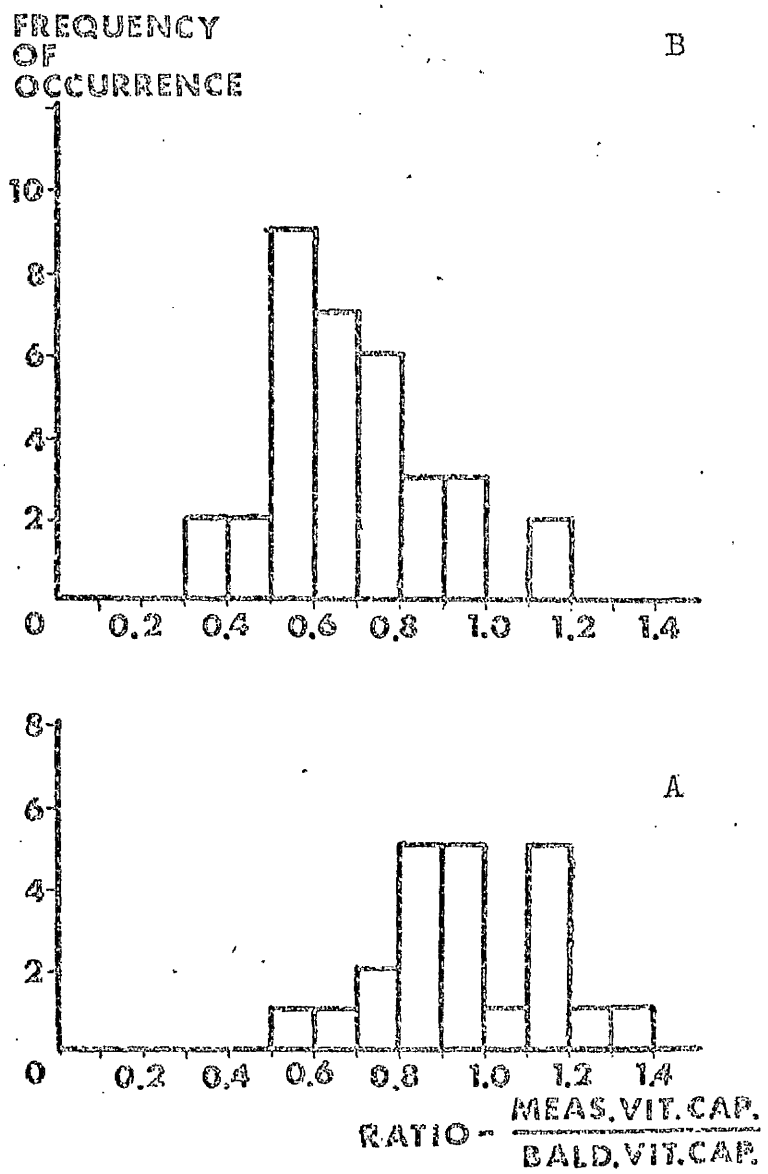


FIG.3.18 FREQUENCY HISTOGRAMS FOR RATIO – (MEASURED VITAL CAPACITY / BALDWIN'S VITAL CAPACITY)

A – CONTROL GROUP B – DYSPHONIC GROUP

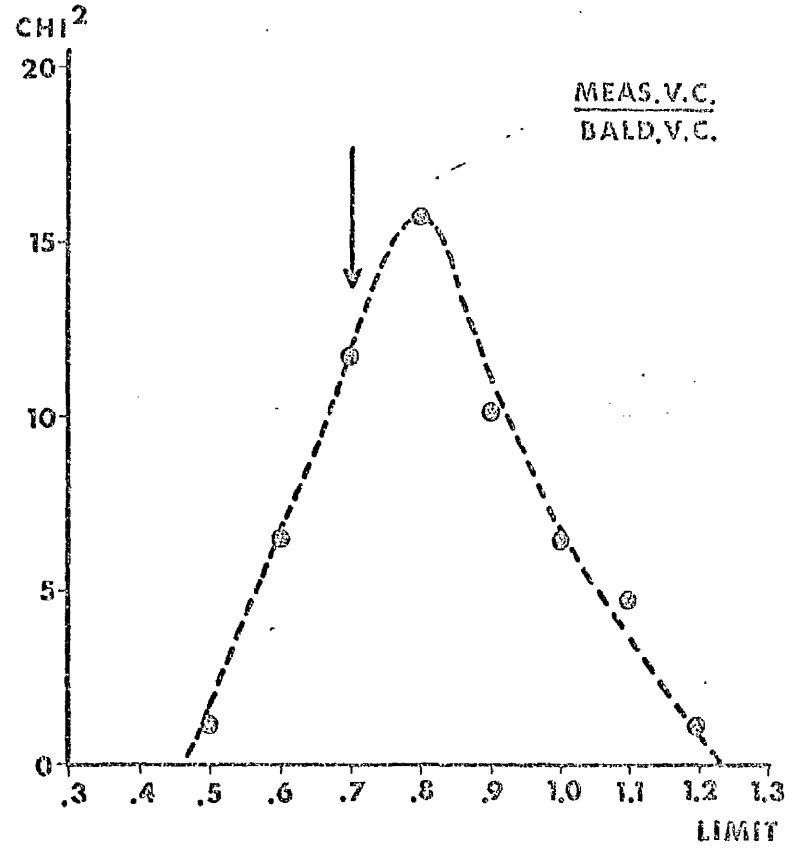


FIG. 3.19 GRAPH OF CHI^2 COEFFICIENT v' LIMIT VALUE FOR RATIO - (MEASURED VITAL CAPACITY / BALDWIN'S VITAL CAPACITY.)

according to whether each result was within or outwith the determined limits. All of the members of both groups were thus allotted total scores in the range 0 - 4, and these results were compared with those of the subjective assessment (TABLE 3.2). Every member of the control group was subjectively assessed as "normal". This group consisted of 28 subjects, of whom 16 were assigned scores of 0, 11 of 1, and 1 had a total score of 2. However, this series of tests relied much more on the absolute values of the parameters, e.g. MFR, phonation volume and vital capacity, than did the previous series, and also much more on subject co-operation. Thus a greater degree of variability was expected, even from the control group, and so a total score of 0 or 1 would not be considered to indicate clinical abnormality. Eight of the 11 members of this group with total scores of 1 were the subjects who had abnormally high MFR values resulting from atypical phonations, although this did not effect the other measures. The dysphonic group consisted of 39 subjects, of whom 4 were allotted scores of 0, 8 of 1, 14 of 2, 9 of 3 and 4 were given scores of 4. (TABLE 3.2A). The comparison of the objective results with those obtained from the subjective assessment is given in TABLE 3.2B. Twelve members of the dysphonic group would not be considered to be "abnormal". However, of the total sample of 67 subjects, in 54 of the cases the objective and subjective assessments agreed. The subjects who were subjectively assessed to have the greater degrees of abnormality also tended to be assigned the higher total scores. The contingency coefficient was again calculated to test for the independence of the objective and subjective assessments, ($C=0.65$) and the CHI^2 value (47.8) indicated that there was a high degree of correlation between the results of the two methods ($p < 0.0005$, 12 degrees of freedom). Again, however, the objective assessment gave a numerical

TABLE 3.2 (A) RESULTS OBTAINED FOR CONTROL AND DYSPHONIC GROUPS USING OBJECTIVE ASSESSMENT OF PHONATION, AND (B) COMPARISON BETWEEN OBJECTIVE AND SUBJECTIVE ASSESSMENTS.

(A)

		TOTAL SCORE				
GROUP	TOTAL NUMBER	0	1	2	3	4
CONTROL	28	16	11	1	0	0
DYSPHONIC	39	4	8	14	9	4

(B)

	SUBJECTIVE ASSESSMENT			
TOTAL SCORE	NORMAL	SLIGHTLY ABNORMAL	ABNORMAL	HIGHLY ABNORMAL
0	16	3	1	-
1	11	2	6	-
2	1	4	9	1
3	-	1	4	4
4	-	-	3	1

value to the degree of abnormality and eliminated the personal bias which must be present in all subjective techniques.

3.7 LIMITS OF NORMALITY

In many instances, the limits of normality were obtained by using the method described in APPENDIX 1.1, to calculate the CHI^2 value as a function of the limit of the required parameter. The high CHI^2 values indicated in each case that a limit was valid at a highly significant level ($p < 0.0005$). However, the limit chosen did not always correspond to the maximum CHI^2 value, as this decision had to be tempered by the numbers of patients which could be successfully coped with clinically. On each occasion when this occurred, the decision was made to optimize the number of false positives, (i.e."normal subjects who would be classified as "abnormal"), but unfortunately, the corollary also took place, producing an increase in the number of false negatives, (i.e. "abnormal" subjects who were classified as "normal"). The decision was taken to allow a level of about 10% for the number of false positive diagnoses resulting from each series of tests, and thus this was similar to using Receiver Operator Characteristic curves for this system. The decision was also made to set equal limits for the tidal volume data, and the same applied for the temporal data in the quiet respiration series of tests.

The method of calculating the limits also depended on the results obtained from the dysphonic group. However, if the data collected from the control group adhered to a particular distribution, then the 95% confidence limits could be used to provide limits of normality which were independent of the results of the dysphonic group. If, for example, the present data obtained from the control group were fitted to Gaussian distribution curves,

the 95% confidence limits would be given by the appropriate value of mean ± 2 S.D. This was carried out for the relevant variables and the results are tabulated in TABLE 3.3. For each curve, the CHI^2 coefficient was calculated from :

$$CHI^2 = \sum_i \frac{\left(f(O)_i - f(E)_i\right)^2}{f(E)_i}$$

where $f_i(O)$ = observed frequency in the i^{th} interval and $f_i(E)$ = expected frequency in the i^{th} interval, with the condition that for all intervals $f_i(O)$ and $f_i(E)$ are greater than 5. (Moroney,1970). If this did not hold, then adjacent intervals were summed until the condition was met. This led to a drastic reduction in the number of degrees of freedom. However, the low CHI^2 values showed that in no case could the Gaussian curve be rejected as being a bad fit to the observed data at the 95% level of confidence. The test limits showed a reasonable agreement with those obtained from the Gaussian assumption, but it would be desirable to obtain data from a much larger control group before this method would be suitable for calculating the limits.

3.8 CLINICAL VALUE OF THE OBJECTIVE ASSESSMENTS

The results obtained using the objective assessments which have been developed, agreed well with those obtained by the standard speech therapy techniques. However, not only is the subjective nature of the orthodox assessment eliminated, but quantitative values of respiratory control and phonatory function are presented for each subject, and these factors are of prime importance in deciding upon the optimum treatment required. In particular, if a patient exhibits an abnormal quiet respiratory pattern, there is a strong possibility that he will also have abnormal phonatory function, but if this alone is treated it is probable that little or no improvement will be ob-

TABLE 3.3 RESULTS OBTAINED ASSUMING GAUSSIAN DISTRIBUTIONS FOR DATA FROM CONTROL GROUP.

PARAMETER	MEAN	SD	CHI^2	DEGREES OF FREEDOM	95% LIMIT	TEST LIMIT
SD/M%, INSP. TIDAL VOL.	9.32	3.50	3.92	3	16.32	16.0
SD/M%, EXP. TIDAL VOL.	9.59	4.29	3.45	2	18.17	16.0
SD/M%, PERIOD	5.48	2.34	2.48	3	10.16	14.0
SD/M%, INSP. TIME FRACTION	7.34	2.87	2.89	3	13.08	14.0
MFR	7.55	1.47	0.86	3	4.61, 10.49	5.0, 11.0
M/P	0.75	0.08	2.74	3	0.59	0.65
MVC/BVC	0.95	0.20	0.59	3	0.55	0.7

served. The respiratory control must firstly be improved.

These tests are useful not only in providing an initial assessment of a subject's condition, but also in monitoring a patient's progress throughout a course of treatment. The respiratory outputs shown in Figure 3.20 were obtained from a subject who was suffering from hypertension due to psychological factors. The breath pattern during quiet respiration was very irregular with great variation from cycle to cycle, and during the sustained phonation, the sound level was variable, as shown by the microphone output, and the air flow rate was very unsteady. At the end of each phonation, the patient expired about 0.25l. of air which should have been utilised for phonation, thus the phonation volume was smaller than expected. The patient underwent a course of treatment consisting of breathing exercises and relaxation therapy, and after several sessions, a marked improvement was evident in some of the measured parameters. The variation in the quiet respiratory pattern as measured by the (SD/M%) values for tidal volumes, decreased from about 64% to 28%, which although still abnormally high, was a significant improvement. The air flow rate during sustained phonation became more steady, with the (M/P) ratio increasing from 0.60 to 0.71, and the mean flow rate decreasing from 13.8l/min. to 10.5l/min. Thus, although some of the values were still outside the "normal" range, the overall trend towards "normality" was evident.

These series of tests are now used routinely to assess every patient attending the speech therapy department, and where possible to monitor the patient's progress. The results are calculated using a Wang 700 desk computer, and the program is given in APPENDIX 2. The program was subsequently rewritten in Fortran IV to suit an IBM 1100 series computer. (Appendix 2.3).

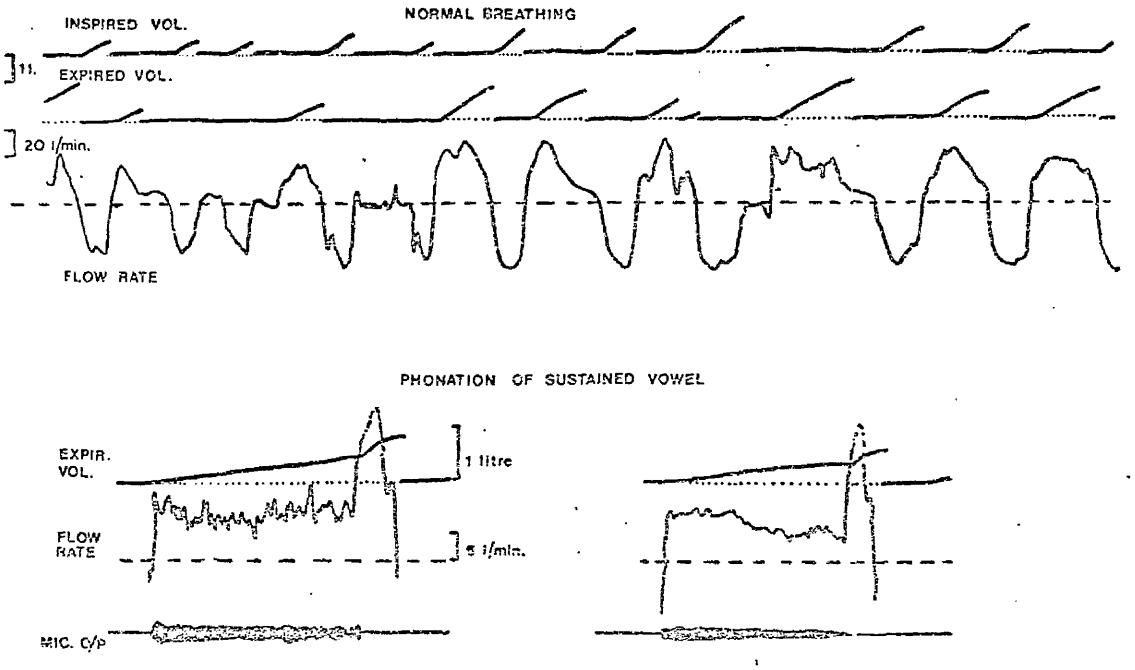


FIG.3.20 RESPIROMETRY OUTPUT OBTAINED FROM SUBJECT SUFFERING FROM HYPERTENSION.

The frequency distributions resulting from the next 140 patients are shown in Figures 3.21 - 3.28, and, where appropriate, the Gaussian curves expected on the basis of the original control group are also shown (dashed curves). The data obtained during quiet respiration, Figures 3.21 - 3.24 showed that a considerable number of the group had results well above those that would be expected by comparison with the control group. This led to the frequency distributions having lower peak values, and extending to higher (SD/M%) values, (apart from that for inspiration fraction, Figure 3.24, which showed few patients with abnormally high values).

The results obtained during the sustained phonation were similar, with many of the subjects having data points outwith the limits of "normality", (Figures 3.25 - 3.28). The assessment scores assigned for the quiet respiration recordings are given in TABLE 3.4. 55 patients received scores of 0, 20 of 1, 43 of 2, 21 of 3 and 1 of 4. The most common "abnormality" in the group with a score of 1 was in either of the tidal volume data sets, and this factor was common throughout. Of the 43 subjects with scores of 2, 35 had the combination of both tidal volume sets scored as "abnormal". (This relationship was also present in the original group of dysphonic subjects where there was a significant correlation between the (SD/M%) values for inspiration tidal volume and expiration tidal volume ($r = 0.61$, $p < 0.001$)). The least common "abnormal" parameter was inspiration fraction.

The results of the second assessment for the same group are shown in TABLE 3.5. The main relationship shown is the close association between the factors assessing vital capacity and phonation volume and also between the factors assessing air flow rate, MFR and (M/P). The factor (MVC/BVC) rarely appears without the factor PHON.VOL. v V.C. For both assessments combined, only 14 subjects (10%) were considered to be not abnormal, 3 having scores of (0,0) and 11 of (0,1).

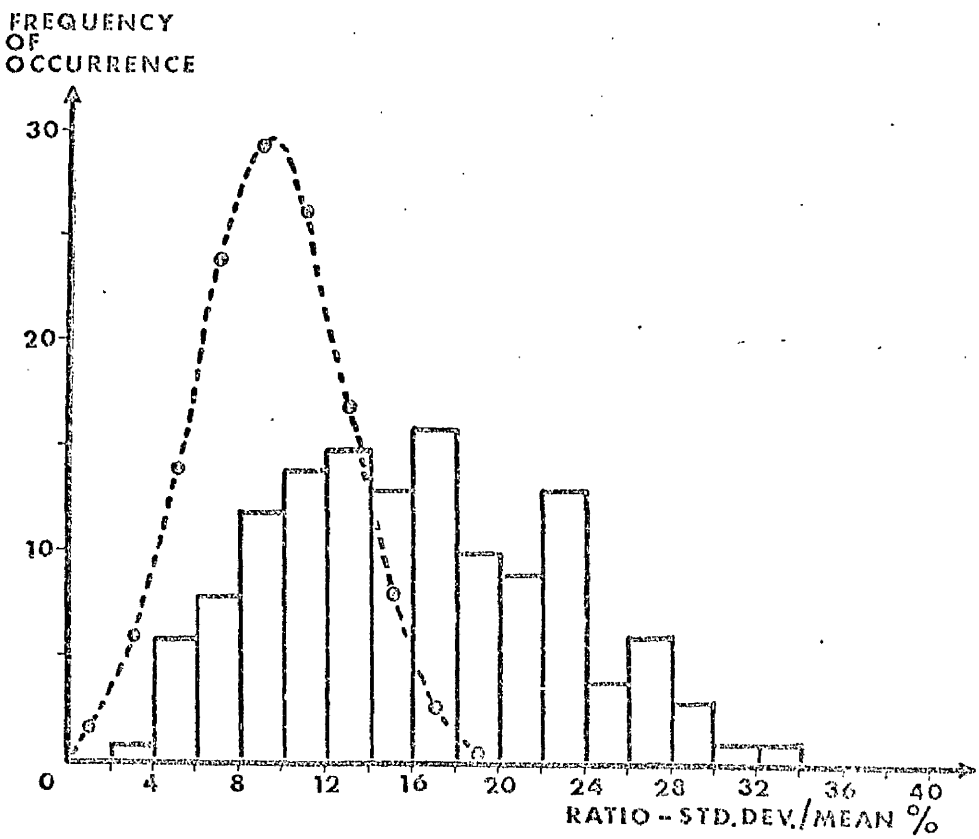


FIG. 3.21 EXTENDED DATA FOR INSPIRATORY TIDAL VOLUMES.

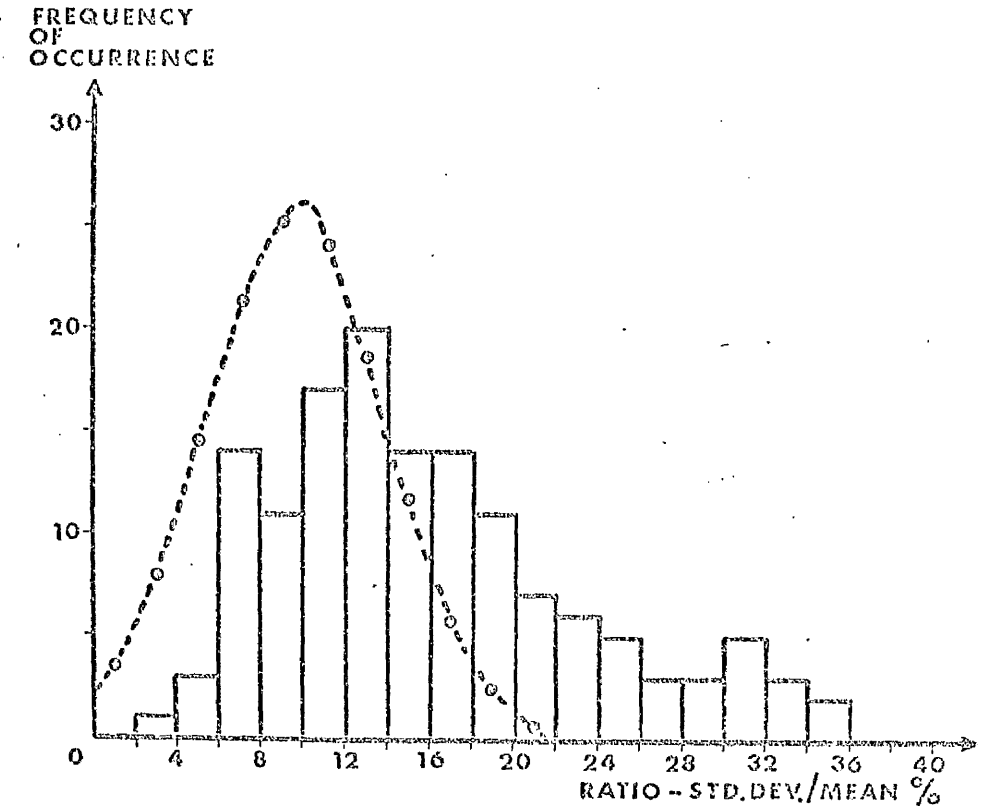


FIG.3.22 EXTENDED DATA FOR EXPIRATORY TIDAL VOLUMES.

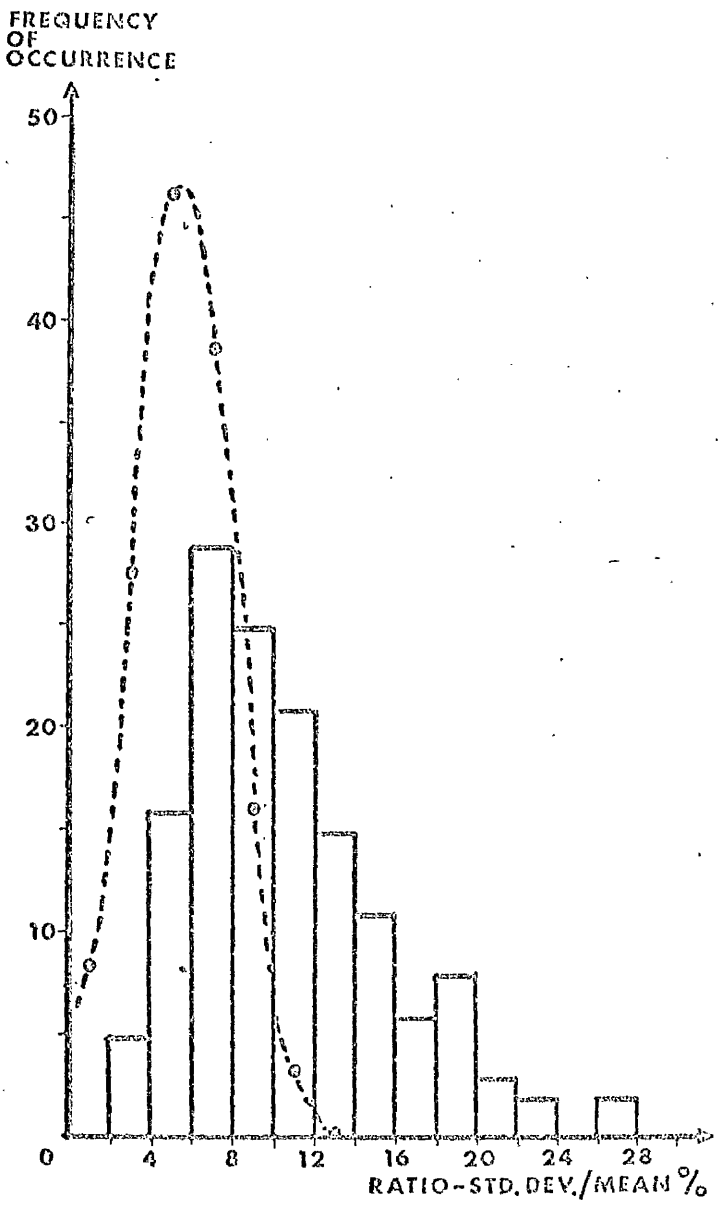


FIG.3.23 EXTENDED DATA FOR RESPIRATORY PERIOD.

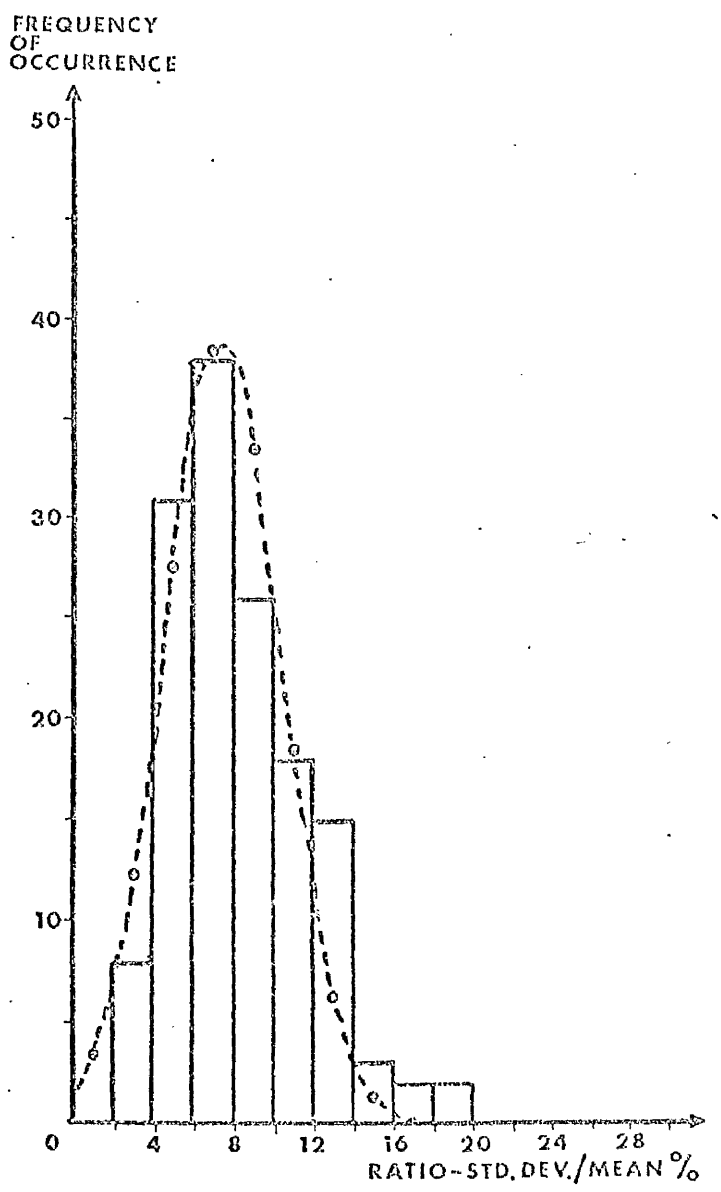


FIG.3.24 EXTENDED DATA FOR FRACTION OF TIME PER PERIOD SPENT ON INSPIRATION.

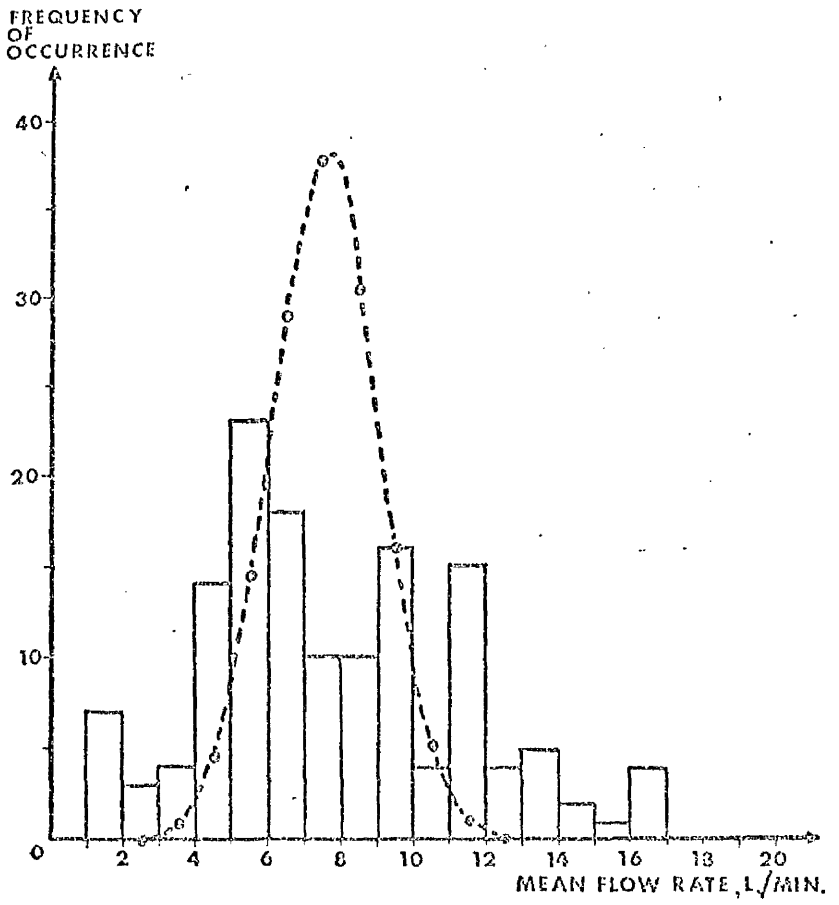


FIG.3.25 EXTENDED DATA FOR MEAN FLOW RATE DURING SUSTAINED PHONATION.

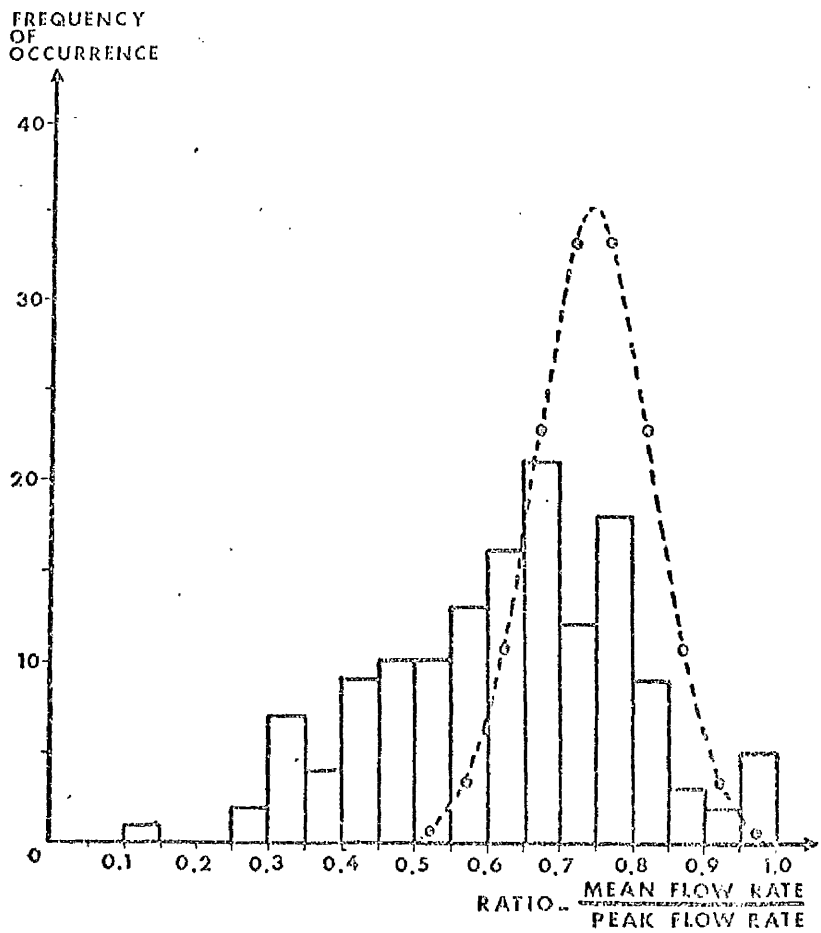


FIG.3.26 EXTENDED DATA FOR RATIO - (MEAN FLOW RATE / PEAK FLOW RATE).

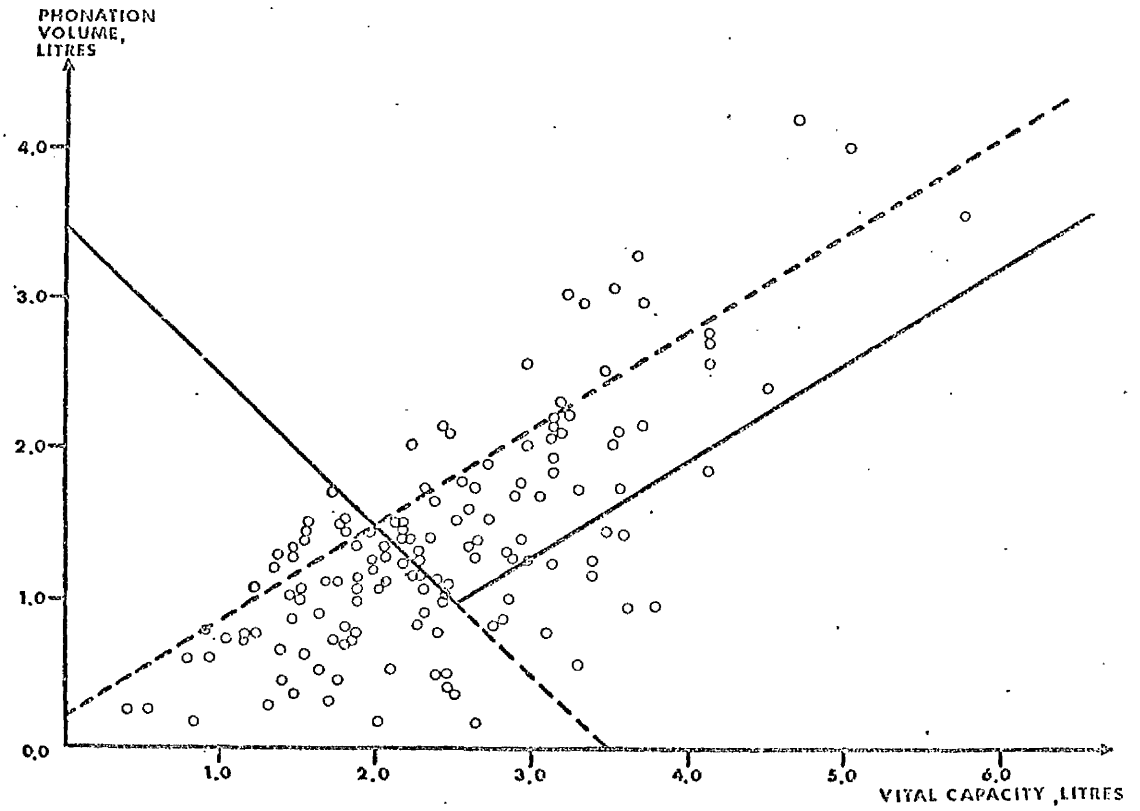


FIG.3.27 EXTENDED DATA FOR GRAPH OF PHONATION VOLUME v VITAL CAPACITY.

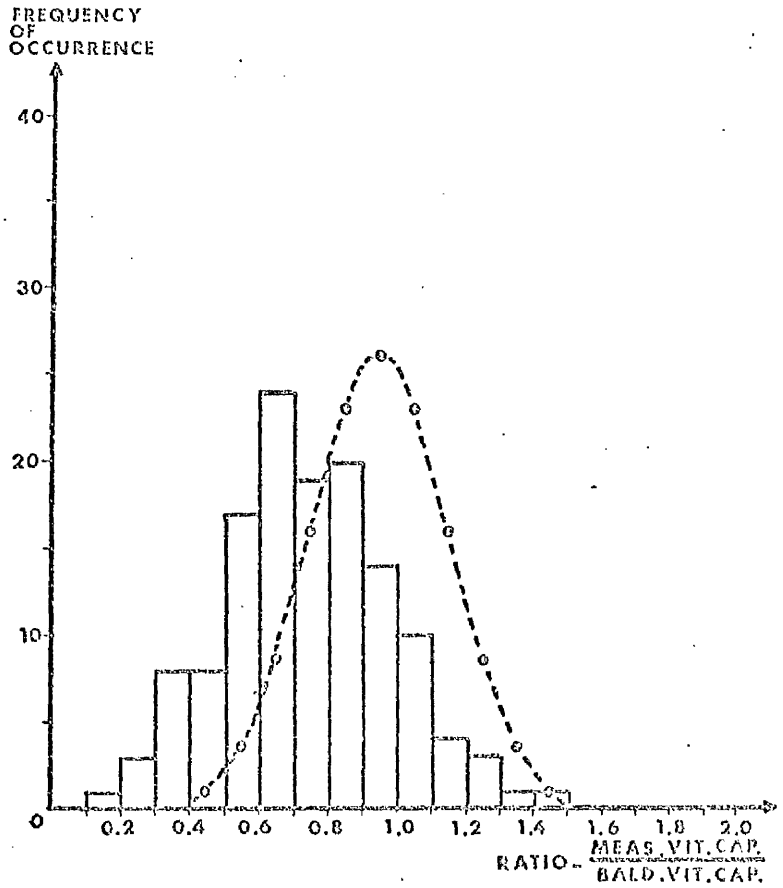


FIG.3.28 EXTENDED DATA FOR RATIO - (MEASURED VITAL CAPACITY / BALDWIN'S VITAL CAPACITY).

TABLE 3.4 RESULTS OBTAINED FROM QUIET RESPIRATION ASSESSMENT FOR EXTENDED GROUP OF PATIENTS.

TOTAL SCORE	COMBINATION ASSESSED AS ABNORMAL	NO.	TOTAL NO.
0	-	55	55
1	INS.T.VOL.	11	20
	EXP.T.VOL.	6	
	PERIOD	3	
	INSP.FRACT.	-	
2	INS.T.VOL.,EXP.T.VOL.	35	43
	INS.T.VOL.,PER.	3	
	INS.T.VOL.,INSP.FRACT.	2	
	EXP.T.VOL.,PER.	1	
	EXP.T.VOL.,INSP.FRACT.	-	
	PER.,INSP.FRACT.	2	
3	INS.T.VOL.,EXP.T.VOL. PER.	19	21
	INS.T.VOL.,EXP.T.VOL. INSP.FRACT.	2	
	EXP.T.VOL., PER., INSP.FRACT.	-	
	INS.T.VOL., PER., INSP.FRACT.	-	
4	INS.T.VOL.,EXP.T.VOL., PER.,INSP.FRACT.	1	1
			140

TABLE 3.5 RESULTS OBTAINED FROM PHONATORY ASSESSMENT FOR EXTENDED GROUP OF PATIENTS.

TOTAL SCORE	COMBINATION ASSESSED AS ABNORMAL	NO.	TOTAL NO.
0	-	11	11
1	MFR	10	43
	M/P	21	
	PH.V.v V.C.	10	
	MVC/BVC	2	
2	MFR, M/P	6	35
	MFR, PHON.V. v V.C.	6	
	MFR,MVC/BVC	0	
	M/P, PHON.V. v V.C.	7	
	M/P, MVC/BVC	1	
	PHON.V. v V.C., MVC/BVC	15	
3	MFR,M/P,PHON.V. v V.C.	7	38
	MFR, M/P, MVC/BVC	2	
	MFR,PHON.V. v V.C.,MVC/BVC	14	
	M/P,PHON.V.v V.C.,MVC/BVC	15	
4	MFR,M/P,PHON.V.v V.C. MVC/BVC	13	13
			140

Thus 90% of this group of patients were considered to show some degree of abnormality, and the objective assessment provided guidelines to possible courses of treatment.

3.9 SUMMARY

The standard techniques of assessing vocal performance are unsatisfactory because of their subjective nature. In this CHAPTER, the development of two series of tests was described, both based on aerodynamic measurements. The first series of tests was used to assess respiratory control during a period of quiet respiration, and consisted in measuring the coefficient of variation, (SD/M%), for several parameters, i.e. inspiration and expiration tidal volumes, period and fraction of time per cycle spent on inspiration. The assumption was made that subjects showing good respiratory control would show less variability in these parameters than subjects having a lack of control. In the second series of tests, measurements relating to the control of air flow and air usage were made during the sustained phonation of the vowel /a/. In both series of tests, statistical techniques were used to obtain "limits of normality", by comparing the results of a control group with those of a dysphonic group. The presentation of the results of the relevant tests to the clinician and speech therapist was made simpler by developing a score system, whereby the subject received a score of 1 for each test result which was outwith the prescribed "limits of normality". There was a good agreement between the results obtained using the objective tests and those of the orthodox subjective techniques, but the numerical data from the former, could be used to give a better guide to a possible course of treatment. The objective assessment is now applied routinely in the Speech Therapy Department.

CHAPTER 4.

THE FREQUENCY ANALYSIS OF VOWEL SOUNDS.

CHAPTER 4. THE FREQUENCY ANALYSIS OF VOWEL SOUNDS.

4.1. INTRODUCTION

It was explained in CHAPTER 1 that the sound source for voiced sounds, (e.g. vowel sounds), resulted from the periodic interruption of the air flow from the lungs by the vibration of the vocal folds. This sound is modified by the resonating properties of the cavities of the upper vocal tract, these properties being dependent on the size and shape of the individual cavities, and also upon the coupling between them. Thus, the voice output signal is also periodic, having the same period as the source function, i.e. that of the vocal fold vibration. Any periodic function can be described as the summation of a series of harmonically related frequency components by means of Fourier Analysis, and this is particularly valid for voice signals, since hearing also involves a process of frequency analysis. The frequency spectrum of a vowel sound consists of a component at the fundamental frequency, F_o, (equal to the reciprocal of the period of the vocal fold vibration), and a series of harmonically related components at the frequencies $2F_o$, $3F_o$, -- etc.

In this chapter, the frequency spectra of vowel sounds are discussed, and compared to results obtained by Peterson and Barney, (1952). The presence of non-harmonic components is used to assess vocal performance (using the classification suggested by Yanagihara (1967)), and these results are compared to those obtained by the aerodynamic techniques described in CHAPTER 3. The non-harmonic frequency components consist of several types, including those appearing as random "noise" components, which form the basis of the Yanagihara classification, and also those which arise as sub-multiples of the main harmonic series (i.e. at

frequencies $\frac{1}{2}F_o$, $\frac{3}{2}F_o$ etc.). The latter type of frequency components are found to be quite common in the spectra obtained from dysphonic subjects, and several examples are used to illustrate this feature.

4.2. HARMONICS AND FORMANTS

The harmonic nature of the frequency spectrum of the source function is illustrated in the upper diagram of Figure 4.1. The spectrum consists of a series of equally spaced frequency components, and in this case, the intensity decreases steadily as the frequency increases. Let the envelope curve of the source spectrum be described by the frequency dependent function $S'(f)$. The vocal tract transfer function, $T'(f)$, is continuous with respect to frequency and consists of a series of peaks due to the resonances of the cavities of the upper vocal tract, so that frequency components occurring near these resonant peaks will be transmitted more strongly, (i.e. amplified) than components further away from the peaks. The envelope curve of the voice output function, $P'(f)$, is given by the result of $T'(f)$ acting on $S'(f)$, such that :

$$P'(f) = T'(f). \; S'(f) \qquad (4.1)$$

In acoustic measurements, it is usual to express such functions on a logarithmic scale (decibel or dB scale). Thus if :
$P(f) = \log P'(f)$, $T(f) = \log T'(f)$ and $S(f) = \log S'(f)$ equation 4.1 becomes

$$P(f) = T(f) + S(f) \qquad (4.2)$$

as shown in figure 4.1. The peaks in the envelope curve of the voice output function are called FORMANTS, and are numbered sequentially from low to higher frequencies. The formant frequencies do not coincide with the resonant frequencies of the vocal tract transfer function because of the frequency dependence

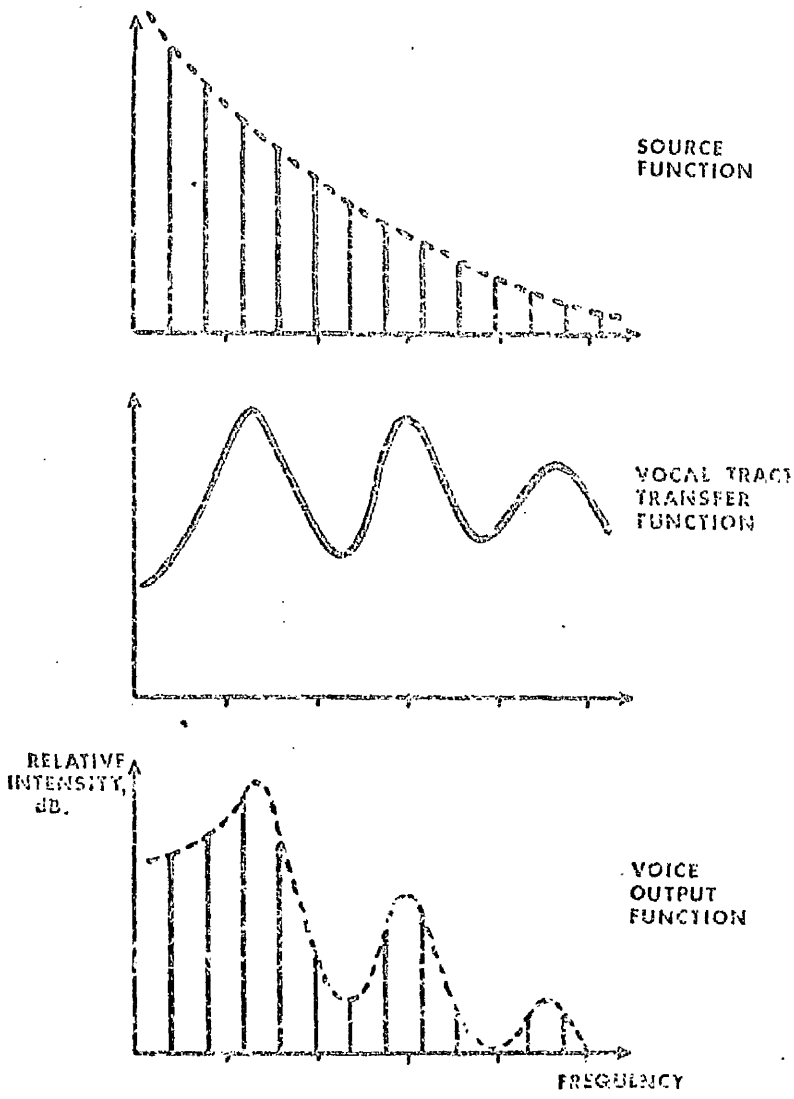


FIG.4.1. ILLUSTRATION OF HARMONIC NATURE OF SOUND SOURCE AND VOICE OUTPUT FUNCTIONS, AND THE ORIGIN OF FORMANTS.

of the source function, S(f). The spectrum of the voice output function also consists of a series of harmonically related components, and it is unlikely that a harmonic component will correspond exactly to a formant frequency. There is, therefore, often great difficulty in identifying formant frequencies, especially if the fundamental frequency is high and the number of harmonic components small, e.g. in the spectra of female subjects.

Examples of the frequency analysis of the vowels /i/, /e/ and /a/ carried out using the Sonagraph are shown in Figure 4.2. The filter bandwidth was 300Hz., and since the fundamental frequency was about 110Hz., the harmonic nature of the sounds was not evident. However, the horizontal dark bands indicated the presence of energy peaks or formants, e.g. between 2 KHz. and 4 KHz. for the vowel /e/. The formant structures of six vowel sounds are illustrated in the sectioner displays shown in Figure 4.3, the 45 Hz. filter having been used. The harmonic nature of each sound is evident from the discrete frequency components which are present, and the formants are labelled F1, F2, etc. It can be seen from the figure that each vowel appears to have a distinctive structure with respect to the formant frequencies.

Peterson and Barney, (1952), reported an experiment in which a group of American English speaking subjects recorded a series of words of the form hVd, where V represents different vowel sounds, (e.g. had, hid, hud, etc.), and the vowel portion of each word was analysed to give the frequencies of the formants, F1, F2 and F3. A plot of F1 v. F2 was made and demonstrated that each vowel tended to occupy a particular area in the F1 v. F2 space (Figure 4.4), indicating the importance of the formant structure in the discrimination between different vowel sounds. In later experiments, (Peterson, 1961), it was shown that vowel recognition depended closely on the ratio F2/F1, and

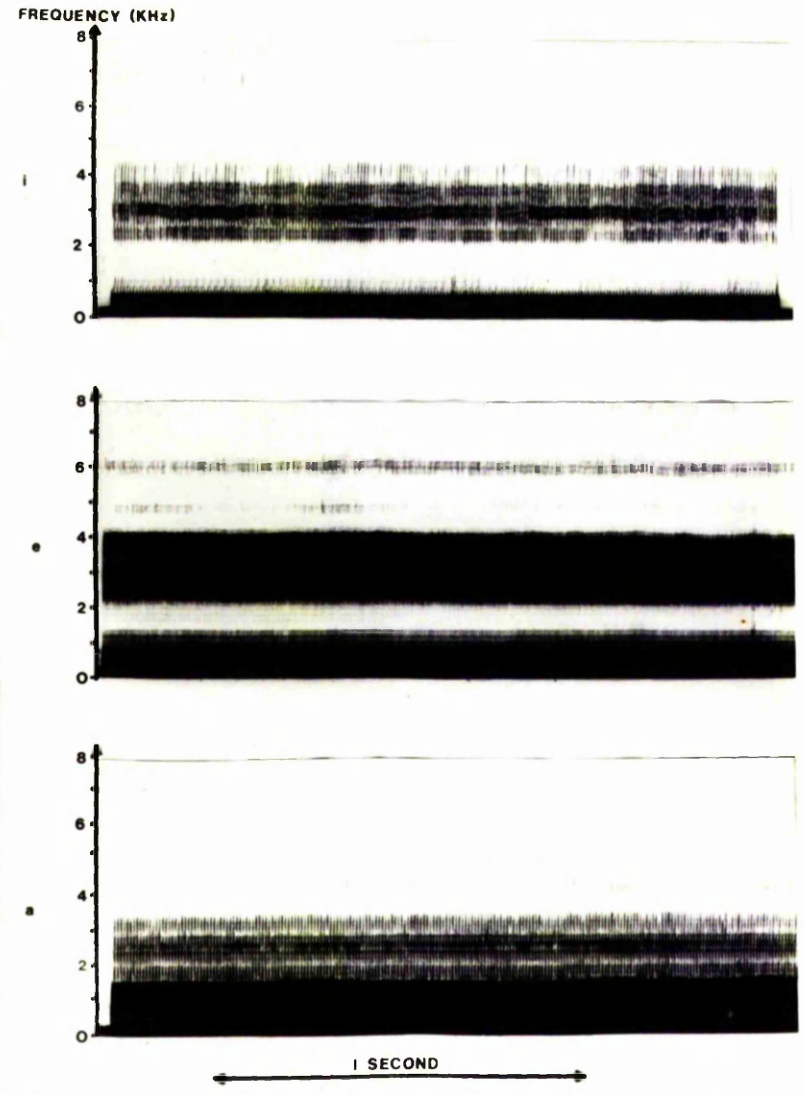
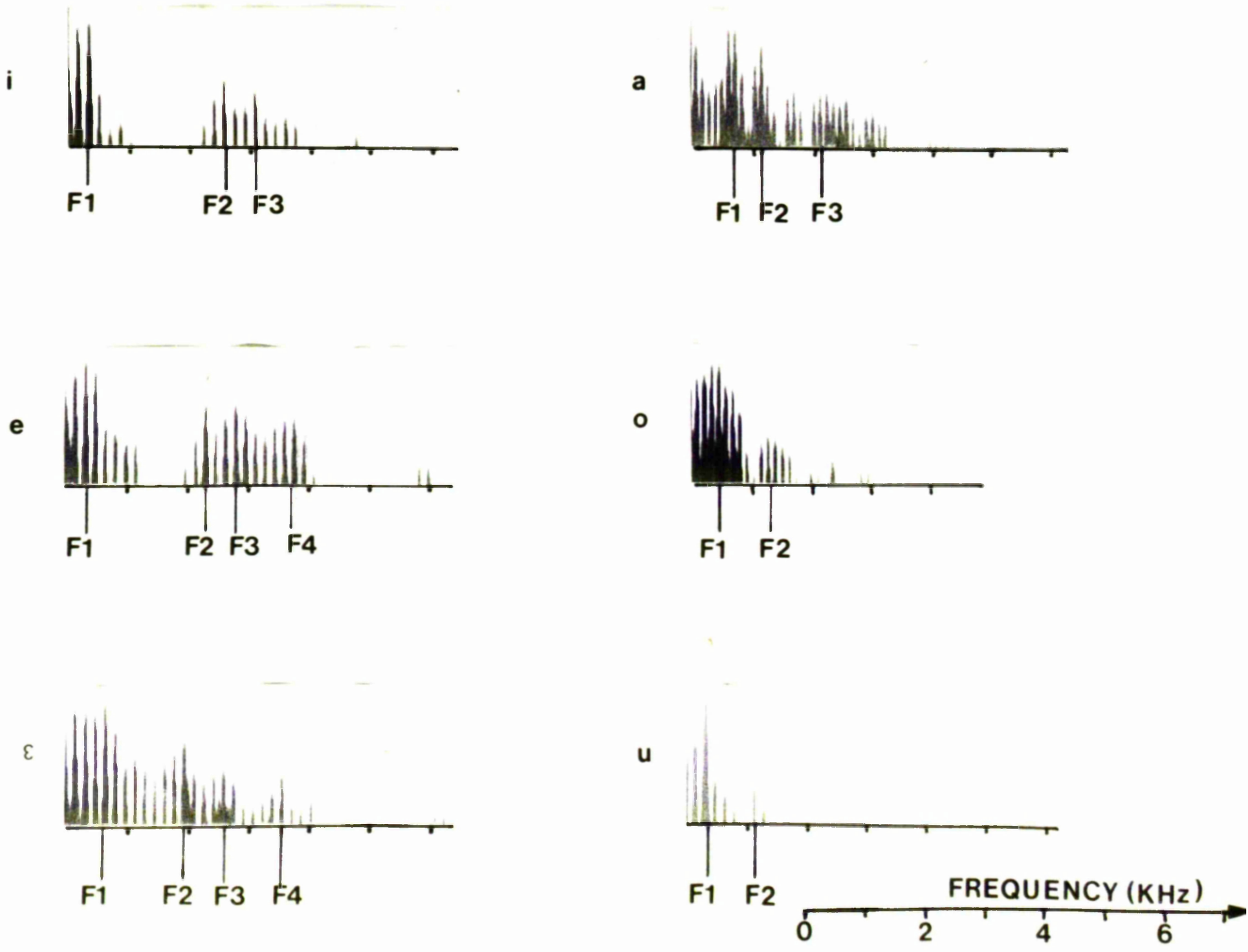


FIG.4.2 SONAGRAPHIC DISPLAY OF VOWELS /i/, /e/ and /a/ using 300Hz.Filter.


i
F1
F2
F3
a
F1
F2
F3
e
F1
F2
F3
F4
o
F1
F2
ε
F1
F2
F3
F4
u
F1
F2
FREQUENCY (KHz)
0
2
4
6

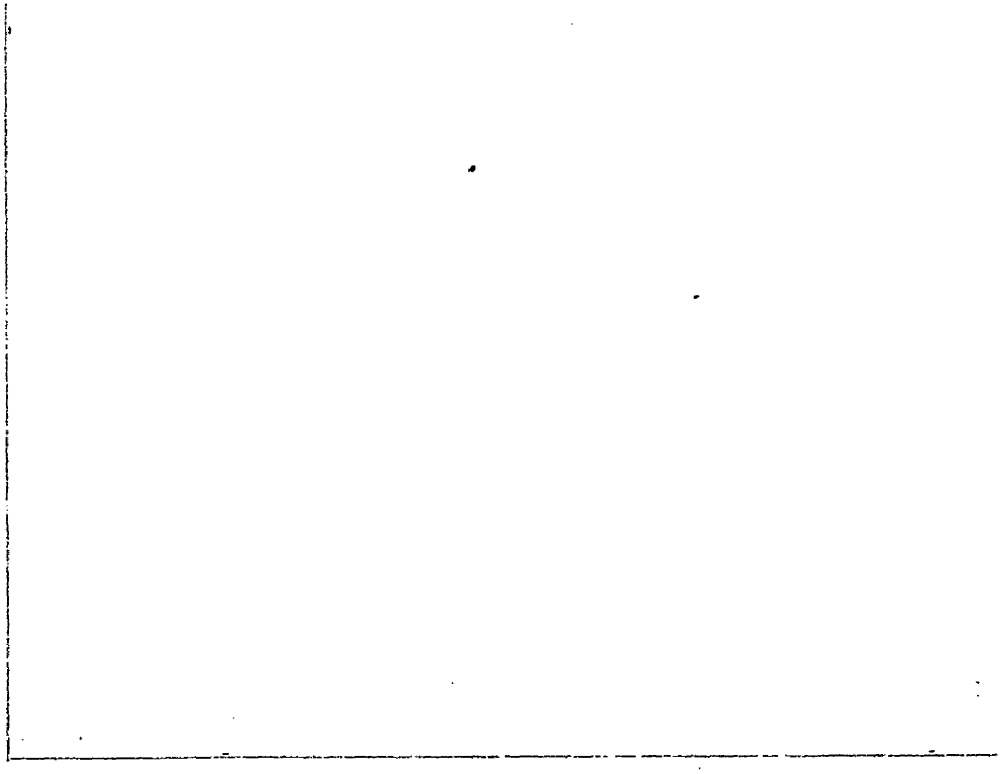


FIG.4.3 SECTIONER DISPLAYS OF 6 VOWEL SOUNDS SHOWING POSITIONS OF FORMANTS.

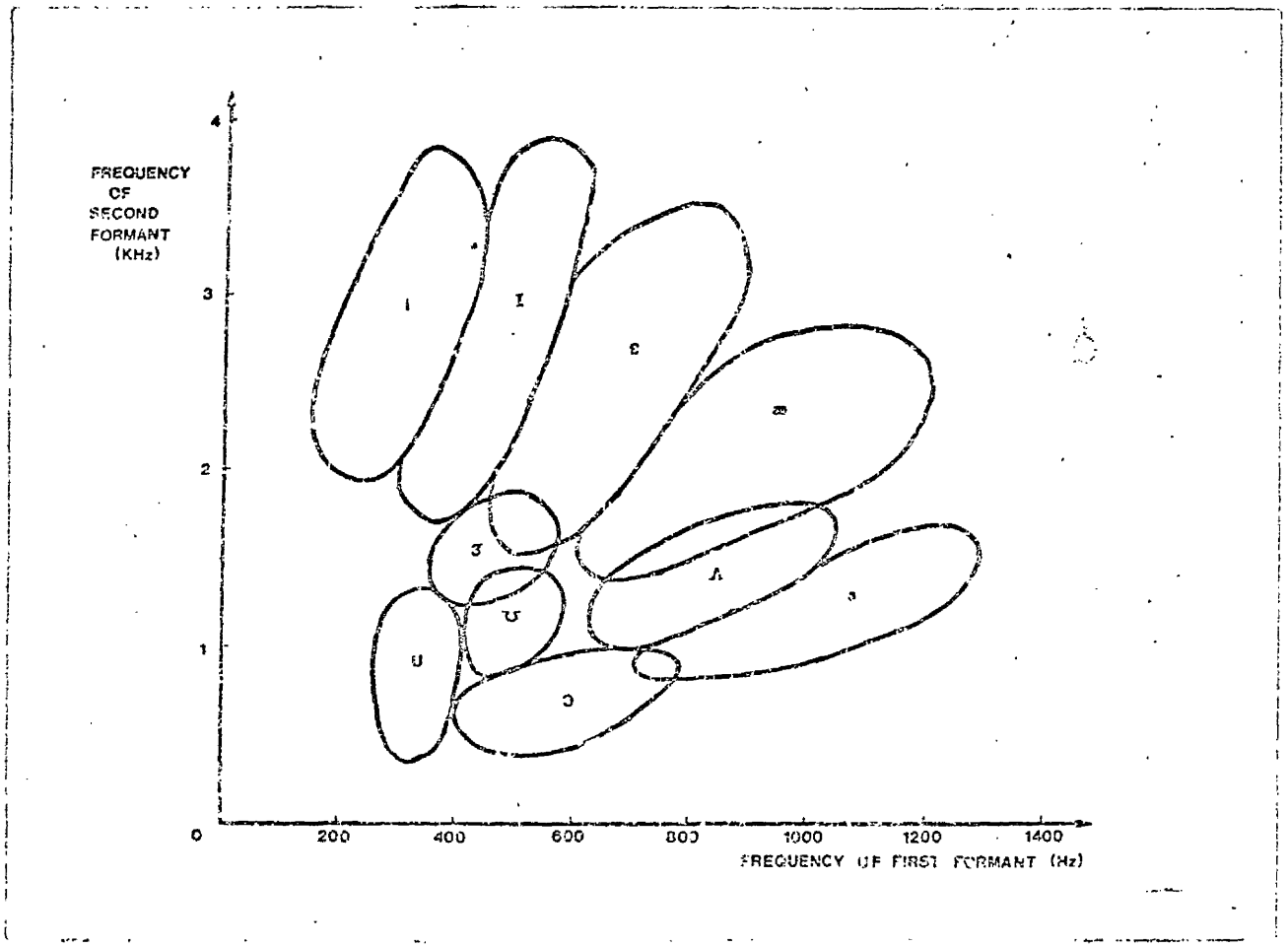


FIG.4.4 PLOT OF FREQUENCY OF FIRST FORMANT v FREQUENCY OF SECOND FORMANT FOR VARIOUS VOWEL SOUNDS. (PETERSON AND BARNEY, 1952).

that the absolute values of F1 and F2 were related to the dimensions of the vocal tract, i.e. children with shorter vocal tract lengths had resonances, and therefore formants, at higher frequencies than older males, who had larger vocal tract lengths.

In order to verify that these results could be reproduced using the present apparatus, a similar experiment was carried out. The series of vowels /i/, /e/, /ɛ/, /a/, /o/ and /u/ were recorded using the condenser microphone (as described in 3.2) from a group of 24 normal speaking adult subjects. Each subject was placed 20cms. in front of the microphone and required to sustain each vowel at a comfortable sound level and normal fundamental frequency for a period of at least 5 seconds. The recording of each vowel was analysed using the sonagraph. Sections were obtained with the 45Hz. filter and envelope curves were fitted by hand, and the formant frequencies measured. A plot of F1 v F2 was drawn and is shown in Figure 4.5. The results obtained were similar to those reported by Peterson and Barney, with the different vowel sounds occupying particular areas in the F1 and F2 space. However, in some cases there was considerable overlap between the areas corresponding to adjacent vowels, for example, the data for the vowel /e/ are not shown because of the overlap with those of the vowel /i/. The degree of overlap present suggested the importance of other factors, perhaps the third and higher formants, in the discrimination between similar vowel sounds. The formant values obtained for each vowel for several male members of the group are shown in Figure 4.6, and the closeness of the F1 and F2 values for the vowels /i/ and /e/ is still evident. However, the factors involved in vowel discrimination were beyond the scope of the present investigation, but the experiment demonstrated that the results reported previously could be reproduced by the available recording and

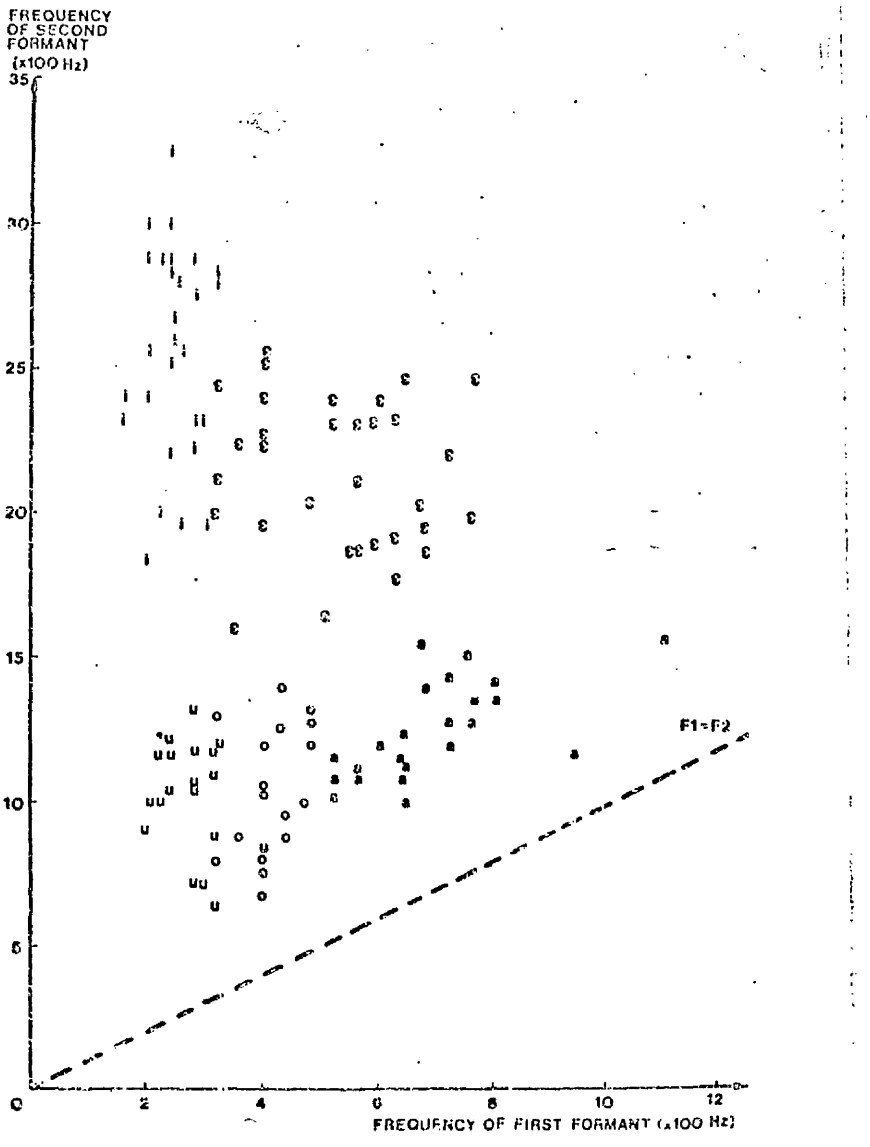


FIG.4.5 PLOT OF FREQUENCY OF FIRST FORMANT V FREQUENCY OF SECOND FORMANT FOR DIFFERENT VOWEL SOUNDS.

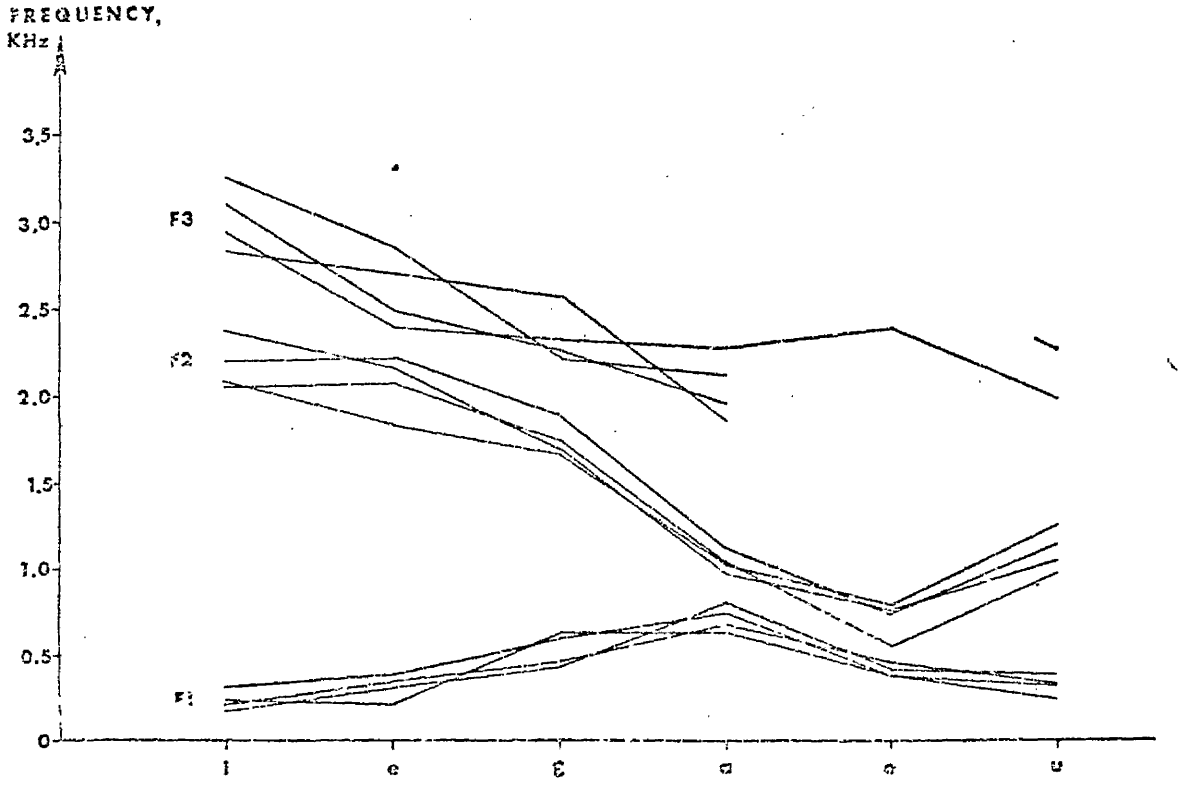


FIG.4.6 EXAMPLES OF FORMANT TRANSITIONS OF SOME OF THE MALE SUBJECTS.
M

analysis system. The differences observed were caused by differences in articulation between American and British subjects, and also contextual effects due to the transitions within the hVd words. The main errors were due to difficulties in fitting the envelope curve from which the formant frequencies were obtained. These were especially noticeable in the vowels of the subjects with high fundamental frequencies, e.g. females who thus showed a smaller number of harmonic components in a formant, and also in vowels where F1 and F2 were close, e.g. the vowel /a/. The experiment provided an initial background of experience in observing and interpreting the frequency spectra of vowel sounds.

4.3 VOICE ASSESSMENT FROM FREQUENCY SPECTRA

It was noted earlier that the frequency spectrum of a vowel sound normally consists of a series of harmonic components with a superimposed formant structure, and it was suggested that vowels subjectively classed as abnormal would show some alteration in the frequency components present in their spectra. According to Yanagihara (1967), these alterations were due to either the presence of random "noise" components between the harmonic ones, or to small but random changes in the fundamental frequency of phonation. ("Noise" components are noticeable in the spectrum of the vowel /a/ (Figure 4.3) especially at F1 and F2, and also at F2 and F3 of the vowel /ɛ/.) Yanagihara developed a classification of hoarseness on the basis of the spectral changes as seen on Sonagraphic outputs. "Hoarse" vowels were divided into four categories, defined as follows :

Type I: noise components are mixed with harmonic components, chiefly in the formant region.

Type II: the noise components in the second formant of the vowels /i/ and /ɛ/ predominate

over the harmonic components, and there are slight additional noise components in the high frequency region, above 3KHz.

Type III: the second formant of /i/ and /ɛ/ are replaced entirely by noise components, and there is also an increase in the abundance and intensity of the noise components above 3KHz.

Type IV: the first formants are replaced by noise components (especially for the vowels /i/, /ɛ/ and /a/), and there is an even greater degree of noise components at high frequencies. (Yanagihara, 1967).

The greater degrees of hoarseness corresponded to the gradual replacement of the harmonic components by noise components until at the most severe level, the harmonic structure, and even the formant structure was no longer evident. This type of classification was extended by the use of the sonagraph contour module to produce the "voice print" type of display (Iwata and Von Leden, 1970) which allowed a more quantitative analysis of the signal intensity in different frequency bands, and also of the continuity in time of the energy distribution. The increase in noise components and random variations in the fundamental frequency were observed as a broadening of the formant bars, which also became irregular, especially at the margins.

Three examples of standard sonagrams (300 Hz. bandwidth filter) of vowels which were subjectively described as hoarse are shown in Figure 4.7. The increase in the noise components are indicated by the irregularity in the pattern (C/f figure 4.2) especially in the ranges 6-8KHz. for /i/, and 2-3KHz. and 6-8KHz. for /ae/.

Figure 4.8 demonstrates the voice print output obtained from a subject suffering from myasthenia

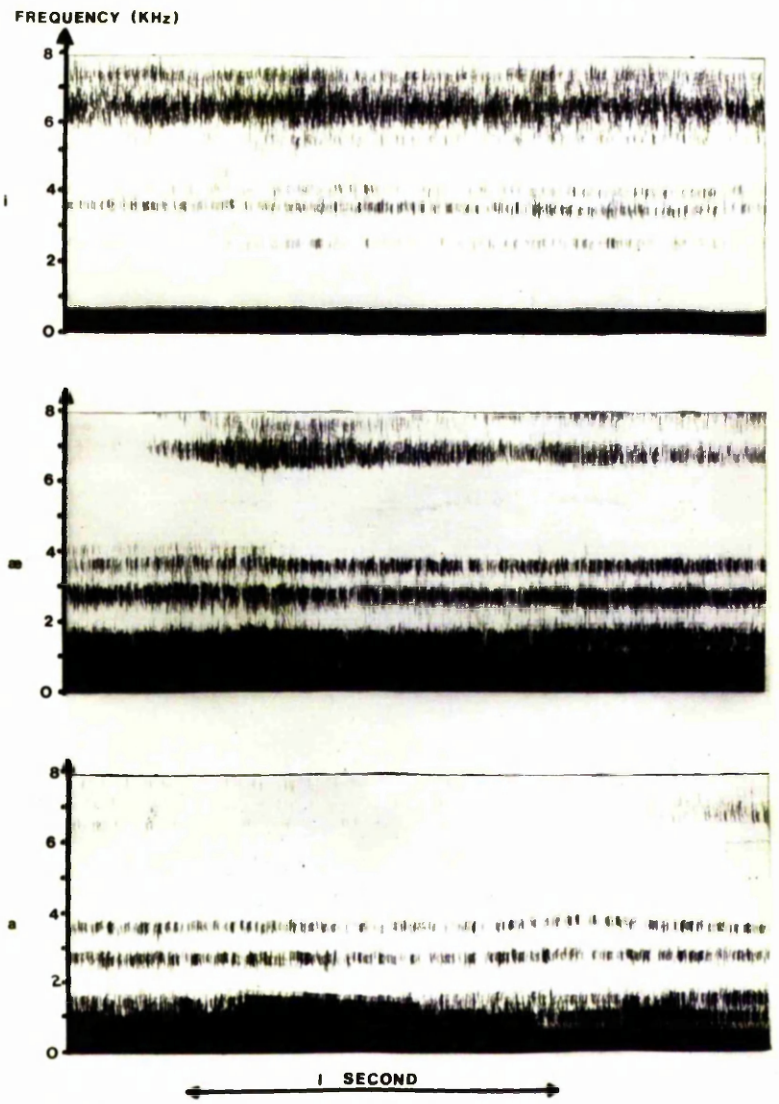


FIG.4.7 SONAGRAMS DEMONSTRATING INCREASE IN NOISE COMPONENTS IN "HOARSE" VOWELS.

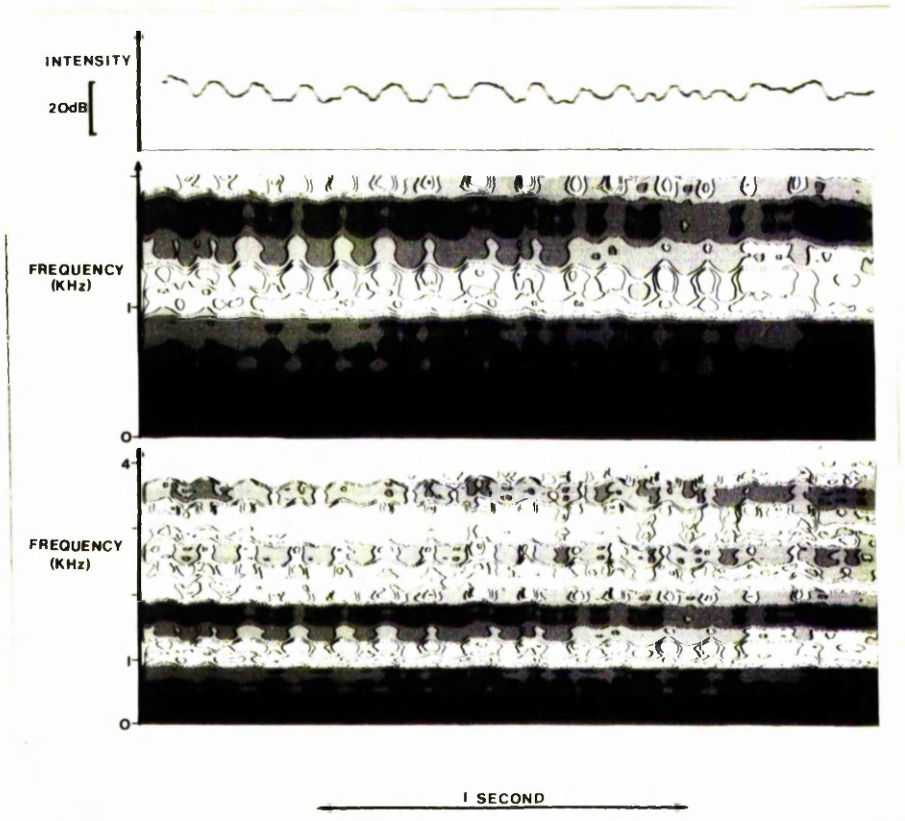


FIG.4.8 VOICE PRINT DISPLAY FOR SUBJECT SUFFERING FROM MYASTHENIA GRAVIS (/e/).

gravis, a neuro-muscular disease characterised by a rapid onset of weakness during sustained or repeated muscular contraction. Among the muscles first affected are those involved in jaw movement, swallowing, and the muscles of the face and neck. The sonagrams seen in the diagram show irregularities in the formant bands, especially in the second formant (1-2KHz.), and also in the overall intensity. However, the variations are not wholly random, but are at least partially periodic (frequency about 8Hz.) and may be due to the breakdown of a feedback control loop as a result of the disease.

The usefulness of this type of analysis in the routine assessment of abnormal voice production was investigated by firstly comparing the spectra obtained from a control group with those obtained from a group of dysphonic patients. The members of the second group were then classified according to Yanagihara's method, and the results compared with those obtained on the basis of the aerodynamic assessment developed in CHAPTER 3.

4.4 TYPES OF SPECTRAL ABNORMALITIES

The types of spectral changes observed in "abnormal" voice production were investigated by comparing the sonagrams obtained from a control group with those obtained from a group of dysphonic patients. The control group consisted of the 24 subjects referred to in SECTION 4.2, plus an additional 21 subjects who had subjectively normal phonation, giving a total control group of 45 subjects. The dysphonic group consisted of 100 subjects who were attending the speech therapy department for the treatment of various dysphonic conditions.

The series of vowels was recorded from each subject as described in SECTION 4.2 and, in the case of the patients, the recordings were obtained as part of the routine initial assessment described in CHAPTER 3. A frequency spectrum of each recording was obtained

using the sonagraph. Standard sonagrams were made using the 300Hz. bandwidth filter to allow the overall formant pattern to be observed, and more detailed information about the harmonic structure was gained using the 45Hz. bandwidth filter and the scale magnifier module. Section displays were obtained using the 45Hz. filter and the overall intensity was also recorded.

Each subject in the control group displayed a formant pattern similar to those shown in Figure 4.2 and 4.3, but 12 subjects in this group demonstrated the presence of some noise components between the harmonic ones. This occurred mainly in the vowels /i/, / ɛ/ and /a/, and the frequencies at which these components were evident were measured. The difference between the intensity of the noise component and that of the adjacent harmonic components was also measured in each case, and the data are tabulated in TABLE 4.1. It was seen from the table that, in general, the differences between the intensities of the noise components and the harmonic ones decreased with increasing frequency. This was consistent with the fact that the spectrum of a fricative noise source tended to increase in the high frequency region, whereas the harmonic source spectrum was a slowly varying function of frequency in this range, (Stevens and Klatt,1974). It also provided a guide to the amount of "noise" which could be present in the frequency spectrum, although still not be detectable by the ear, since no member of this group was subjectively classed as having an abnormal phonation. The overall voice intensities remained fairly constant (to within about 3dB) during each phonation.

In the dysphonic group of 100 subjects, 89 showed the presence of noise components to a greater or

TABLE 4.1. DATA ON SPECTRAL NOISE LEVELS OBTAINED FROM THE VOWELS RECORDED FROM THE CONTROL GROUP

SUBJECT	FREQUENCY AROUND WHICH NOISE COMPONENTS WERE EVIDENT (KHz)	INTENSITY DIFFERENCE BETWEEN NOISE COMPONENTS AND ADJACENT HARMONIC COMPONENTS (dB)
C 10	1.5	10
C 1	1.5	20
C 5	1.5	16
C 11	2.0	14
C 7	2.0	18
C 4	2.5	10
C 18	2.5	16
C 23	3.0	8
C 31	4.0	8
C 41	4.0	8
C 44	4.0	8
C 28	4.0	6

lesser degree and, in agreement with the findings of Yanagihara, in general, the higher noise levels were present in the sonagrams of the subjects who were subjectively assessed as having greater degrees of abnormality. 9 subjects showed spectral noise levels at all frequencies, even including the first formant, but it proved impossible to quantify the levels of noise components in the spectra, firstly because of their great abundance in some cases, and also because of their variability in time, even during the short time of the standard sonagraph display (2.4 or 4.8 seconds).

Several other types of spectral abnormalities were observed in the recordings made from this group, and some of these are best illustrated by the following examples.

4.4.A REGULAR MODULATIONS IN FUNDAMENTAL FREQUENCY AND/OR INTENSITY

Figure 4.9 shows the sonagrams of the vowel /o/ of a patient who was subjectively assessed as only slightly abnormal. There was no evidence of noise components, but the intensity display showed the presence of a regular modulation in the intensity of the phonation. The scale magnified output in the upper part of the figure indicated that there was a similar modulation in the fundamental frequency, and that this modulation was almost in phase with that of the intensity. Figure 4.10 shows sonagrams of the vowel /i/ phonated by the same subject, and similar phenomena were noted, with the difference that in this case, the modulations in intensity and fundamental frequency were almost exactly out of phase, i.e. as the fundamental frequency increased the intensity decreased.

These examples illustrated the interrelationship between fundamental frequency, and thus harmonic structure, and the formant pattern in the control of

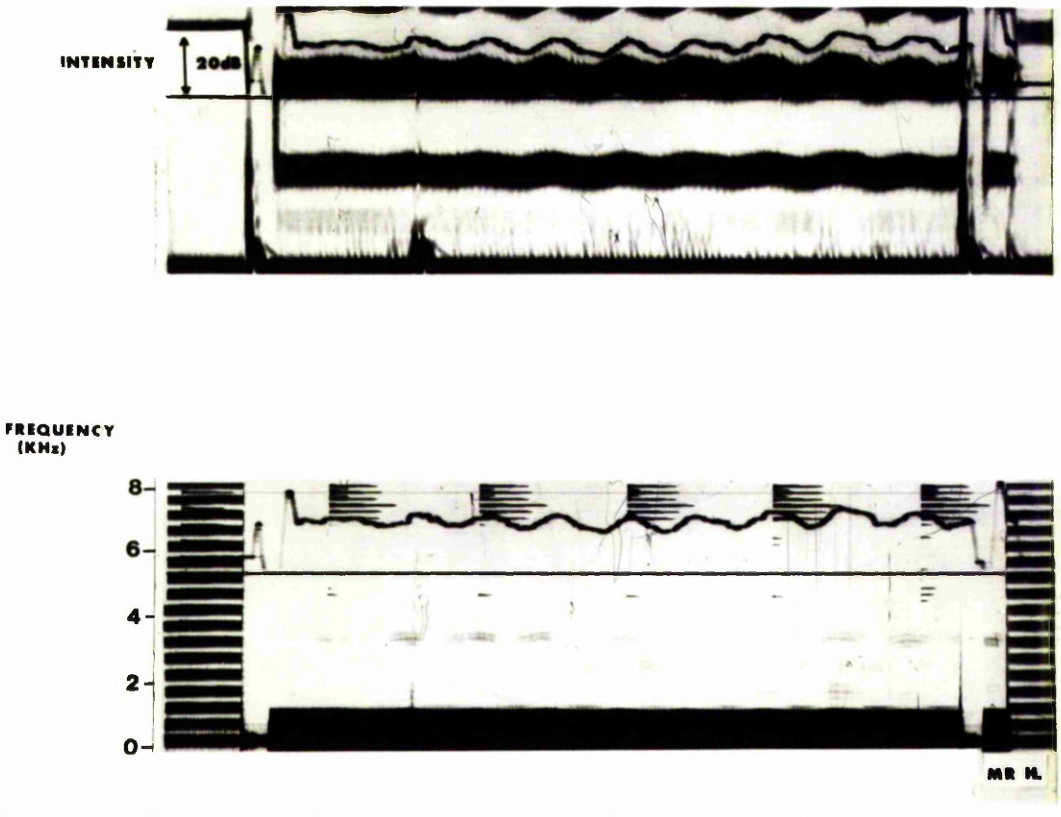


FIG.4.9 SONAGRAMS OF VOWEL /O/ SHOWING IN-PHASE MODULATION IN SOUND LEVEL AND FUNDAMENTAL FREQUENCY.

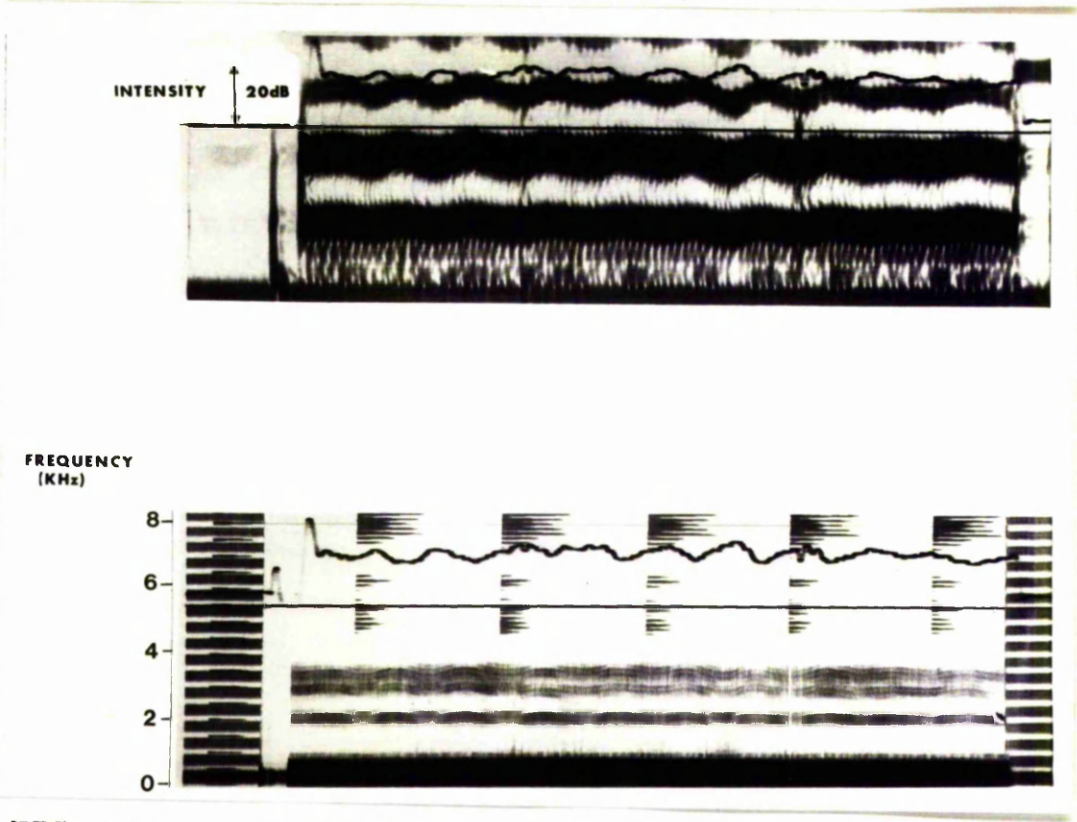


FIG.4.10 SONAGRAMS OF VOWEL /i/ SHOWING OUT-OF-PHASE MODULATIONS IN SOUND LEVEL AND FUNDAMENTAL FREQUENCY.

sound intensity. Since the most intense harmonic component need not correspond to the peak of the resonance or formant, then a small change in the fundamental frequency could result in the harmonic scanning the formant peak. As the frequency of the harmonic component approached the peak frequency, the intensity of the harmonic was amplified with respect to the adjacent harmonic components, and resulted in a net increase in the overall intensity. The reverse occurred as the harmonic moved away from the peak frequency, i.e. there was a net decrease in intensity. Thus if the initial harmonic frequency was lower than the formant frequency, an increase in the fundamental frequency would produce an increase in the intensity, and the modulations would be in phase. However, if the initial harmonic frequency was higher than the formant frequency, an increase in the fundamental frequency would cause a decrease in the intensity, and the modulations would be out of phase. The situation was further complicated by the presence of other harmonics and formants, which might act to nullify any intensity variations.

4.4.B RANDOM CHANGES IN FUNDAMENTAL FREQUENCY

Figure 4.11 illustrates the sonagrams obtained from the recordings of a patient who was subject to "pitch breaks", i.e., sudden changes in the fundamental frequency of phonation. In this case the vowel phonated was /a/, and it could be seen from the lower part of the figure, which was obtained using the 300Hz. bandwidth filter, that the phonation was irregular in frequency content, especially in the 1-3KHz. range during the latter portion of the recording. This was confirmed in the sectioner displays (shown in the central part of the figure) by the marked increase in the amount of noise components evident in the last three sections. These components were present even in the first formant, and virtually masked out the harmonic components. The upper part

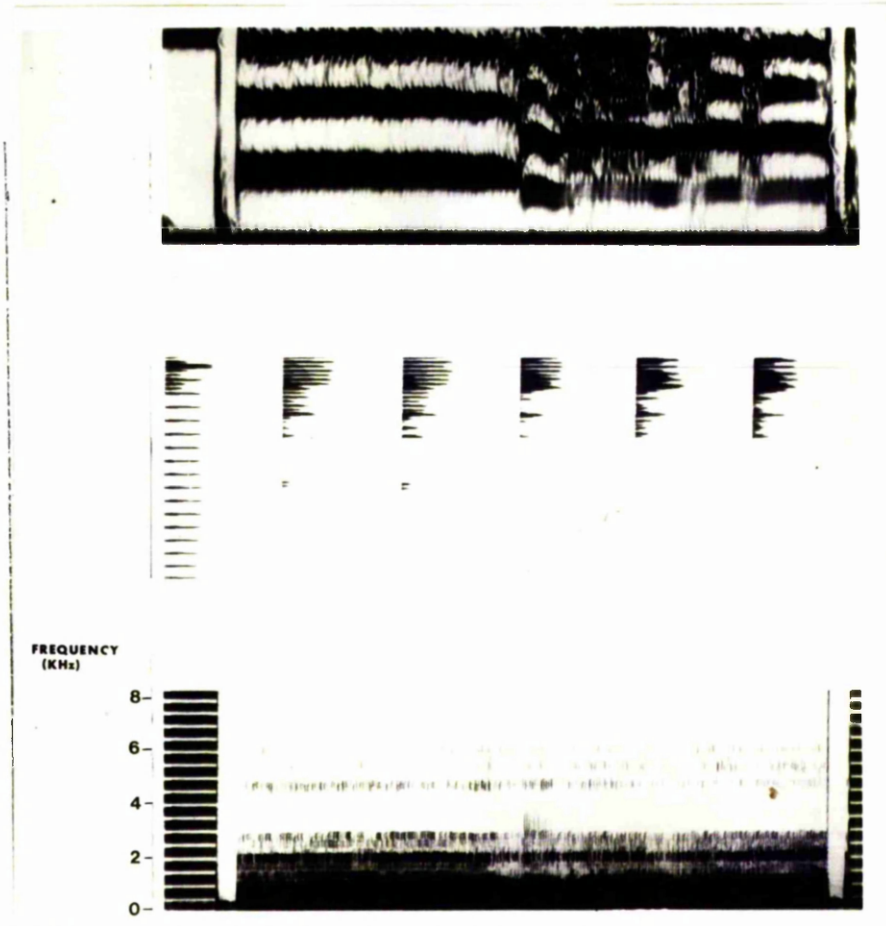


FIG.4.11 EXAMPLE OF SONAGRAM ILLUSTRATING A "PITCH BREAK".

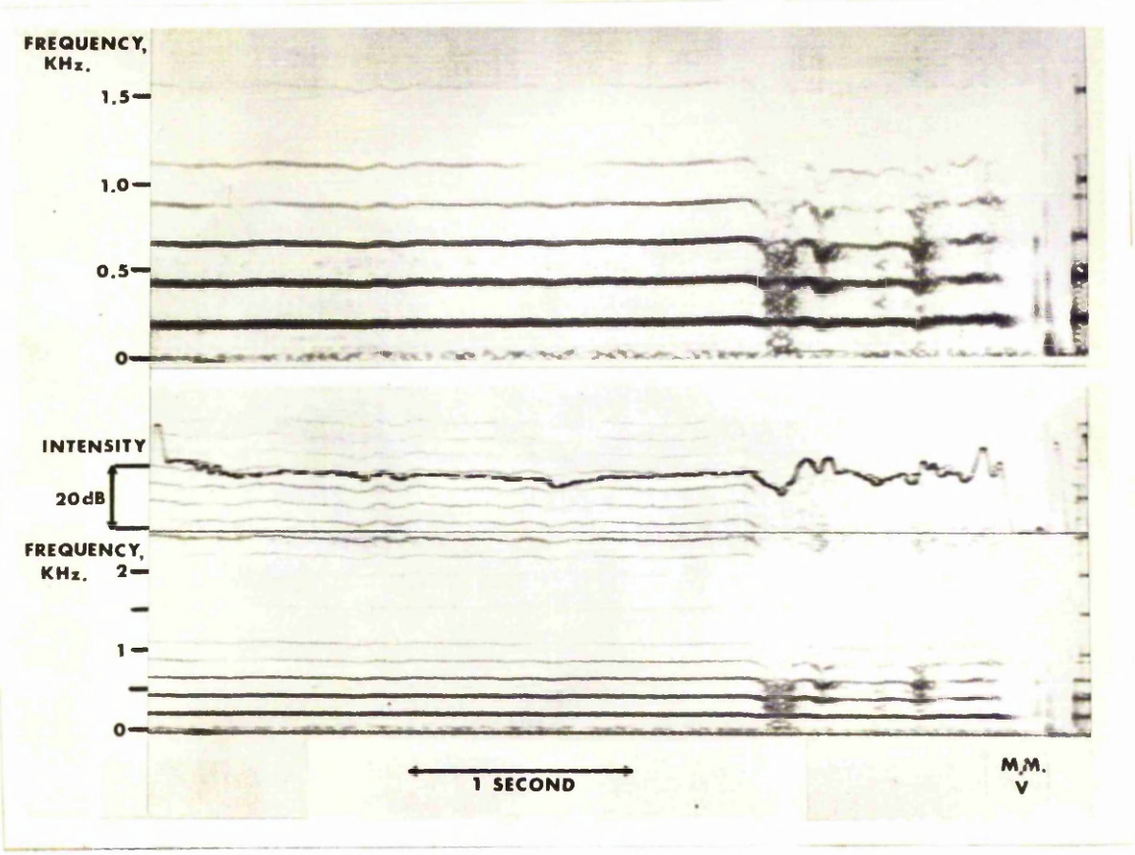
of the diagram showed the magnification of the frequency range 0-500Hz. using the 45Hz. bandwidth filter. This allowed the first few harmonic components to be observed, and it was noted that the increase in noise components coincided with a sudden change in the fundamental frequency of phonation from about 156Hz. to 111Hz. The phonation at the lower frequency was less stable than the higher one, as shown by its irregularity during the later stages of the phonation, as well as the increase in the random noise components.

4.4.C VOICE BREAK

A slightly different phenomenon is illustrated in Figure 4.12, that of a "voice break". The patient was being treated for mechanical dysphonia caused by vocal misuse of long standing. The phonation was initially maintained at a fairly constant sound level and fundamental frequency (220Hz.). At the first "voice break" (A), the harmonic components degenerated into noise components, with the intensity becoming uncontrolled and showing variations of the order of 10dB. A normal frequency pattern was resumed for a short time, to be followed by another voice break and the eventual cessation of phonation. Voice breaks can result from an abnormality of either or both vocal folds, which is severe enough to cause the vibration of the folds to cease, usually in the closed position, e.g. vocal nodules or some other irregular growth. Phonation is only resumed when the subglottic pressure is sufficient to force the folds apart and allow the vibration to be maintained.

4.4.D STABLE FREQUENCY COMPONENTS (INCLUDING SUB-HARMONIC COMPONENTS)

The diagrams shown have illustrated two types of spectral components, the standard harmonic components which were multiples of the fundamental frequency, and the random noise components which were completely independent of the fundamental frequency.



A

FIG.4.12 EXAMPLE OF SONAGRAMS ILLUSTRATING A "VOICE BREAK".

However, other regular spectral components were also observed, and some of these were at least partially stable, lasting for times of the order of seconds, and one such example is shown in Figure 4.13.

The phonation was of the vowel /a/ produced by a patient who was being given treatment for phonating with her false vocal folds, although this was never confirmed during a laryngeal examination. During the first part of the phonation, the fundamental consisted of a double component, (at about 500Hz. and 600Hz.) each giving rise to its own harmonic series, as shown in the section displays. The 500Hz component eventually coalesced with the 600Hz. one, and a more normal frequency spectrum resulted.

An example of the phonation of the vowel /ɛ/ by the same subject is given in Figure 4.14. The lowest frequency component initially had a frequency of 550Hz. but decreased to about 500Hz. as the phonation progressed (C). The component, initially at about 1.15KHz. divided into two components, only one of which (E) was harmonically related to C. The component A was maintained at a frequency of about 800Hz. and was related to component F, and perhaps B. Thus, two harmonic series were present, although in neither case was a component found which corresponded to the fundamental frequency, probebly due to a sharp fall off in the vocal tract transfer characteristics.

The origin of the unusual frequency components in the spectra of this patient was not clear. One possibility was that in this case the mechanical coupling produced between the vocal folds during vibration, was temporarily reduced, thus allowing the folds to vibrate fairly independently at different frequencies. Another mechanism was that there existed sufficient coupling between the vocal folds and the false vocal folds to bring the false folds into a state of vibration, and also produce excitation of the vocal tract. The fact that neither of these

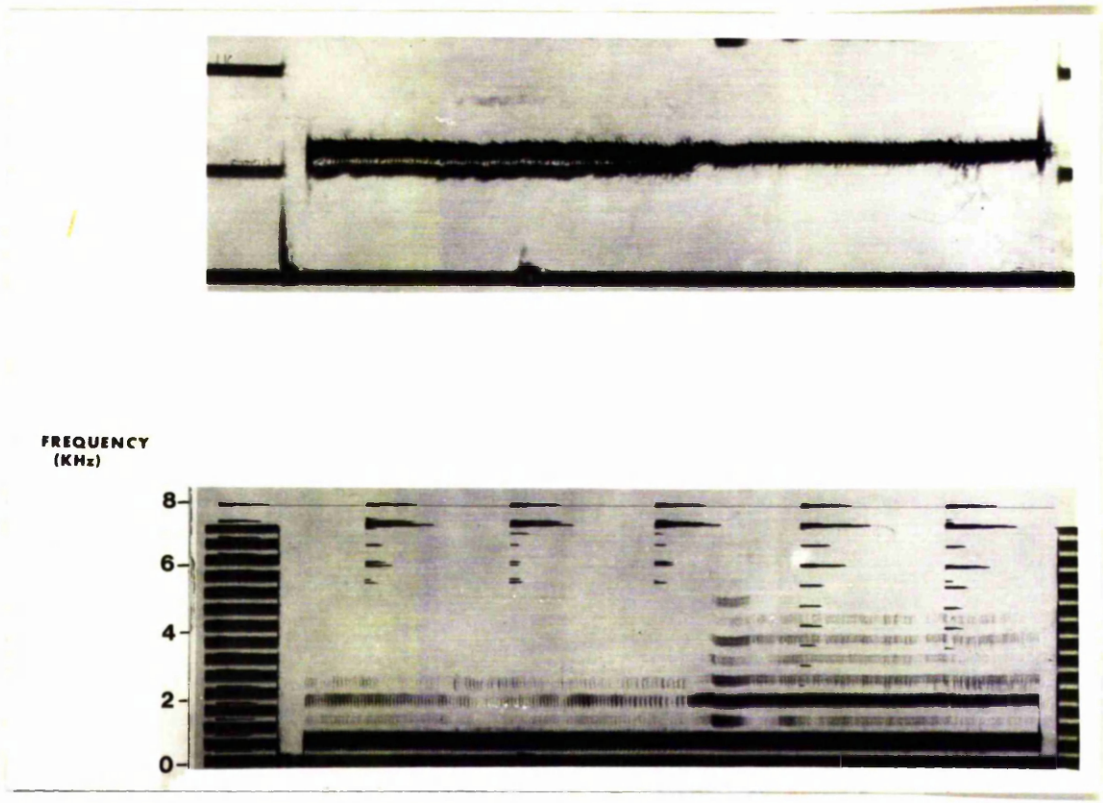


FIG. 4.13 SONAGRAM ILLUSTRATING A DOUBLE FUNDAMENTAL FREQUENCY COMPONENT.

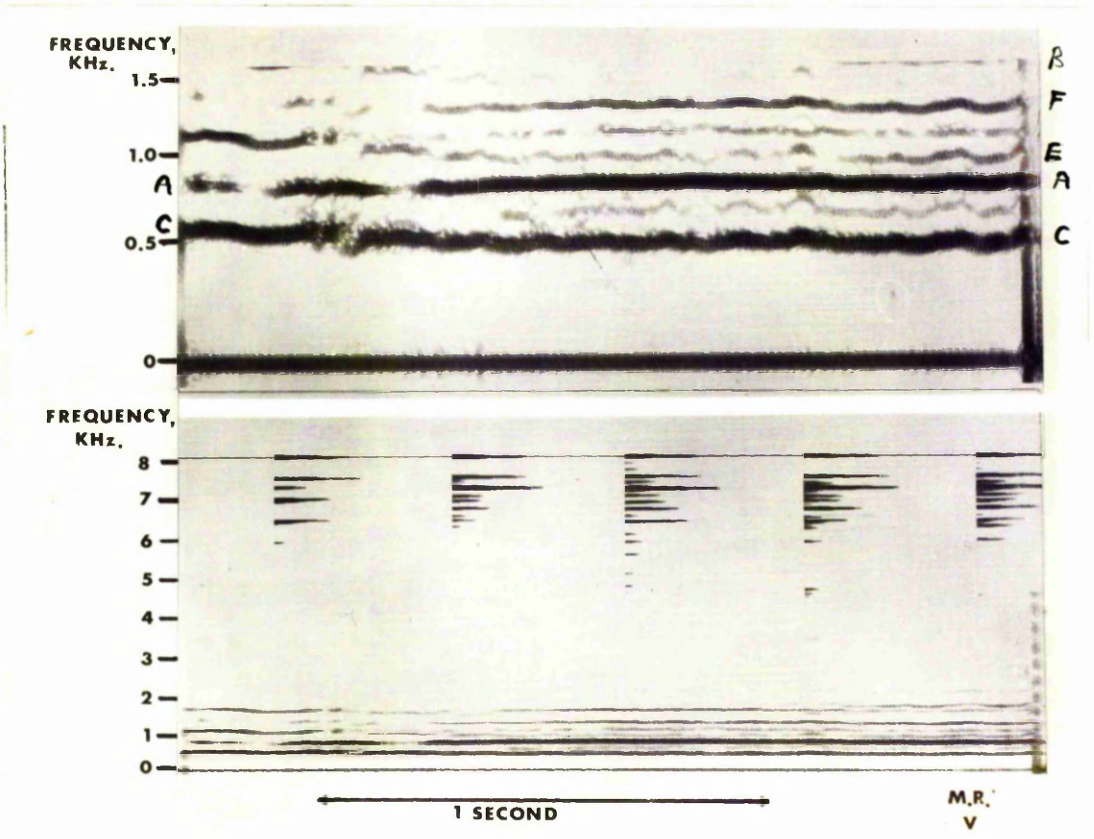


FIG.4.14 SONAGRAM ILLUSTRATING TWO INDEPENDENT HARMONIC SERIES.

hypotheses could be verified during a laryngeal examination did not invalidate them, since a laryngeal examination produced phonation under abnormal conditions, and could alter the vibratory pattern of either the vocal folds or the false vocal folds.

The sonagraphs of the vowel / ɛ/ shown in Figure 4.15 were obtained from the recordings of a patient who was subjectively assessed as being slightly adnormal. The fundamental frequency was 220Hz. and the sonagrams were fairly normal, apart from the component present between the second and third harmonics at a frequency of about 550Hz. Components between the first and second harmonics, and also below the fundamental were also evident in the section displays. These components which were fairly regular and much more intense than any noise components, seemed to be consistent with a series of components at sub-multiples of the fundamental frequency of the main series.

A similar phenomenon is illustrated in Figure 4.16. This patient was subjectively assessed as being highly adnormal, and the sonagrams of his recorded vowels showed the presence of noise components, voice breaks and a very irregular intensity of phonation. Also evident in the sonagrams were components which were maintained for about 0.5 second, (between A and B) and were consistent with sub-multiples of the fundamental frequency, i.e. sub-harmonic components.

Figures 4.17, 4.18 and 4.19 give further examples of the presence of sub-harmonic components. In these cases the sub-harmonic components were present only intermittently, but lasted for several seconds, e.g. A-B in Figure 4.17, C-D in Figure 4.18 and E-F in Figure 4.19, showing that the vibratory pattern was at least semi-stable. These components had intensities much greater than any noise components. It was also noted, by comparing adjacent

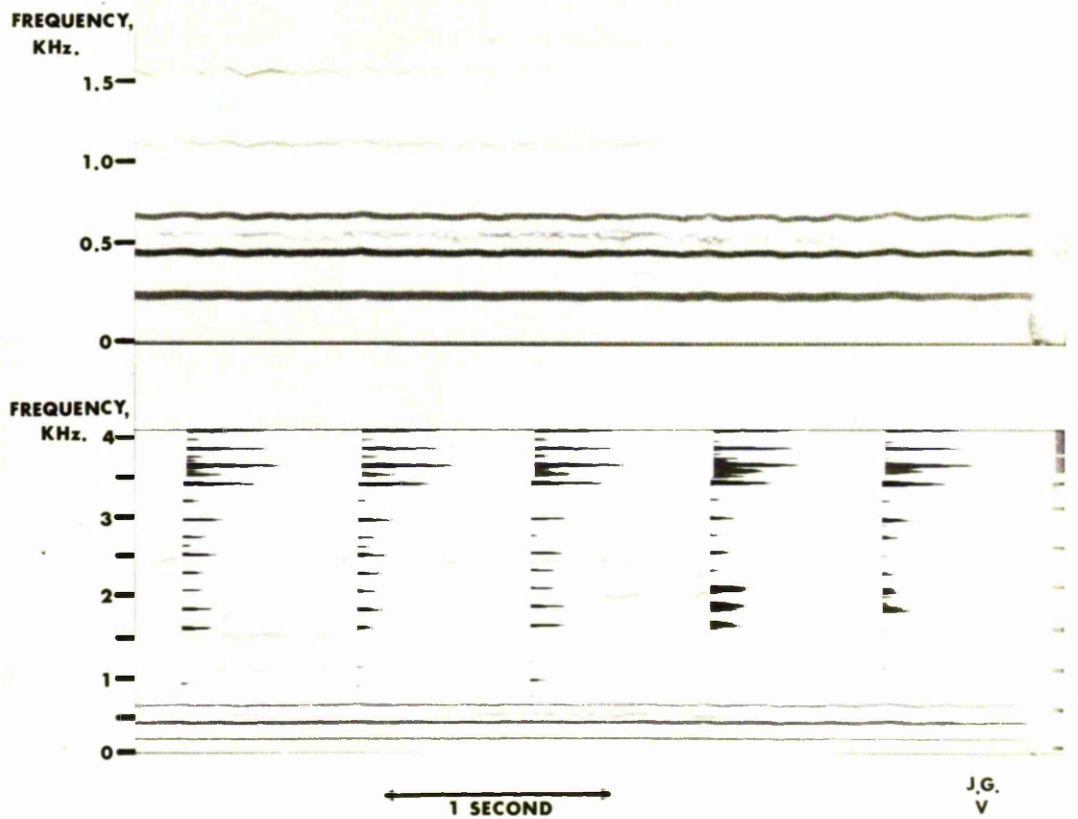


FIG.4.15 EXAMPLES OF SONAGRAMS ILLUSTRATING SUB-HARMONIC FREQUENCY COMPONENTS.

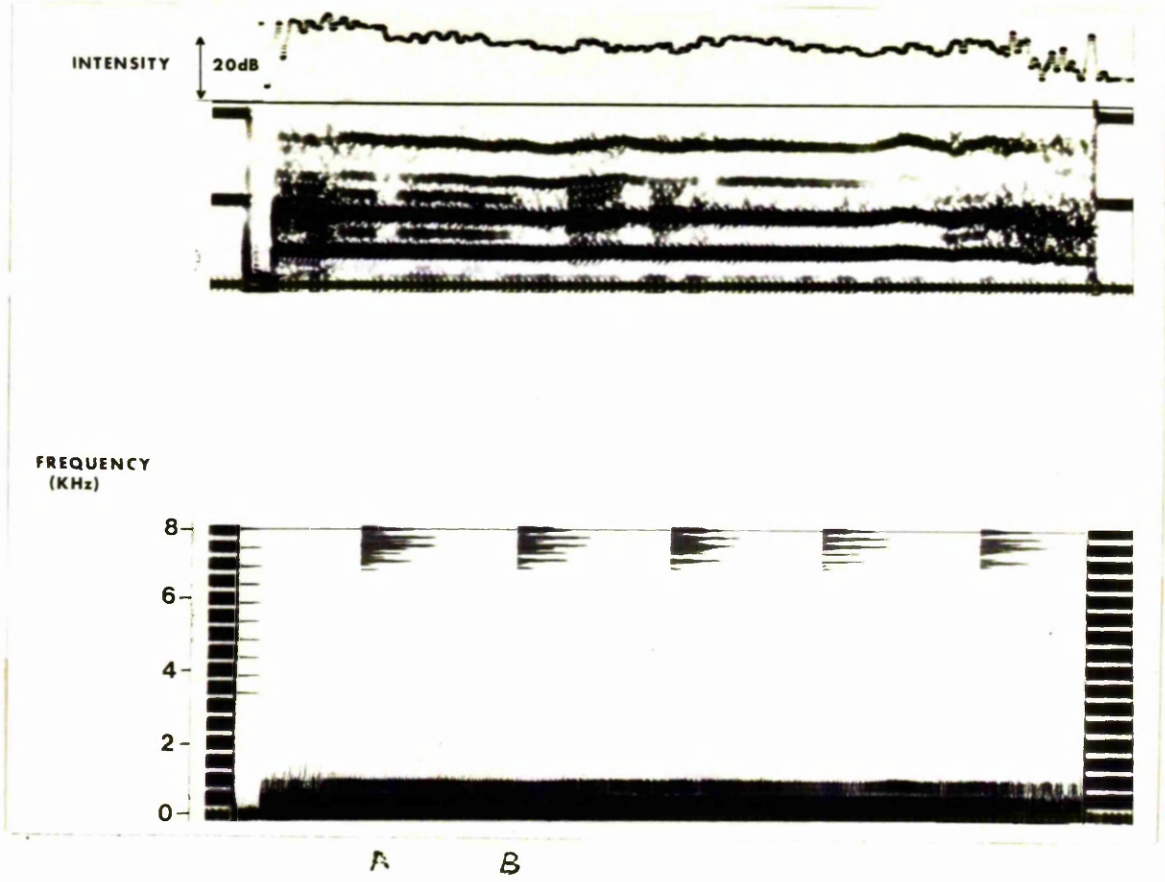


FIG.4.16 SONAGRAMS SHOWING NOISE COMPONENTS, VOICE BREAK, IRREGULAR INTENSITY AND ALSO THE PRESENCE OF SUB-HARMONIC COMPONENTS (A-B).

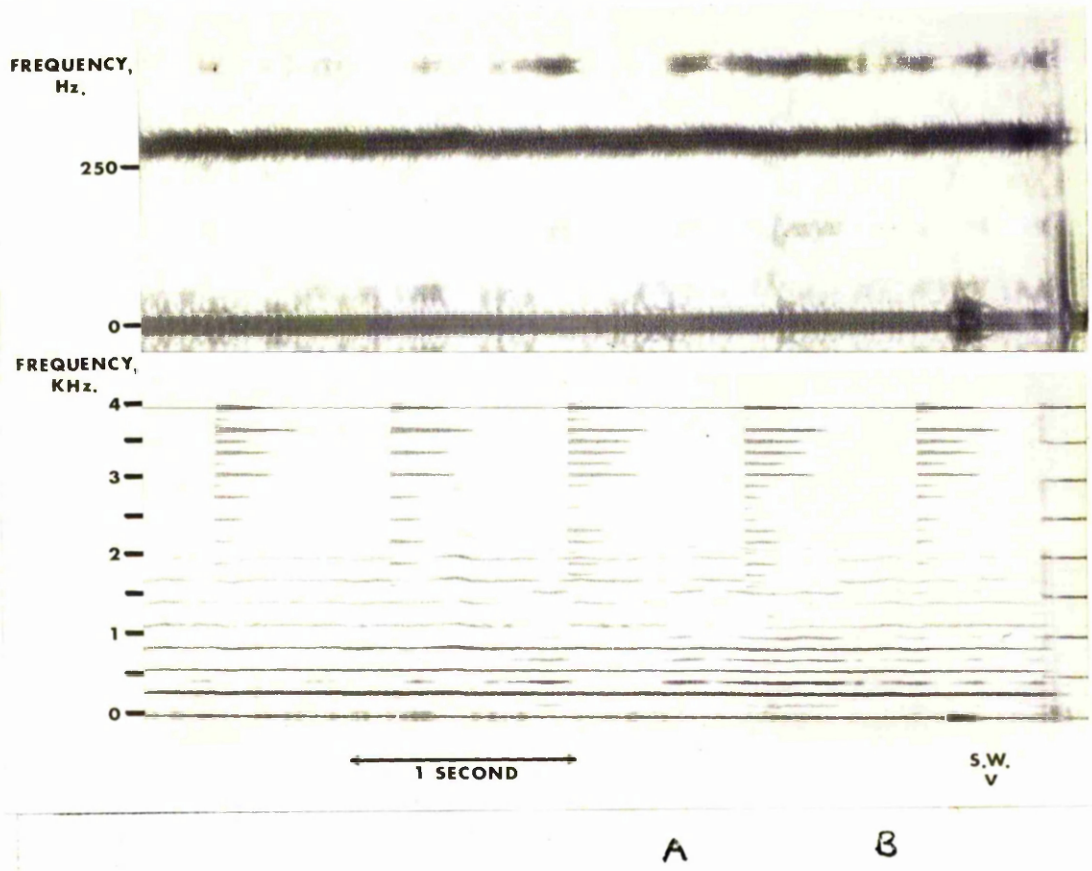


FIG.4.17 SONAGRAMS SHOWING PRESENCE OF SUB-HARMONIC COMPONENTS OF $\frac{1}{2}$Fo SERIES.

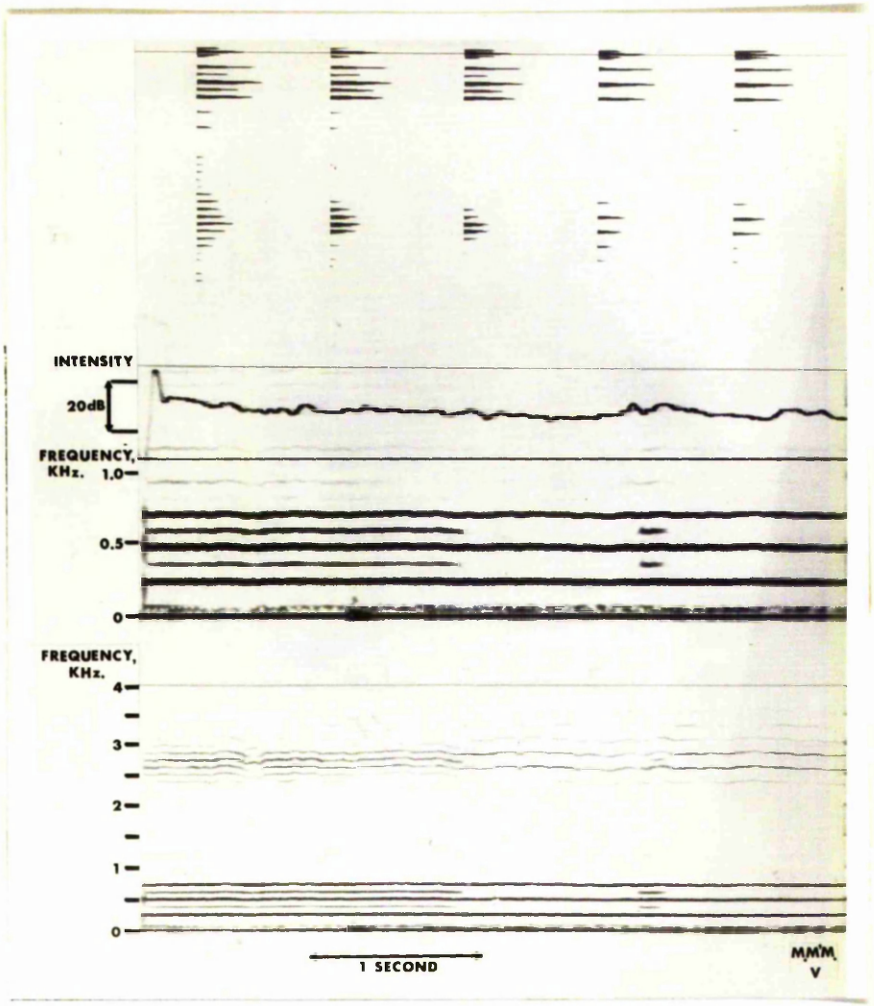


FIG.4.18 SONAGRAMS SHOWING PRESENCE OF SUB-HARMONIC COMPONENTS OF $\frac{1}{2}$Fo SERIES.

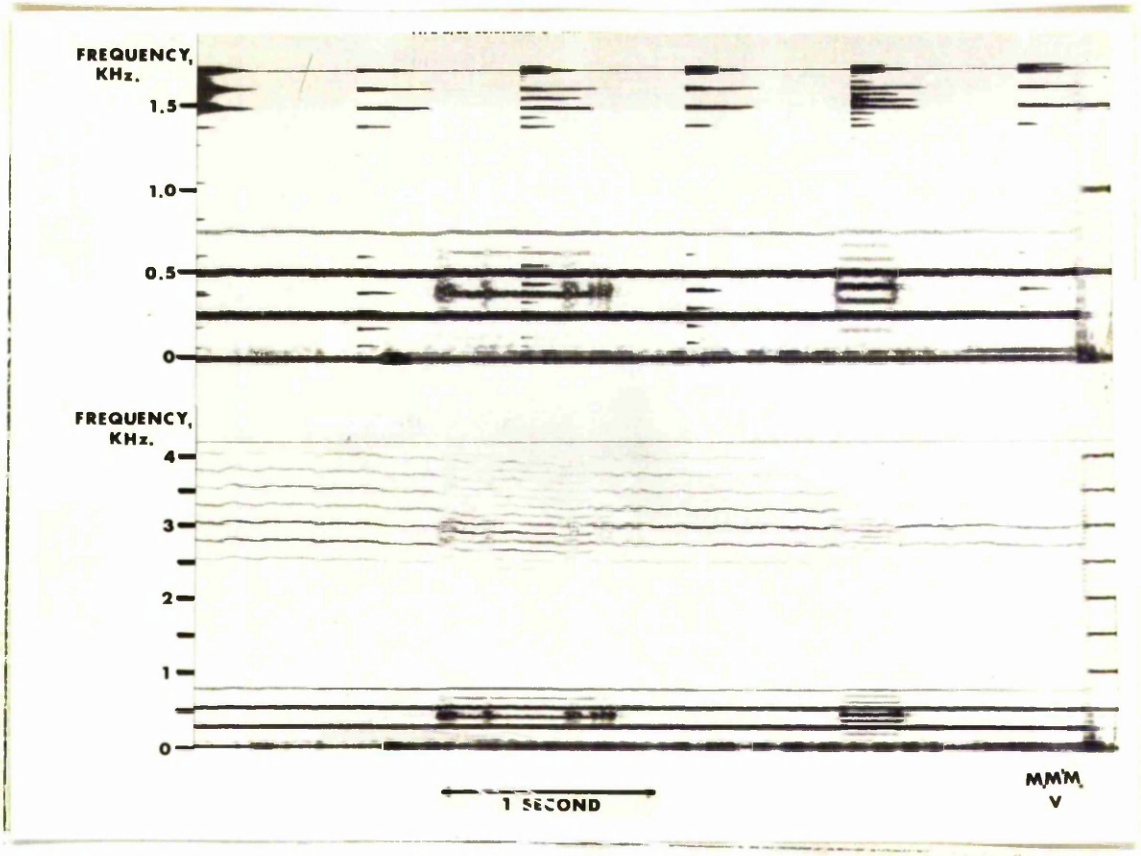


FIG.4.19 SONAGRAMS SHOWING PRESENCE OF SUB-HARMONIC COMPONENTS OF $\frac{1}{2}$Fo AND $\frac{1}{3}$Fo SERIES.

section displays that the presence of sub-harmonic components had no effect on the intensity of the standard harmonic components. There were only two occasions in which the frequencies of the sub-harmonic components did not correspond to the series $\frac{1}{2}F_o$, $\frac{3}{2}F_o$, $\frac{5}{2}F_o$, etc., and one is shown in Figure 4.19. At the point G, for about 0.25 sec., components of the series $\frac{1}{3}F_o$, $\frac{2}{3}F_o$, $\frac{4}{3}F_o$, $\frac{5}{3}F_o$, etc., were noted. On one occasion components of the $\frac{1}{4}F_o$ series were observed (SECTION 5.6), but sub-harmonics of higher order were never observed, but if any were ever present, it was unlikely that they would have been resolved with the 45Hz. filter unless the fundamental frequency was very high. Sub-harmonic components were reported by Seymour (1972) in some spectra of sung notes, but these occurred at frequencies around 5.7KHz. for the male voices and 6.8KHz. for the female voices, well above the frequencies discussed in the present investigation. The origin of sub-harmonic components will be discussed in CHAPTER 5.

It was found that sub-harmonic components were not uncommon in the spectra of the vowels recorded from the dysphonic group (being present in the spectra of 31 of 100 subjects). The presence of these and other spectral irregularities was noted for each member of the group, and the results compared with those obtained by using the aerodynamic assessment.

4.5 COMPARISON OF FREQUENCY SPECTRUM ASSESSMENT WITH AERODYNAMIC ASSESSMENT.

The spectrograms of the vowels recorded by the control group showed no unusual features apart from the presence of the noise components already discussed (SECTION 4.4). The members of the dysphonic group underwent aerodynamic assessment as discussed in CHAPTER 3., but there were incomplete data for 15 members of this group, which was thus reduced to 85.

The sonagraph outputs of each member were assessed with regard to Yanagihara's classification (SECTION 4.3), so that each subject was classified as Type I, II, III or IV according to the amount and distribution of noise components present in his spectra. If no noise components were present, the subject was classified type 0. The presence of the spectral abnormalities illustrated in SECTION 4.4 was also noted, these being

(a) regular changes in fundamental frequency
(b) random changes in fundamental frequency
(c) the presence of sub-harmonic components
(d) random changes in the overall intensity

The data from the Aerodynamic Assessment and Yanagihara classification for the dysphonic group are given in TABLE 4.2 A and B. As a result of the aerodynamic assessment of phonatory function, 6 subjects were assigned scores of 0, 20 of 1, 28 of 2, 23 of 3 and 8 had a score of 4. The Yanagihara spectral noise classification assessed 24 as being of type I, 29 of type II, 18 of type III and 6 as type IV. 8 subjects showed no evidence of spectral noise components and were classified as type 0. A direct comparison between the two assessments was made using the data tabulated in TABLE 2C. The contingency coefficient was calculated as C = 0.59 (APPENDIX 1.2) corresponding to a CHI^2 value of 45.6 ($p < 0.0005$, $CHI^2 = 41.3$ for 16 degrees of freedom). Thus there was a highly significant correlation between the results obtained by each assessment.

The data regarding the presence of the other spectral abnormalities (apart from random noise components) were tabulated according to the subject's aerodynamic assessment score, and these are given in TABLE 4.3A and B for those having scores of 0 and 1 respectively, TABLE 4.4 for those being assigned scores of 2 and TABLE 4.5A and B for those with scores of 3 and 4 respectively. It was noted from

TABLE 4.2 RESULTS OBTAINED ACCORDING TO AERODYNAMIC ASSESSMENT AND YANAGIHARA CLASSIFICATION.

(A) AERODYNAMIC ASSESSMENT OF PHONATORY FUNCTION.

TOTAL SCORE	0	1	2	3	4
NO.OF SUBJECTS	6	20	28	23	8

(B) SPECTRAL NOISE ASSESSMENT (YANAGIHARA).

TYPE	0	I	II	III	IV
NO.OF SUBJECTS	8	24	29	18	6

(C)

		YANAGIHARA ASSESSMENT TYPE				
		0	I	II	III	IV
AERODYNAMIC ASSESSMENT SCORE	0	2	4	--	--	--
	1	1	12	5	2	--
	2	4	6	12	4	2
	3	1	2	10	9	1
	4	--	--	2	3	3

TABLE 4.3 DATA ON PRESENCE OF SPECTRAL ABNORMALITIES IN SPECTRA OF SUBJECTS WITH AERODYNAMIC SCORES OF (A)0 and (B) 1.

	SUB-JECT	YANAGIHARA ASSESS. TYPE	REGULAR CHANGES IN Fo	RANDOM CHANGES IN Fo	PRESENCE OF SUB. HARM.COMP.	RANDOM CHANGES IN INTENSITY
A	18	I				
	48	I				
	169	I				1
	199	0				
	224	I				1
	229	0				
B	20	I			1	
	21	I			1	
	22	I				
	26	I			1	
	27	I				1
	36	I			1	
	42	II			1	
	58	I				
	66	III	1			1
	129	II			1	1
	171	I				1
	175	II			1	1
	186	II				
	187	II		1		
	189	I	1			1
	198	I				
	206	0				
	210	III			1	
	228	I				
	233	I				

TABLE 4.4 DATA OF PRESENCE OF SPECTRAL ABNORMALITIES IN SPECTRA OF SUBJECTS WITH AERODYNAMIC SCORES OF 2.

SUB-JECT	YANAGIHARA ASSESS. TYPE	REGULAR CHANGES IN Fo	RANDOM CHANGES IN Fo	PRESENCE OF SUB-HARMONIC COMPS.	RANDOM CHANGES IN INTENSITY
19	II	1			
25	0				
32	0				1
35	II				
41	II				
44	I		1		1
45	0				
49	II				
53	II	1			
55	III				
57	II				
69	III			1	
110	I				
112	II				1
116	III				
125	I				1
126	III				
173	II	1			1
181	0				
191	IV	1			1
192	IV				1
194	II	1			
195	II				1
200	II				1
201	I				1
205	I				
221	I				1
222	II			1	1

TABLE 4.5 DATA ON PRESENCE OF SPECTRAL ABNORMALITIES IN SPECTRA OF SUBJECTS WITH AERODYNAMIC SCORES OF (A)3 and (B)4.

	SUB-JECT	YANAGIHARA ASSESS. TYPE	REGULAR CHANGES IN Fo	RANDOM CHANGES IN Fo	PRESENCE OF SUB. HARMONIC COMPS.	RANDOM CHANGES IN INTENSITY
A	23	0	1		1	
	24	III		1		
	28	IV		1	1	1
	39	II		1		
	47	III		1		
	54	III		1	1	
	63	II		1		
	108	III				1
	109	II		1	1	
	111	II		1		1
	130	III				1
	178	II				1
	182	III	1		1	1
	183	II			1	1
	184	I				
	185	II				
	188	II	1			1
	193	III		1		1
	197	III				1
	208	III	1			1
	223	II			1	1
	225	II				
	235	I	1		1	1
B	33	II			1	
	38	IV				
	43	III				
	137	III				1
	190	IV		1	1	1
	203	IV		1	1	1
	226	II			1	1
	227	III	1			1

the tables that the presence of sub-harmonic components was fairly uniformly distributed over the groups with scores of 1, 3 and 4. Random changes in the overall intensity of the phonation were also common in subjects with aerodynamic assessment scores of 1, 2, 3 and 4. Patients with the higher aerodynamic scores of 3 and 4 tended to exhibit more than one of the spectral irregularities, and the presence of random changes in the fundamental frequency of phonation was particularly common in this group. This was consistent with the fact that the greater the degree of abnormality of the vocal folds themselves, then the greater would be the disruption of their vibratory pattern. This could be first indicated by the presence of sub-harmonic components alone, and then in combination with a loss in the control of the intensity, and in the worst cases to random changes in the fundamental frequency of phonation. The presence of sub-harmonic components alone did not seem to be of clinical value.

4.6 DISCUSSION

It was confirmed that abnormalities in vocal function produced detectable changes in the frequency spectra of the phonated vowel sounds. The results obtained from a group of dysphonic patients using Yanagihara's classification, which was based on the amount of random noise components present in the spectrum, showed close agreement with those based on the aerodynamic assessment of phonatory function (SECTION 3.6). However, the latter assessment was in general more rapidly carried out, and gave results which were more meaningful clinically, e.g. the aerodynamic assessment of phonatory function gave a direct indication of the adequacy of either breath control or phonation volume. Thus the use of the aerodynamic assessment was of greater value as a routine clinical investigation. However, the spectral studies did reveal the presence of abnormal frequency components

in the spectra of the vowels recorded from many dysphonic patients. Sub-harmonic components were found to be not uncommon, and their presence indicated an unusual vocal fold vibratory pattern. This is discussed further in CHAPTER 5. The occurrence of sudden changes in the fundamental frequency of phonation and large random fluctuations in sound level was found to be indicative of a greater degree of abnormality, and may also indicate an abnormal vocal fold vibratory pattern. Thus, although the presence of random noise components in the spectrum could be developed into a classification of hoarseness, there were other spectral abnormalities which could also give useful information about the phonatory function.

4.7 SUMMARY

In this chapter the general form of the frequency spectra of vowel sounds was discussed, including the origin of the formants. The roll of the formants in vowel discrimination was described. The Yanagihara classification of hoarseness was used to detail the degree of abnormality of a group of dysphonic patients and the result compared with those obtained using the aerodynamic assessment of phonatory function (SECTION 3.6). There was a close correlation between the two methods of vocal assessment ($p < 0.0005$), but the aerodynamic method was much more practicable as a clinical technique. However, the frequency analysis did lead to the observation of several types of spectral abnormality (as well as random noise components) which could be related to the vibratory pattern of the vocal folds.

CHAPTER 5.

THE VIBRATORY PATTERN OF THE VOCAL FOLDS.

CHAPTER 5. THE VIBRATORY PATTERN OF THE VOCAL FOLDS.

5.1 INTRODUCTION

The sound source during voicing arises from the periodic interruption of the air supply from the lungs, by the regular vibration of the vocal folds. Thus, it was of prime importance to the study of phonation to obtain a measure of the vibratory pattern of the folds, and to be able to relate the phenomena observed in such a pattern to the events occurring in the voice output. In SECTION 2.4 several methods of obtaining such information were briefly described, and it was concluded that the laryngograph technique was the only one suitable for use in the present investigation.

In this chapter, the interpretation of the laryngograph waveform (Lx) is discussed, and some of its characteristics investigated. The Lx waveform was used to study the onset of vocal fold vibration during the initiation of phonation. Frequency spectra of the Lx waveforms were obtained, and the properties of these were related to the closing phase of the vocal fold vibration. Abnormal frequency components present in the spectra of the voice output were attributed to the vocal fold vibratory pattern, and an explanation of the occurrence of subharmonic frequency components was suggested by the use of a simple model.

5.2 INTERPRETATION OF THE LARYNGOGRAPH WAVEFORM (Lx).

The operation of the laryngograph was discussed in CHAPTER 2. This was a current measuring device which measured the change in AC current (at a frequency of 1MHz) passing between two planar electrodes placed on the neck, as the vocal folds opened and closed. The Lx waveform consisted of three main portions; a steep increase, a less steep decrease and a relatively flat base, Figure 5.1. During the closing phase, the vocal folds were coming together

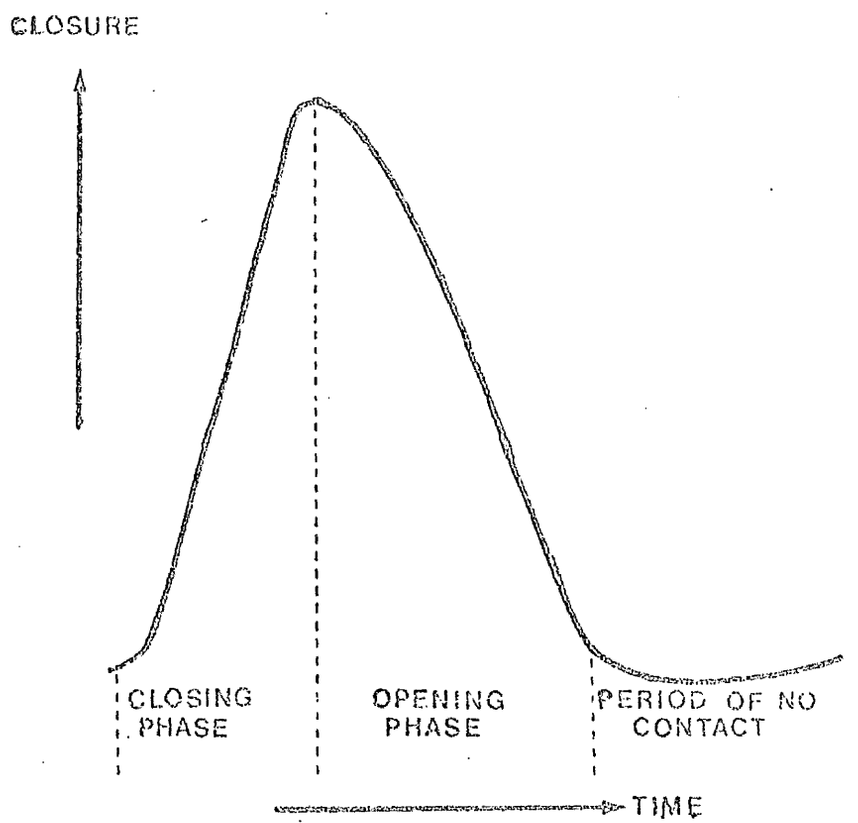


FIG.5.1 IDEALISED DIAGRAM OF 1 CYCLE OF Lx WAVEFORM.

causing an increase in their area of contact, and a decrease in the electrical impedance between the external electrodes, giving rise to an increase in the amount of AC current passed. In the opening phase, the vocal folds were moving apart, resulting in a decrease in the area of contact, and an increase in the electrical impedance between the electrodes, with a corresponding decrease in the amount of current being passed. The flat base of the Lx waveform corresponded to the open phase of the vocal fold vibration, during which time there was no contact between the folds. Thus the Lx waveform was a measure of vocal fold contact and not of either glottal area or glottal width, with which it had an anti-phase relationship. The Lx output was meaningful only during the phase when the folds were in contact.

The general correctness of the above interpretation of the Lx waveform was confirmed by Fourcin (1974), who correlated the different parts of the Lx waveform with glottal area as obtained by stroboscopic photography. However, Lecluse et al (1975) indicated that the relationship between the Lx waveform and vocal fold closure was more complex, as the glottis closed slightly before the maximum of the Lx and remained closed until slightly after the Lx peak. This was interpreted to reflect the fact that there was a small vertical motion of the vocal folds which may also have resulted in a change in the area of contact. However, from the Lx waveform, it was impossible to define precisely the moment of total closure, and in this thesis the definitions illustrated in Figure 5.1 were broadly adhered to, (except for one example given in SECTION 5.3).

Figure 5.2 shows an oscilloscope photograph of simultaneous microphone and Lx outputs from a subject who was phonating the vowel /a/ at a fundamental frequency of 110Hz (period 9.1msec). It was noted that the closing period was short compared to the opening period, (0.9msec), and that the closing period corresponded to the time of maximum excitation of the vocal

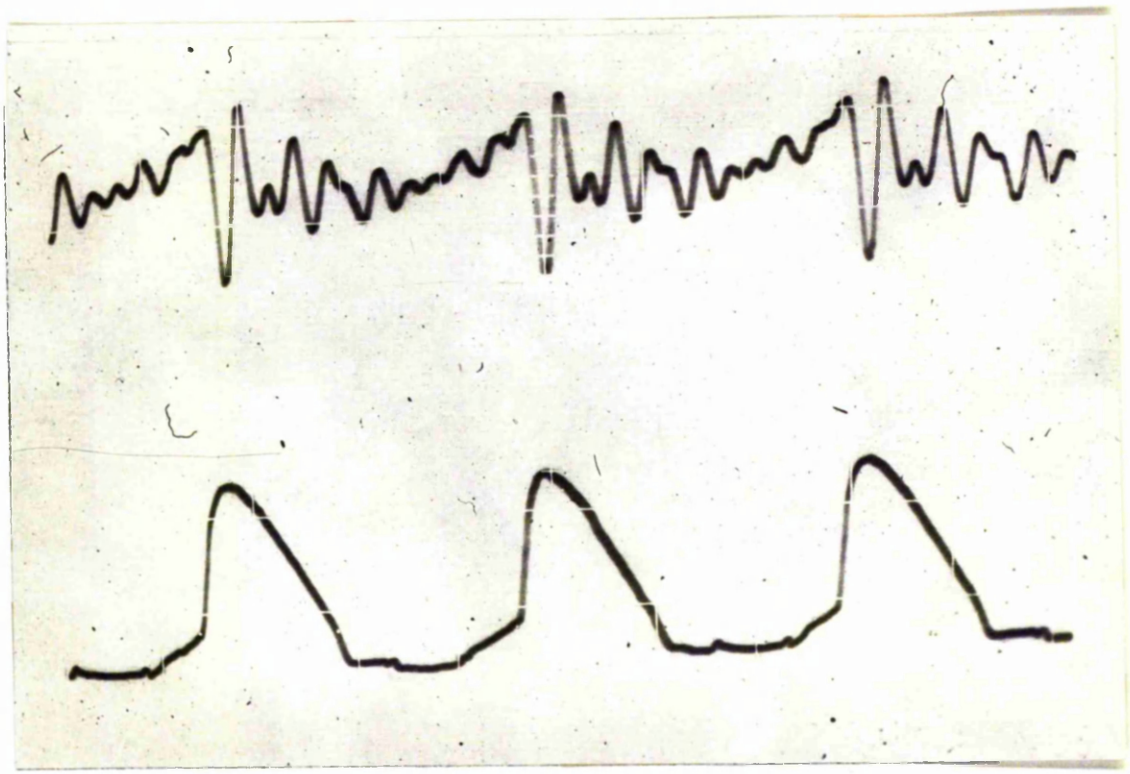


FIG.5.2 OSCILLOSCOPE PHOTOGRAPH, SIMULTANEOUS MICROPHONE OUTPUT (UPPER TRACE) AND Lx WAVEFORM (LOWER TRACE) DURING PHONATION OF THE VOWEL /a/.

tract (as seen by the large amplitude oscillations on the microphone output). This was found to be exclusively the case, and indicated the prime importance of vocal fold closure to the phonatory process.

5.3 TEMPORAL CHARACTERISTICS OF THE Lx WAVEFORM.

The Lx waveform could be described by the parameters of fundamental frequency (or period), length of the closing phase and the length of the opening phase. However, since the other methods of obtaining details of the vocal fold vibratory pattern measured either glottal width or glottal area, available data could not be used for direct comparison with the Lx parameters, so that normative data on these parameters had to be obtained.

A group of seven normal speaking adults (three males and four females) was investigated, and microphone and Lx recordings were made from each subject during the phonation of the vowels /i/, /e/, / ɛ /, /a/, /o/ and /u/. Each subject was placed 20cms from the condenser microphone and maintained each vowel in turn for between 5 and 10 seconds at a sound level of about 70db, and at his "normal" fundamental frequency in chest register. The microphone and Lx outputs were also recorded on an FM tape recorder (which preserved the shape of the Lx waveform), and these recordings were subsequently replayed through a U-V recording oscilloscope to obtain photographic records. Six cycles of the recording of each vowel were randomly selected and the closing times, opening times and open times measured, with the sweep speed of the oscilloscope being known. The mean results for each phonation are given in TABLE 5.1. The data indicated that there was no relationship between the Lx parameters and the different vowel sounds, and also that the parameters remained fairly constant for each subject. There was little intersubject variation

SUB-JECT	VOWEL	FUND. FREQ. (Hz)	CL.T. (M.SEC.)		O.T. (M.SEC)		OPEN TIME (M.SEC.)	
			MEAN	S.D.	MEAN	S.D.	MEAN	S.D.
A (M)	i	121	1.07	0.10	4.66	0.10	2.51	0.19
	e	119	1.02	0.07	5.10	0.10	2.31	0.15
	ɛ	119	0.91	0.05	4.75	0.21	2.75	0.19
	a	118	0.86	0.12	4.78	0.33	2.82	0.43
	o	119	1.32	0.05	4.59	0.12	2.49	0.06
	u	123	1.23	0.13	4.24	0.29	2.67	0.36
B (M)	i	132	0.56	0.05	5.21	0.12	1.83	0.15
	e	126	0.45	0.04	5.06	0.11	2.42	0.11
	ɛ	123	0.51	0.04	5.31	0.11	2.30	0.07
	a	125	0.43	0.02	5.09	0.20	2.47	0.21
	o	123	0.51	0.03	4.35	0.14	3.29	0.18
	u	127	0.50	0.01	4.74	0.11	2.61	0.09
C (M)	i	108	0.51	0.02	5.17	0.09	3.62	0.18
	e	112	0.68	0.07	5.13	0.27	3.13	0.34
	ɛ	114	0.56	0.04	5.02	0.21	3.20	0.16
	a	106	0.51	0.05	5.39	0.13	3.54	0.22
	o	103	0.42	0.02	5.13	0.17	4.15	0.16
D (F)	i	271	0.70	0.03	1.76	0.03	1.23	0.06
	e	283	0.67	0.04	1.24	0.08	1.62	0.08
	ɛ	275	0.63	0.04	1.44	0.09	1.57	0.08
	a	255	0.64	0.04	1.51	0.11	1.77	0.10
	o	262	0.69	0.02	1.40	0.13	1.73	0.11
	u	256	0.60	0.30	1.86	0.14	1.45	0.12
E (F)	i	282	0.56	0.03	1.42	0.03	1.56	0.07
	e	277	0.58	0.04	1.46	0.07	1.57	0.05
	ɛ	264	0.60	0.03	1.57	0.05	1.62	0.05
	a	245	0.59	0.03	1.57	0.06	1.92	0.10
	o	253	0.58	0.03	1.26	0.07	2.12	0.10
	u	245	0.41	0.03	1.46	0.13	2.12	0.14

TABLE 5.1

(CONTINUED ON NEXT PAGE)

SUB-JECT	VOWEL	FUND. FREQ. (Hz)	CL.T. (M.SEC.)		O.T. (M.SEC.)		OPEN TIME (M.SEC.)	
			MEAN	S.D.	MEAN	S.D.	MEAN	S.D.
F (F)	i	252	0.61	0.06	2.10	0.06	1.26	0.05
	e	251	0.60	0.08	1.88	0.09	1.50	0.12
	ɛ	254	0.49	0.03	1.94	0.11	1.51	0.12
	a	254	0.58	0.04	1.84	0.14	1.52	0.13
	o	258	0.67	0.03	1.70	0.10	1.51	0.10
	u	262	0.66	0.06	1.88	0.11	1.28	0.11
G (F)	i	276	0.42	0.03	1.32	0.09	1.88	0.09
	e	272	0.44	0.03	1.55	0.14	1.68	0.10
	ɛ	269	0.51	0.02	1.48	0.03	1.73	0.07
	a	270	0.49	0.04	1.49	0.12	1.73	0.12
	o	274	0.49	0.02	1.19	0.13	1.97	0.11
	u	273	0.46	0.03	1.30	0.14	1.90	0.14

TABLE 5.1 (CONTD).

MEAN VALUES OF CLOSING TIMES, OPENING TIMES AND OPEN TIMES. (S.D. - STANDARD DEVIATION)

M - MALE, F - FEMALE.

(apart from subject A) in the values of the closing time, which were in the range 0.41 - 0.70msec. regardless of the fundamental frequency. Thus the variation in period was accounted for mainly by variations in the opening time and open time, which became shorter at higher fundamental frequencies. The data in TABLE 5.1 were expressed as percentages of the phonation period and these are given in TABLE 5.2. Similarly the results for each subject were fairly constant. The male subjects B and C had closing times in the range 4 - 8% of the period, with opening times in the range 50 - 66%, and open times between 20 - 43%. The opening times were always greater than 7.5 times the closing times. The female subjects D,E,F and G showed similar trends, but the closing times corresponded to a much larger fraction of the period, 10 - 19%, because of the shorter period. The opening times were in the range 31 - 53% of the period, and the open times 31 - 54%. The opening times were always greater than twice, but less than 4.0 times the closing times.

The results obtained for subject A appeared to differ from the other data presented in TABLES 5.1 and 5.2, with the closing times being much longer than those of the other subjects. The main reason for this is illustrated diagrammatically in Figure 5.3. The Lx waveform obtained from this subject had a broad, rounded peak with the rise corresponding to vocal fold closure consisting of two sections, firstly a straight portion rising rapidly, AC, followed by a portion CB, which increased progressively more slowly to the peak at B. The results presented in TABLES 5.1 and 5.2 correspond to the definition of AB as being the closing time. However, the change in gradient at C was quite abrupt, and suggested that C represented the point of closure, and that from C to some point later than B, the change in the area of contact was due to the vertical movement of the vocal folds, as indicated by Lecluse et al (1975). The data for subject A were recalculated

SUBJECT	VOWEL	FUND. FREQ. (Hz)	CL. TIME / PERIOD %	O. TIME / PERIOD %	OPEN TIME / PERIOD %
A (M)	i	121	13.0	56.6	30.5
	e	119	12.1	60.5	27.4
	ɛ	119	10.8	56.5	32.7
	a	118	10.2	56.5	33.3
	o	119	15.7	54.6	29.6
	u	123	15.1	52.1	32.8
B (M)	i	132	7.1	65.9	23.2
	e	126	5.7	63.8	30.5
	ɛ	123	6.3	65.4	28.3
	a	125	5.4	63.7	30.9
	o	123	6.3	53.4	40.4
	u	127	6.4	60.4	33.2
C (M)	i	108	5.5	55.6	38.9
	e	112	7.6	57.4	35.0
	ɛ	114	6.5	57.2	36.4
	a	106	5.4	57.1	37.5
	o	103	4.3	52.9	42.8
D (F)	i	271	19.0	47.7	33.3
	e	283	19.0	35.1	45.9
	ɛ	275	17.3	39.6	43.1
	a	255	16.3	38.5	45.2
	o	262	18.1	36.6	45.3
	u	256	15.3	47.6	37.1
E (F)	i	282	15.8	40.1	44.1
	e	277	16.1	40.4	43.5
	ɛ	264	15.8	41.4	42.7
	a	245	14.5	38.5	47.1
	o	253	14.6	31.8	53.5
	u	245	10.0	35.8	52.0

TABLE 5.2

(CONTINUED ON NEXT PAGE)

SUB-JECT	VOW-EL	FUND. FREQ. (Hz).	CL.TIME % PERIOD	O.TIME % PERIOD	OPEN TIME % PERIOD
F (F)	i	252	15.4	52.3	31.7
	e	251	15.1	47.2	31.7
	ɛ	254	12.4	49.2	38.3
	a	254	14.7	46.7	38.6
	o	258	17.3	43.8	38.9
	u	262	17.3	49.2	33.5
G (F)	i	276	11.6	36.5	51.9
	e	272	12.0	42.2	45.8
	ɛ	269	13.7	39.8	46.5
	a	270	13.2	40.3	46.5
	o	274	13.4	32.6	54.0
	u	273	12.6	35.5	51.9

TABLE 5.2(CONTD)

THE DATA OF TABLE 5.1 EXPRESSED AS PERCENTAGES OF THE PERIOD.

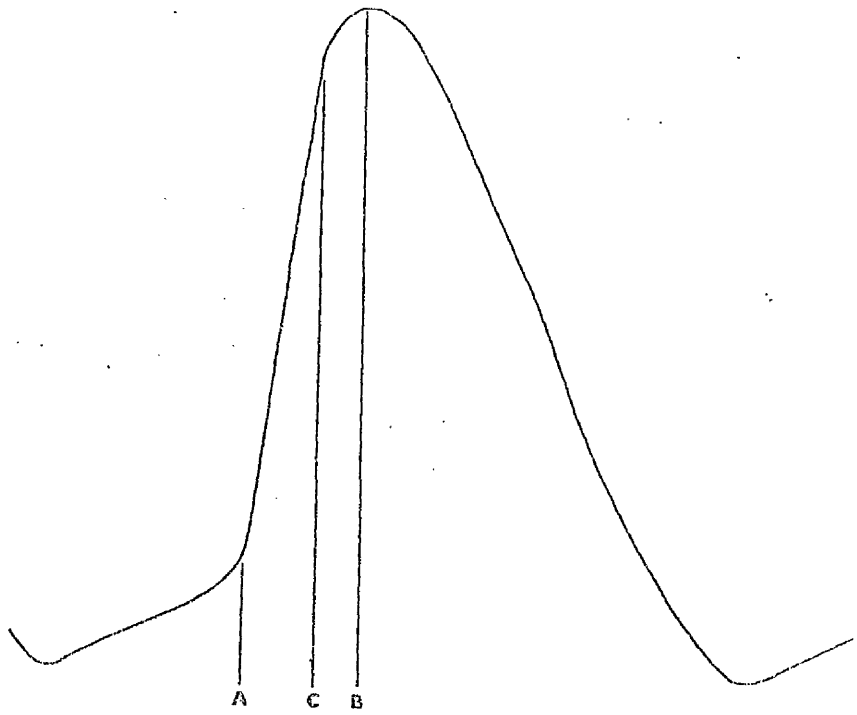


FIG.5.3 IDEALISED EXAMPLE OF THE Lx WAVEFORM OBTAINED FROM SUBJECT A.

using AC as the definition of the closing time, and are given in TABLE 5.3 for the vowels /i/, /e/, /a/, and /o/ in which such an Lx waveshape was evident. The closing times were not found to be in the range 0.66 - 0.74 msec. and represented between 7 and 9% of the period. These data and also those for opening times and open time were very similar to those obtained for the other male subjects. Thus the definition of the closing phase could be modified to be the time between the initial steep rise on the left of the peak to the point of discontinuity in the gradient, where the rise became much less steep. However, it was rarely possible to define such a point in practice, and the original definition was by and large maintained.

In previous studies, several indices to describe the vibratory pattern of the vocal folds have been defined :

$$\text{open quotient (O.Q.)} = \frac{\text{(time during which the glottis is open)}}{\text{period}}$$

$$\text{speed quotient (S.Q.)} = \frac{\text{(time of abduction of vocal folds (opening))}}{\text{(time of adduction of vocal folds (closing))}}$$

(Timcke et al, 1958)

and

$$\text{rate quotient (R.Q.)} = \frac{\text{(time the glottis is closed + time of opening)}}{\text{(time of closing)}}$$

(Kitzing and Sonesson, 1974)

None of these ratios could be related to the Lx parameters because the exact times of closure and of maximum opening were not strictly defined. However, the Lx results showed clearly that in the subjects investigated, the vocal folds closed much more rapidly than they opened, and this agreed with the results of Timcke et al (1958), (who used the high speed photographic technique) and Kitzing and Sonesson (1974) (who used the photoglottograph technique). In view of its relationship

VOWEL	FUND. FREQ. (Hz).	CLOSING TIME (M.SEC)		OPENING TIME (M.SEC)		OPEN TIME (M.SEC)		CL.T / PERIOD %	O.T. / PERIOD %	OPEN TIME / PERIOD %
		MEAN	S.D.	MEAN	S.D.	MEAN	S.D.			
i	121	0.74	0.03	5.14	0.05	2.37	0.14	9.0	62.3	33.4
e	119	0.70	0.02	5.41	0.11	2.29	0.20	8.3	64.4	33.2
a	118	0.74	0.03	4.93	0.40	2.82	0.43	8.7	58.1	27.3
o	119	0.66	0.03	4.94	0.09	2.81	0.03	7.8	58.7	28.7

TABLE 5.3

DATA FOR SUBJECT A, WITH MODIFIED DEFINITION OF CLOSING TIME

to vocal tract excitation, the most meaningful Lx parameter was the closing time (SECTION 5.2).

5.4 THE Lx WAVEFORM DURING VOCAL INITIATION.

Well controlled phonation was dependent upon a well controlled and stable vibration of the vocal folds. The Lx output was used to investigate how soon after the onset of phonation such a stable vibration was achieved.

Recordings similar to those described in SECTION 5.3 were made from a group of 19 normal speaking adults (7 males and 12 females), who were not aware of the purpose of the experiment. The microphone and Lx recordings were then replayed onto the Mingograf ink jet recorder, with the tape recorder being run at 3¾ i.p.s. (half of the recording speed) and the Mingograf at 500 mm/sec., giving a time scale of 10 msec./cm on the chart paper. Two typical examples are illustrated in Figure 5.4, with, in each case, the microphone output being shown on the upper trace and the Lx output on the lower. In both cases the vocal fold vibration, as indicated by the Lx output, built up gradually to a steady amplitude. The point at which the amplitude of the Lx signal became virtually maximal, was taken at the point at which a stable vibratory pattern was achieved. In case A, the vowel /u/ phonated by a male subject at a fundamental frequency of 154Hz, this was attained within 12 cycles (80 msec.) after vocal initiation, as defined by the points U and V. In case B the vowel /i/ produced by a female subject at a fundamental frequency of 205Hz, about 14 cycles (66 msec) were required to reach a steady Lx amplitude (XY). The number of cycles and times required to reach a steady Lx amplitude were measured for each phonation. There was found to be no consistent difference in the results obtained for the different vowels, and so the data were grouped together. Figure 5.5 shows the frequency histogram for the number of cycles required to

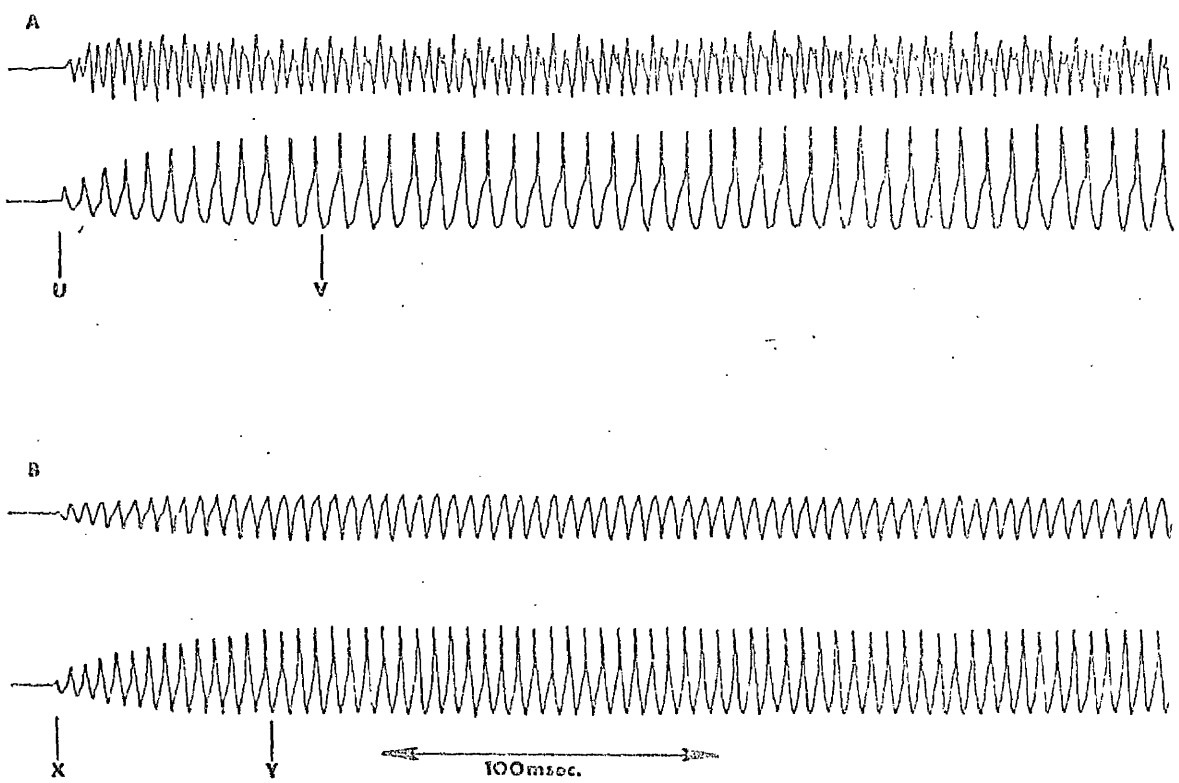


FIG.5.4 TYPICAL EXAMPLES OF MICROPHONE AND Lx OUTPUTS DURING VOCAL INITIATION.

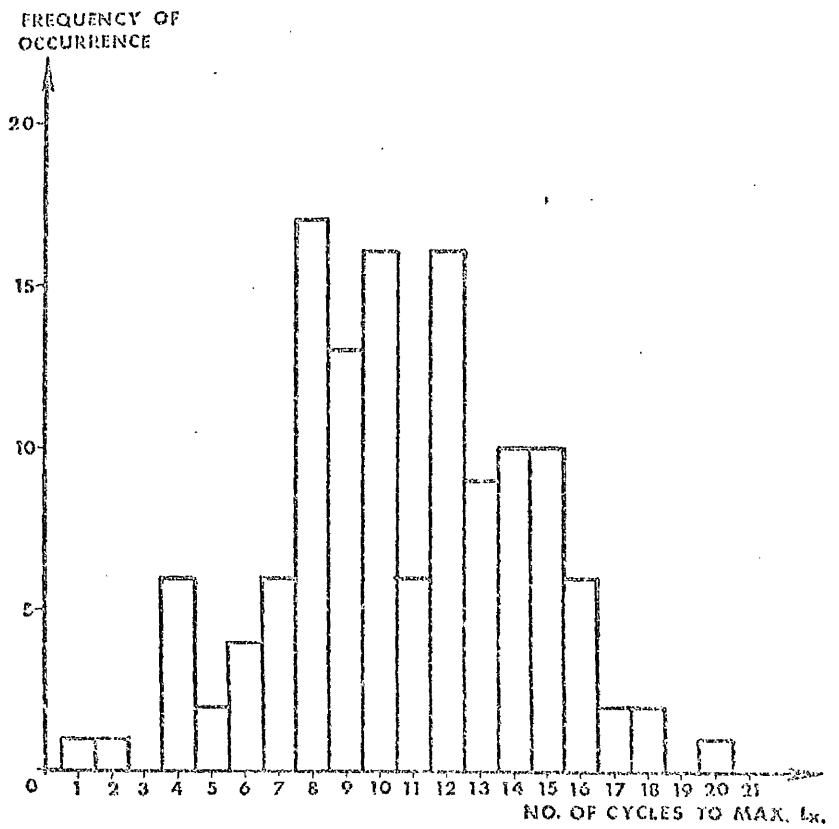


FIG.5.5 FREQUENCY HISTOGRAM OF DATA ON THE NUMBER OF CYCLES TAKEN TO REACH A STABLE VIBRATORY PATTERN AFTER THE ONSET OF PHONATION.

achieve a stable vibratory pattern. In the majority of cases, this was attained between 7 and 16 cycles of vocal fold vibration after the initiation of phonation. There was no significant difference in the frequency distributions obtained from the male and female subjects. The times taken to reach maximum Lx amplitude were also plotted as a frequency histogram (Figure 5.6), and the majority of data points were in the 30 - 90 msec. range. The phonations lying below 30 msec. were all produced by female subjects who had fairly high fundamental frequencies and had also reached the maximum Lx amplitude within relatively few cycles. No male subject achieved maximum Lx amplitude within 50 msec. of the onset of phonation, and no female subject required more than 80 msecs. The greater time taken by the male subjects to reach a steady vibratory pattern was significantly longer than that for females (using the Mann-Whitney U-test, $Z = 7.92$, $p < 0.00003$) and probably resulted from the greater mass, and thus inertia, of the male vocal folds.

In most cases the phonation was initiated at the desired fundamental frequency so that although the amplitude of the vibratory pattern required some time to stabilize, the period of the vibration remained constant. However, this was not exclusively the case, and several examples illustrating this and other types of vocal initiation are given in Figures 5.7 and 5.8.

In Figure 5.7A, the two initial cycles were followed by a long period of no vocal fold contact, after which the vibratory pattern was resumed, but at a low fundamental frequency. The fundamental frequency gradually increased until the final value was reached after about 80 msec. and 12 cycles from the onset of phonation. The examples illustrated in Figure 5.7B, C and D show similar phenomena.

Figure 5.8A shows the initiation of a phonation in which the Lx pattern was stable with respect to

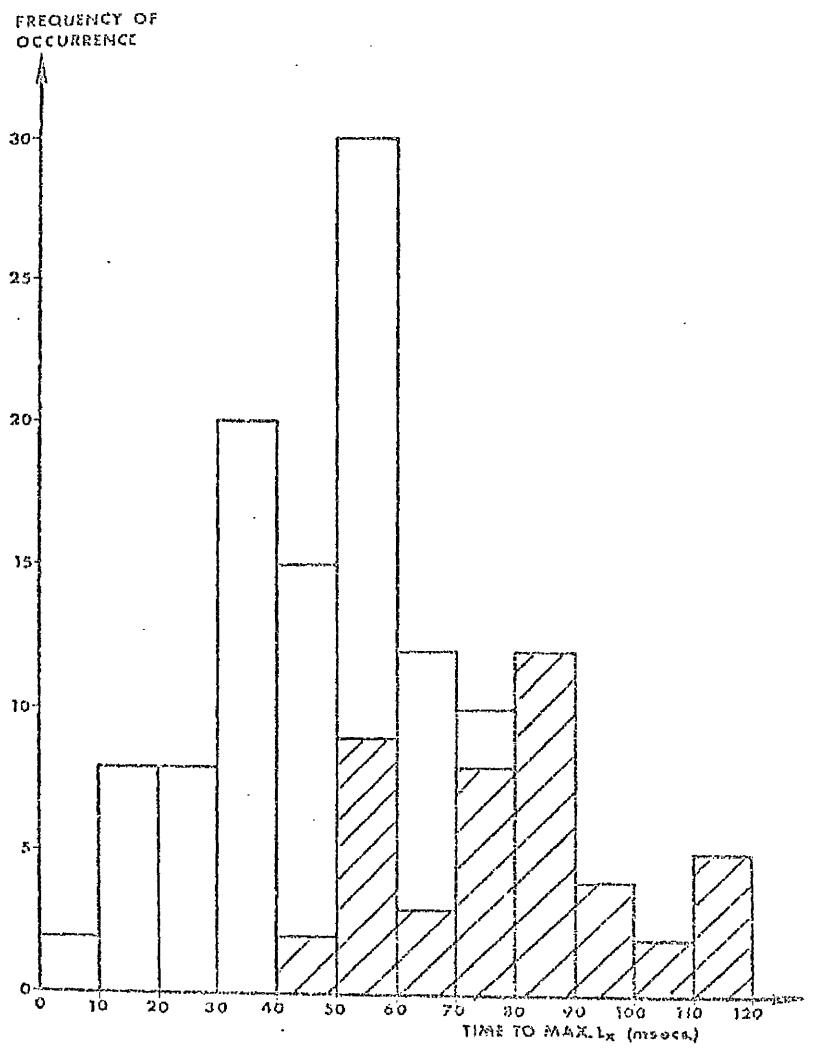


FIG.5.6 FREQUENCY HISTOGRAM OF THE TIMES TAKEN TO REACH A STABLE VIBRATORY PATTERN AFTER THE INITIATION OF PHONATION.

☐ FEMALES , ▨ MALES

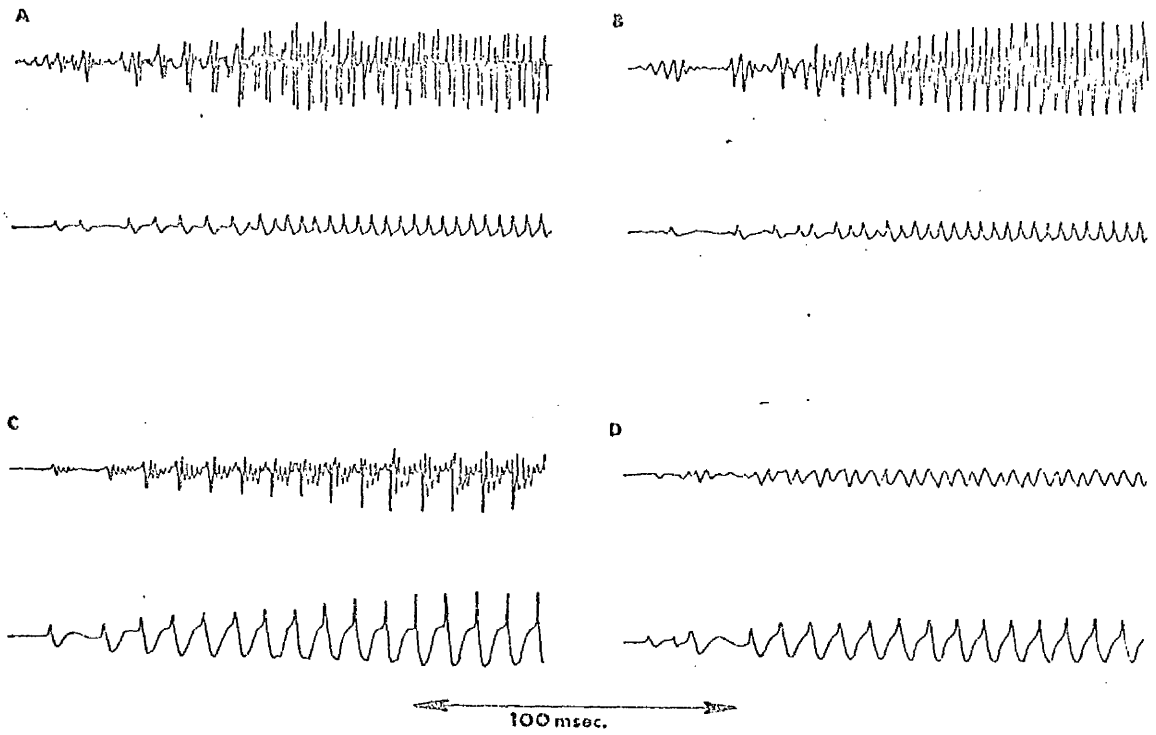


FIG.5.7 EXAMPLES ILLUSTRATING MICROPHONE OUTPUTS (UPPER TRACES) AND Lx OUTPUTS (LOWER TRACES) DURING VOCAL INITIATION.

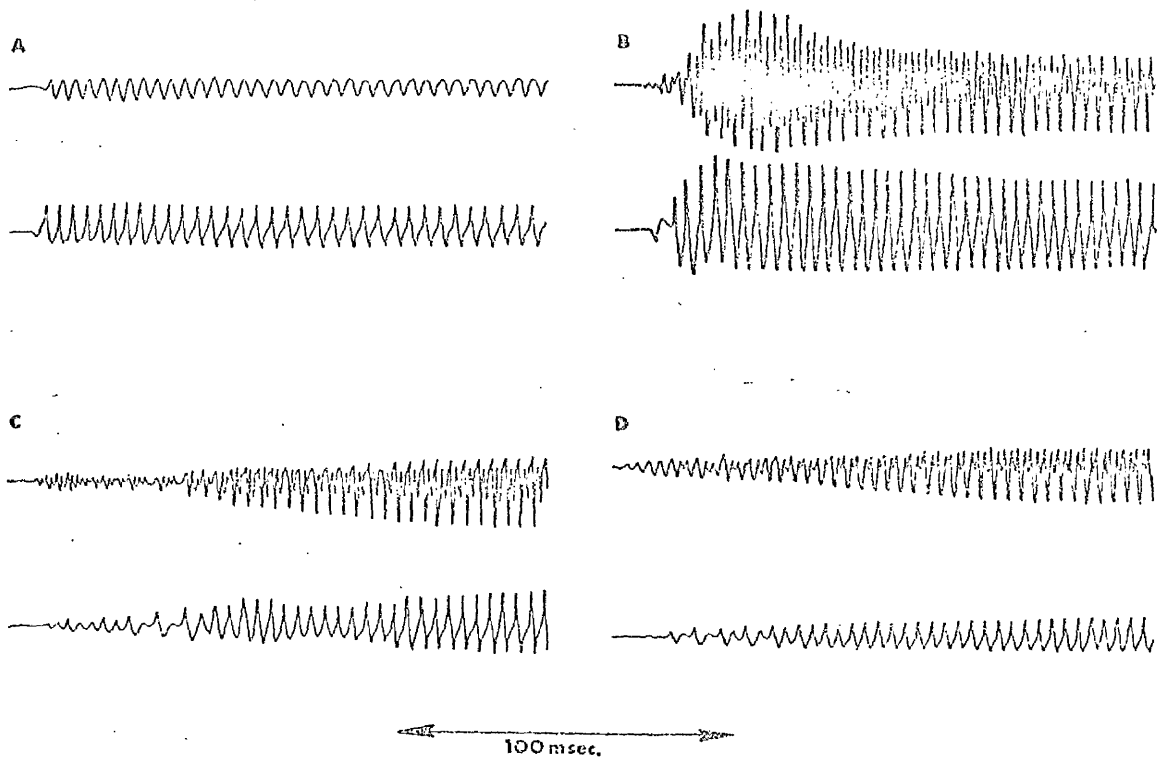


FIG. 5.8 EXAMPLES ILLUSTRATING MICROPHONE OUTPUTS (UPPER TRACES) AND Lx OUTPUTS (LOWER TRACES) DURING VOCAL INITIATION.

both fundamental frequency and amplitude within one cycle of its onset. Figure 5.8B illustrates the "hard attack" type of vocal initiation. A characteristic of this type of vocal initiation was the abrupt rise in the voice output to above the eventual steady level, and a similar appearance was observed for the Lx output. In this type of initiation the vocal folds were set into vibration by a very high subglottic pressure, and during the first few milliseconds, more than average air flow would be noted (Koike et al, 1967). Figure 5.8C and D show similar phenomena to those illustrated in Figure 5.7.

It was found that in almost all of the vocal initiations studied, the desired fundamental frequency was immediately attained. This indicated the efficiency of the control system of the laryngeal musculature, and the independence of the initial fundamental frequency (and therefore vocal fold adjustment) on auditory feedback (although this may have been relevant in some of the phonations illustrated). The time required to achieve a steady amplitude of vibration was a function of the mass of the vocal folds and the driving force (subglottic pressure).

5.5 FREQUENCY ANALYSIS OF Lx WAVEFORMS.

It was observed in CHAPTER 4 that many of the irregularities noted in dysphonic phonations had correlates in irregularities present in the frequency spectra. It was also observed, Figure 5.7, that irregularities in the vocal fold vibratory pattern as measured via the laryngograph, also corresponded to irregularities in the voice output. A series of experiments was undertaken to investigate the properties of the frequency spectra of the Lx output, and to see how far the parameters of the Lx waveform could be predicted from a knowledge of the spectrum.

Some preliminary Lx spectra were obtained using the sonagraph, and an example is shown in Figure 5.9A, where the x-axis was plotted as harmonic number to

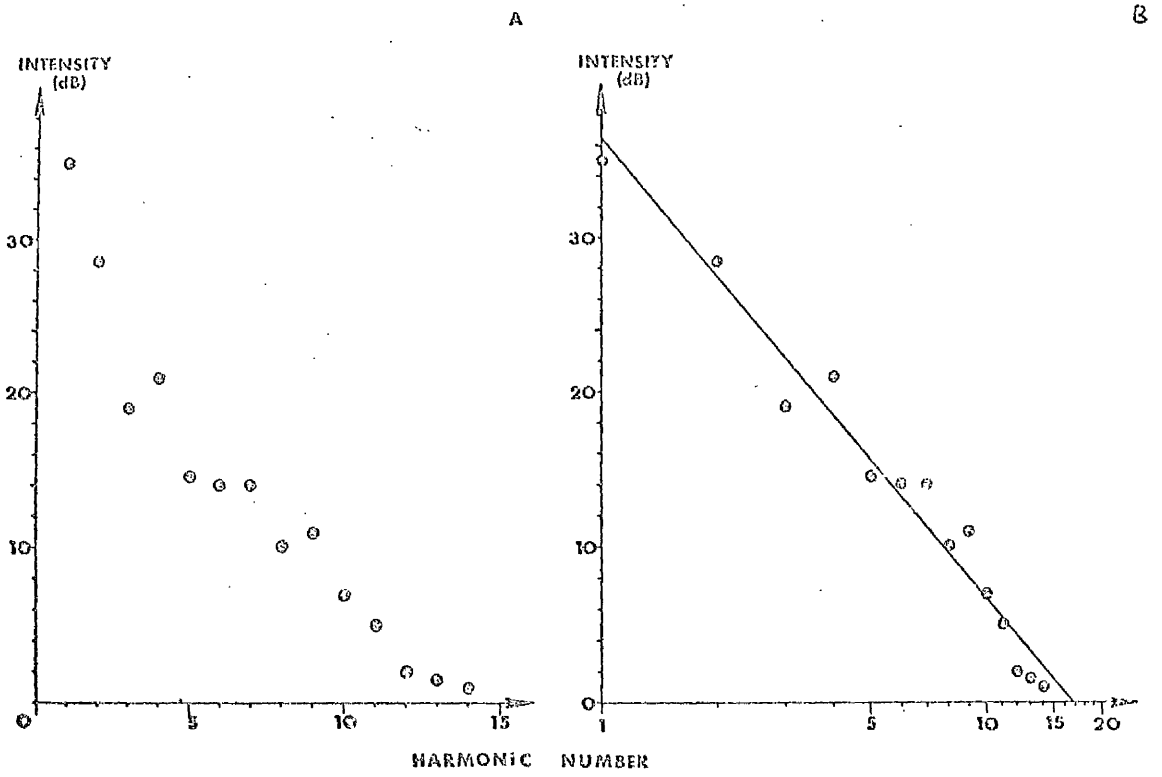


FIG.5.9 EXAMPLE OF THE GRAPHS OF INTENSITY v HARMONIC NUMBER FOR THE SPECTRUM OF AN Lx OUTPUT.
(a) LINEAR FREQUENCY SCALE
(b) LOGARITHMIC FREQUENCY SCALE

make the graphs independent of the absolute value of the fundamental frequency. It was evident that the intensities of the successive harmonic components followed approximately an exponential decrease, and so the spectrum was replotted with a logarithmic frequency scale, Figure 5.9B. The equation of the regression line was calculated, and the value of the correlation coefficient ($r = -0.982$, $p < 0.001$) indicated that the line was a good fit, and had the form:

$$I(f) = I_o + a.\ln(f) \quad , \quad (f=\text{frequency})$$

where I_o and a were the parameters of the regression line. The parameter I_o was only indicative of an overall amplification of the signal, and being dependent on the record and replay levels was of no use in the definition of the Lx characteristics. However, the spectral gradient parameter, $m = a.\ln 2$, was expressed in units independent of frequency, i.e. dB/octave, and could thus be used to define the variation with frequency, of the intensity of the spectral components. This type of analysis was carried out for all the Lx spectra investigated.

In a general way the spectral gradient gave an indication of the intensity of the higher number harmonic components present in the spectra, and thus could be related to the waveshape parameters of the Lx output. The properties of idealised Lx waves were investigated using a Fourier Series program written for a Wang 700 desk top computer (APPENDIX 2.2). The program required as input one cycle of the repetitive waveform, specified by up to 60 consecutive, equally spaced data points, and the Fourier coefficients were calculated. Examples of the type of idealised Lx waveshapes used are illustrated in Figure 5.10. These were approximately triangular in shape, with a linear rise representing the closing phase (defined by the closing time, CL.T), a rounded decreasing section representing the opening phase (defined by the opening time, O.T.), and a flat base for the open phase. The

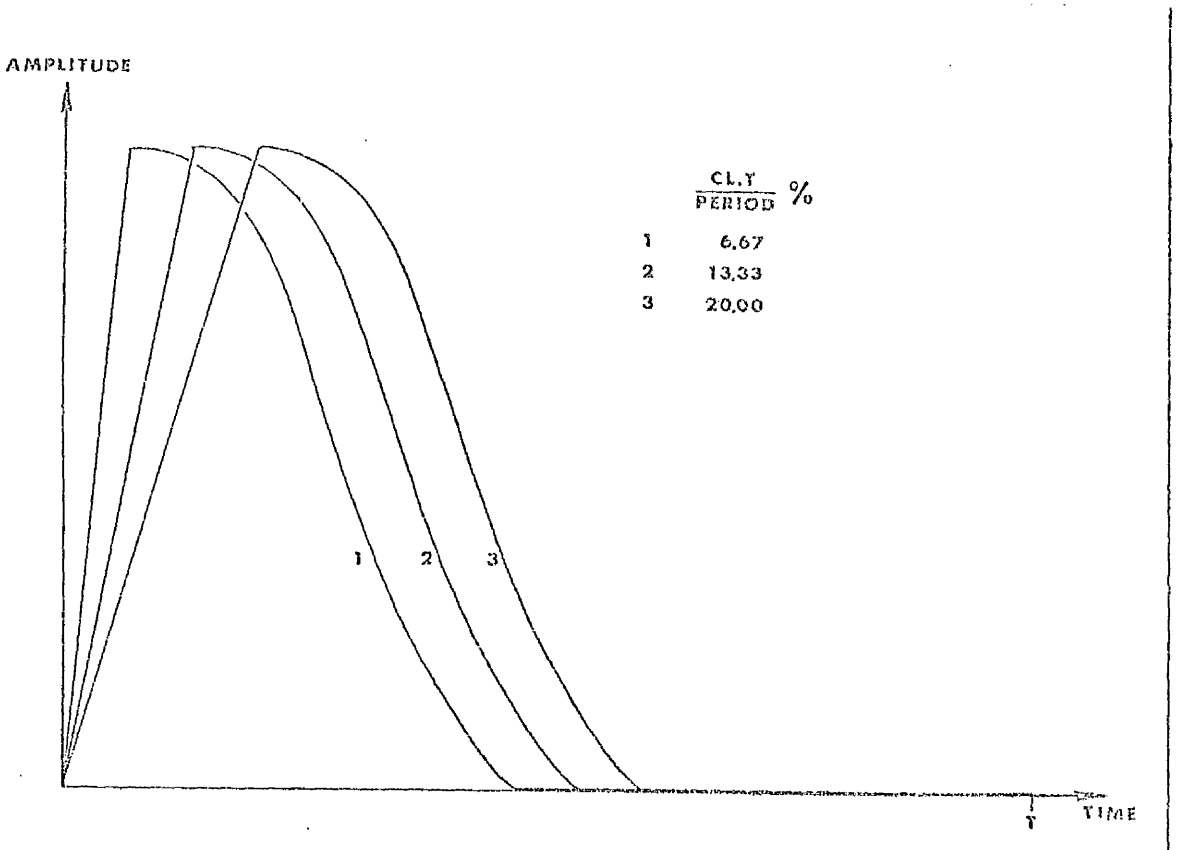


FIG.5.10 EXAMPLES OF THE IDEALISED Lx WAVEFORMS INVESTIGATED.

values of CL.T. and O.T. were chosen so as to adequately cover the ranges for these parameters found for normal subjects in SECTION 5.3. The data for 48 such waveforms were measured, the values of the first 15 harmonic components obtained using the Fourier program, and the spectral gradients calculated in dB/oct.

The spectra shown in Figure 5.11 resulted from waveshapes for which the CL.T. was maintained at 6.7% of the period while the value of O.T. was varied in steps from 26.7 - 50.0% of the period. Thus it appeared that such variations in O.T. had little effect on the resultant spectrum, and this was confirmed by plotting the spectral gradients of these test waveforms against O.T. for different values of CL.T., Figure 5.12. It was noted that for a particular value of CL.T. the spectral gradient remained fairly constant. This was true throughout most of the range, apart from where the CL.T. was close to the value of O.T., and O.T. had a value about 20% of the period.

The variation of spectral gradient with CL.T. can be seen from Figure 5.13. As the value of CL.T. was decreased, the spectral gradient became less negative, reflecting the increase in the intensity of the higher harmonic number components as the rise in the waveshape became sharper. However, the spread of the data points for a given value of CL.T. increased as the value of CL.T. increased. The explanation for this lay in the fact that the spectra of the waveforms with the higher values of CL.T. did not decrease monotonically, but had fine structure consisting of subsidiary maxima and minima. Figure 5.14 shows the spectra obtained for a series of CL.T. values, while the value of O.T. remained constant at 46.7% of the period. Fine structure was observed in all of the spectra with CL.T. greater than 6.7% of the period. A reciprocal relationship was noted between the value of CL.T. and the position of the

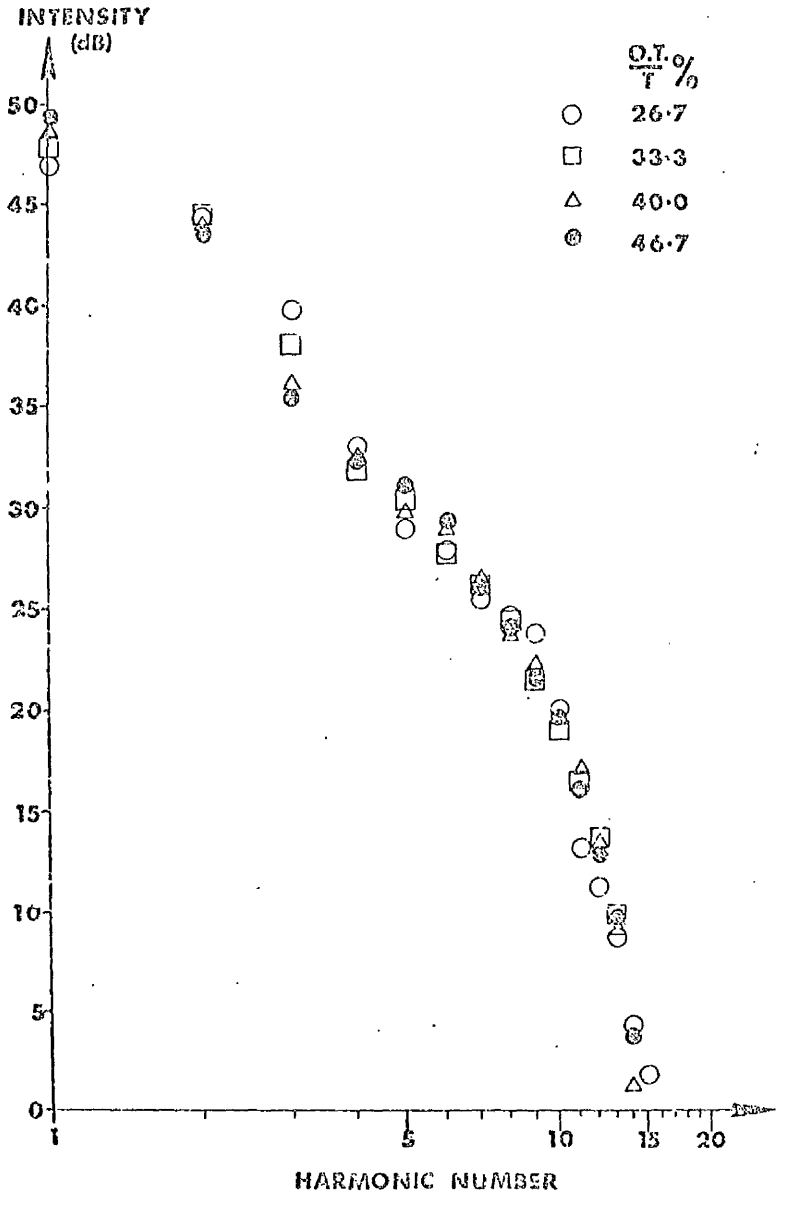


FIG.5.11 SPECTRA OF IDEALISED Lx WAVEFORMS WITH CL.T. = 6.7% OF PERIOD.

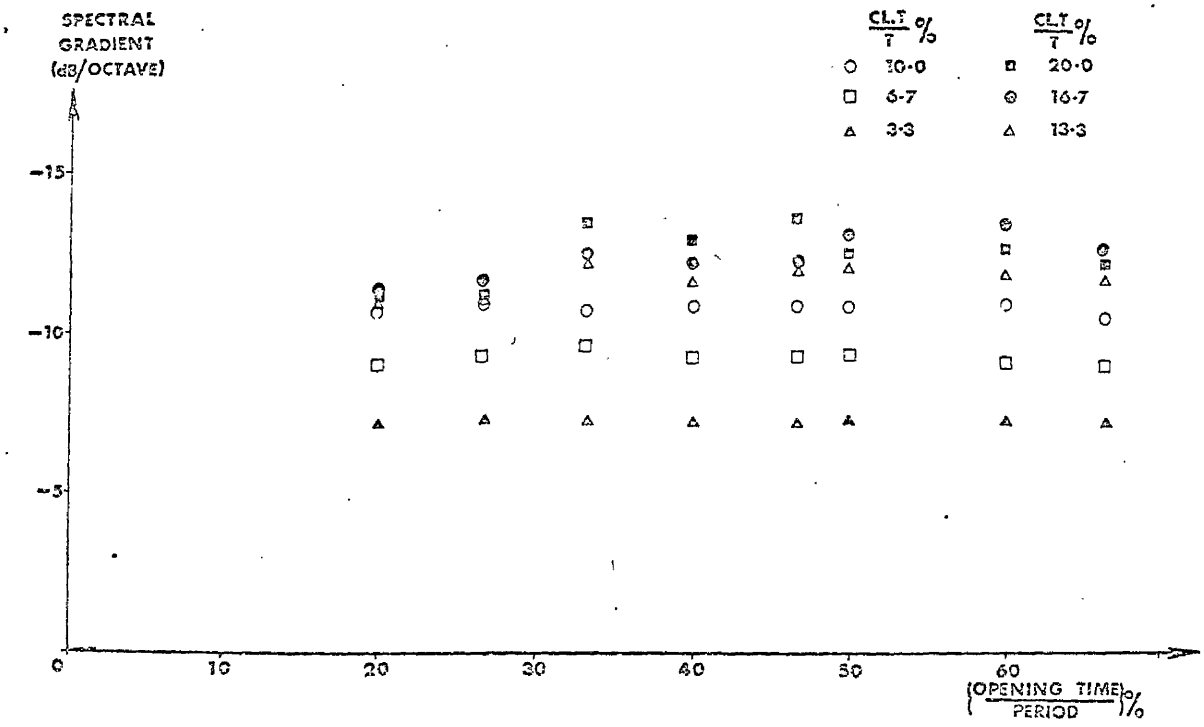


FIG.5.12 GRAPHS OF SPECTRAL GRADIENT v (O.T./T)% FOR IDEALISED Lx WAVEFORMS.

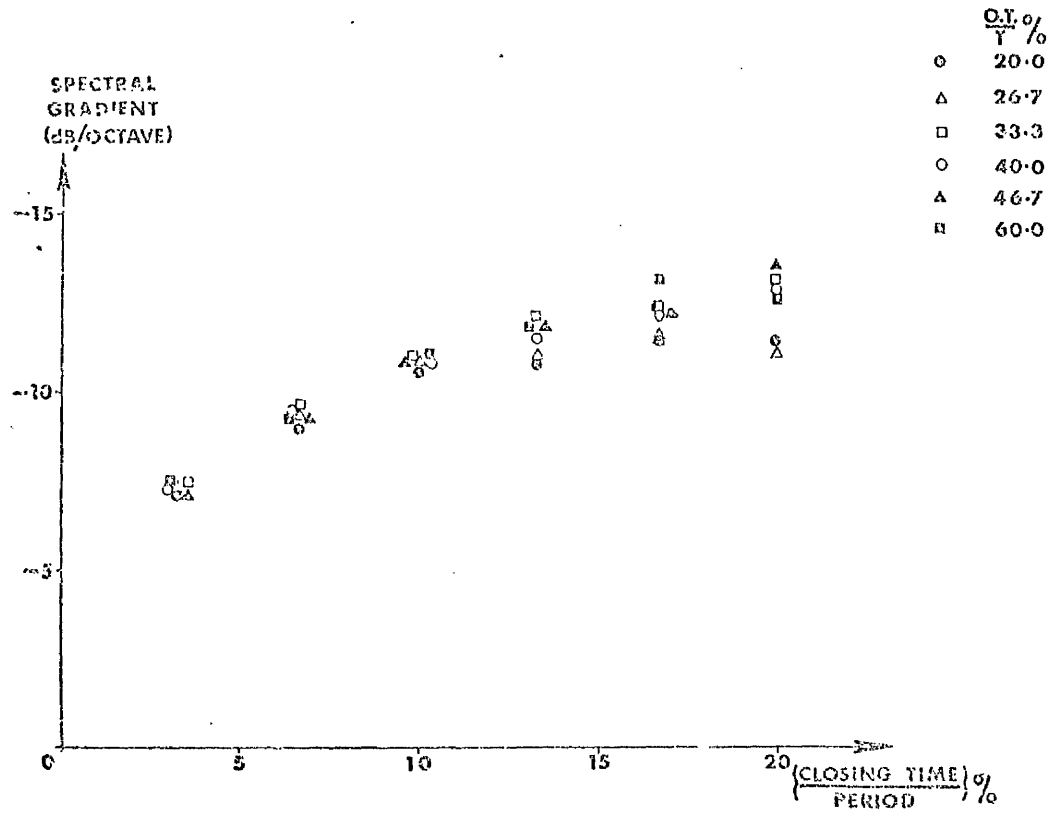


FIG.5.13 GRAPHS OF SPECTRAL GRADIENT v (CL.T./T)% FOR IDEALISED Lx WAVEFORMS.

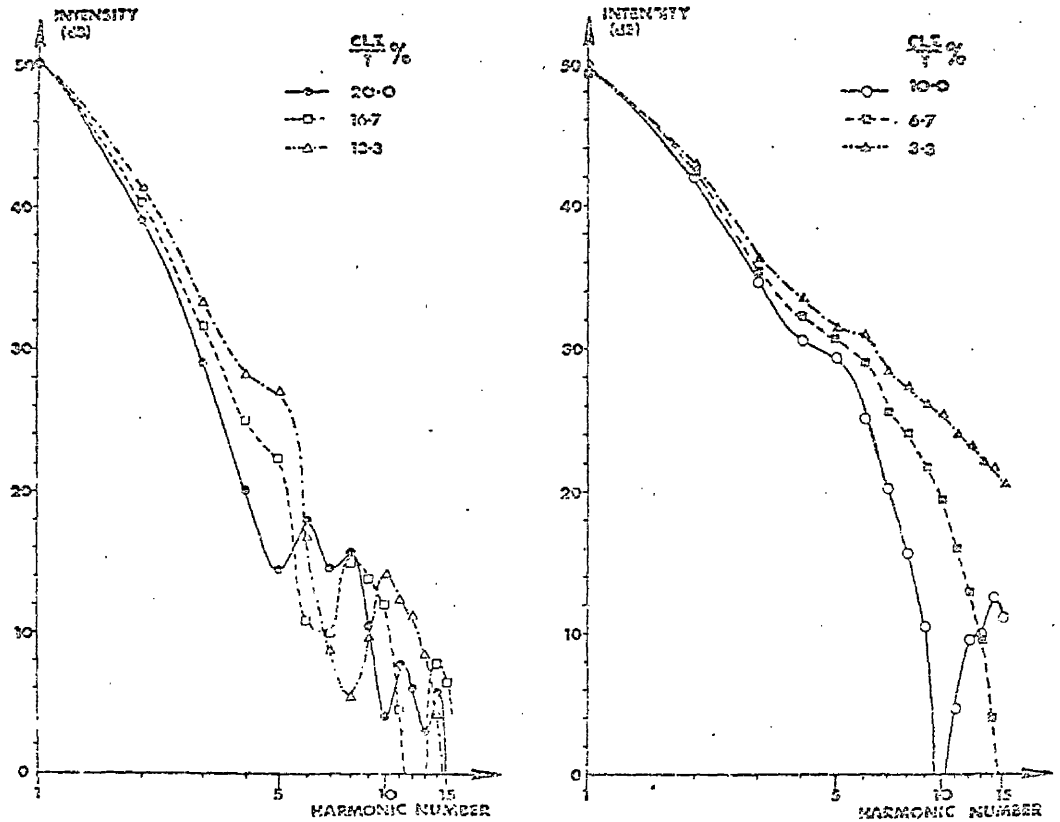


FIG.5.14 SPECTRA OF IDEALISED Lx WAVEFORMS WITH O.T. = 46.7% OF PERIOD

first subsidiary minimum, i.e. if N' = harmonic number of the first subsidiary minimum,

then $N' = \frac{T}{CL.T.}$, where T = period

e.g. if CL.T. = 20% of T, then $N' = 5$.

Such a relationship was not unexpected in view of the analysis of asymmetrical triangular waveshapes (APPENDIX 3). No fine structure was evident in Figure 5.11 since the value of CL.T. was 6.7% of the period, predicting the first subsidiary minimum at $N' = 15$. The results for the test waveforms were tabulated, TABLE 5.4, and a close correlation was evident between the predicted position of the first subsidiary minimum and its experimentally determined value, on the condition that O.T. was greater than about 25% of the period. This condition could be relaxed somewhat if CL.T. was a very small fraction of the period.

A series of the Lx outputs recorded previously (SECTION 5.3) were analysed in a similar fashion. Samples of these recordings were retained in a digital signal store and replayed at slow speed onto an x-t plotter (Bryans series 22000) to obtain large scale drawings of the Lx waveforms. The values of CL.T. and O.T. were measured, and the waveforms specified by up to 60 equally spaced data points. These were used as input data for the Fourier program, and the Fourier coefficients were calculated. The position of the first subsidiary minimum was noted in each case, and the results are given in TABLE 5.5. The values of CL.T. varied from 7.4 - 19.4% of the period, and a strong correlation between $\left(\frac{CL.T.}{T}\right)\%$ and $\frac{100}{N'}$ was evident ($r = 0.97$, $p < 0.001$). Thus, there was good agreement in the values of CL.T. as predicted from the frequency spectra and those measured directly.

The simplest method of obtaining the Lx spectra was to use the sonagraph, and a series of such spectra was obtained from the previous recordings, with the

$\frac{CL.T_{\%}}{T}$	$\frac{O.T_{\%}}{T}$	$\frac{T}{CL.T}$	HAR-MONIC NO.OF 1ST MINIMUM	$\frac{CL.T_{\%}}{T}$	$\frac{O.T_{\%}}{T}$	$\frac{T}{CL.T}$	HAR-MONIC NO.OF 1ST MINIMUM
20	20	5	7	10	20	10	6
20	26.7	5	6	10	26.7	10	10
20	33.3	5	5	10	33.3	10	10
20	40	5	5	10	40	10	10
20	46.7	5	5	10	46.7	10	10
20	50	5	5	10	50	10	10
20	60	5	5	10	60	10	10
20	66.7	5	5	10	66.7	10	10
16.7	20	6	6	6.7	20	15	-
16.7	26.7	6	4	6.7	26.7	15	15
16.7	33.3	6	6	6.7	33.3	15	15
16.7	40	6	6	6.7	40	15	15
16.7	46.7	6	7	6.7	46.7	15	15
16.7	50	6	6	6.7	50	15	15
16.7	60	6	6	6.7	60	15	15
16.7	66.7	6	6	6.7	66.7	15	15
13.3	20	7.5	5	3.33	20	30	-
13.3	26.7	7.5	8	3.33	26.7	30	-
13.3	33.3	7.5	8	3.33	33.3	30	-
13.3	40	7.5	8	3.33	40	30	-
13.3	46.7	7.5	8	3.33	46.7	30	-
13.3	50	7.5	8	3.33	50	30	-
13.3	60	7.5	8	3.33	60	30	-
13.3	66.7	7.5	8	3.33	66.7	30	-

TABLE 5.4

DATA FOR THE IDEALISED Lx WAVEFORMS

$\frac{CL.T_{\%}}{T}$	$\frac{OT.T._{\%}}{T}$	$\frac{H.NO.OF}{1ST\ MINIMUM}$	$\frac{100.}{H.NO.}$
13.0	56.1	7	14.3
12.7	45.5	7	14.3
15.9	49.3	6	16.7
19.4	40.9	5	20
16.0	51.9	7	14.3
7.6	36.9	13	7.7
8.0	40.3	12	8.3
7.1	47.8	13	7.7
13.4	52.1	7	14.3
7.4	49.6	13	7.7
9.7	48.4	11	9.1
11.0	48.5	10	10.0
8.9	52.8	11	9.1
10.8	54.7	9	11.1
11.5	55.8	9	11.1

TABLE.5.5

DATA FOR SAMPLES OF Lx WAVEFORMS SHOWING CLOSE CORRELATION BETWEEN $\frac{CL.T}{T}$ AND $\frac{1}{N'}$ $(r = 0.97)$

positions of the first subsidiary minima being measured. The spectral gradients were calculated, and Figure 5.15 shows the graphs of spectral gradient v $\frac{(CL.T.)}{(T)}\%$ and spectral gradient v harmonic number of first subsidiary minimum (or spectral gradient v $\frac{100}{N'}$). The dashed lines indicate the limits containing the test wave data points (Figure 5.13). The results showed good agreement between the experimental data and those obtained for test waveforms, whether the CL.T. was measured directly, or deduced from a knowledge of the fine structure of the frequency spectrum.

Since the fundamental frequency (and hence the period) could be calculated from the sonagraph outputs, it was possible to calculate the absolute value of the closing time of the vocal fold vibration, since

$$CL.T. = \frac{T}{N'} = \frac{1}{F_o N'}$$

where F_o = fundamental frequency.

The main uncertainty in the value of the closing time as calculated from the fine structure of the Lx spectrum arose from the fact that the frequency of the first subsidiary minimum was restricted to integral multiples of the fundamental frequency. It was not justifiable to interpolate between the harmonic components since the precise shape of the envelope curve was not known, and the maxima and minima were not symmetrical. In the cases where no fine structure was evident in the frequency spectra, it was possible to obtain limits for the closing time from a knowledge of the spectral gradient. The value of the closing time was thought to be of clinical significance as a measure of the tension in the larynx, e.g. if the tension produced in the vocal folds was too great, then a greater force, i.e. subglottic pressure would be required to maintain vibrations at a suitable amplitude. This could increase the air flow rate during the period of glottal opening, thus producing a greater Bernoulli force, and a more rapid vocal

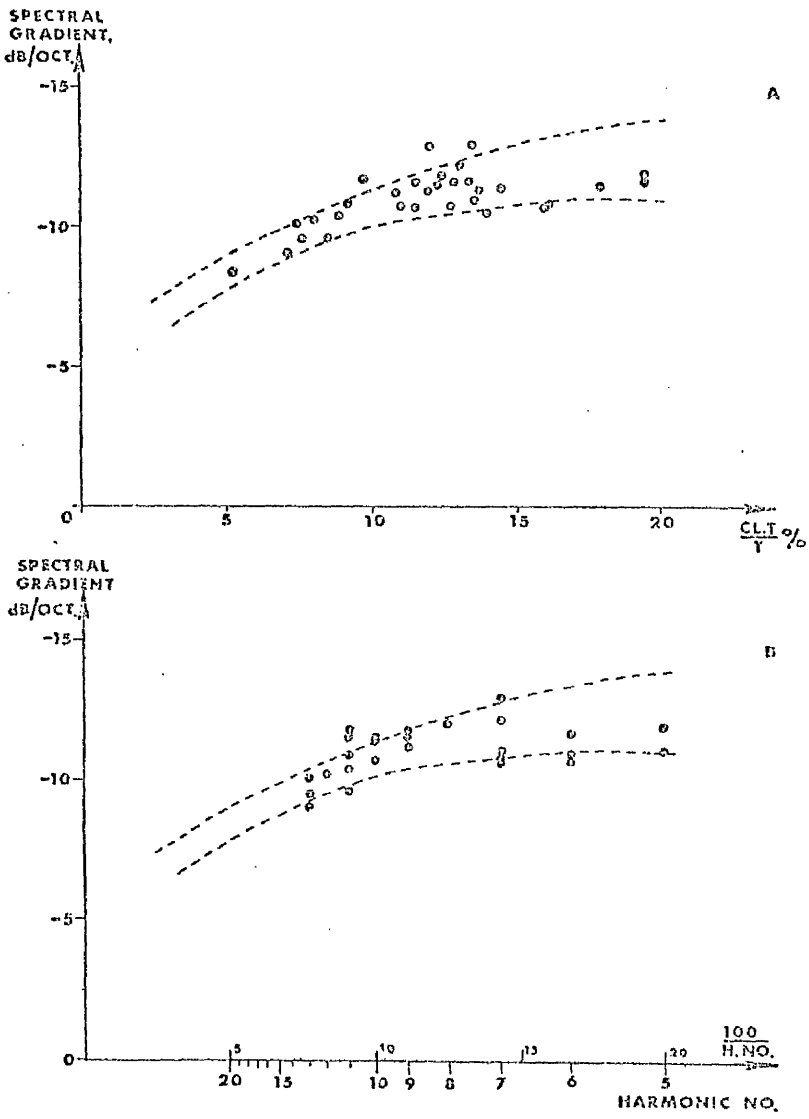


FIG.5.15 GRAPHS OF (A) SPECTRAL GRADIENT v $\frac{CL.T}{T}$ % AND

(B) SPECTRAL GRADIENT v HARMONIC NUMBER OF FIRST SUBSIDIARY MINIMA.

fold closure. This would result in a more violent vocal fold contact and the possible development of such pathological conditions as vocal nodes or contact ulcers etc. However, the further investigation of this aspect was beyond the scope of this thesis.

5.6 ABNORMAL PHENOMENA IN Lx SPECTRA.

In CHAPTER 4 (SECTION 4.4) the presence of abnormal frequency components in the voice spectra of phonations produced by dysphonic subjects was discussed. In the present section this theme will be developed by investigating the Lx spectra of some dysphonic subjects, and noting how far such abnormalities, especially the production of subharmonic frequency components, can be attributed to the vibratory pattern of the vocal folds, as measured by the frequency spectra of the Lx outputs. A two channel sonagraph was available for this part of the study, thus allowing the analysis of simultaneous portions of the voice and Lx outputs.

Abnormalities in the basic frequency pattern, such as changes in the fundamental frequency of phonation, as expected, could be attributed directly to corresponding changes in the frequency of the vocal fold vibration. This is demonstrated in Figure 5.16 in which the laryngograph output (upper diagrams) and voice output (lower diagrams) produced by a dysphonic subject were analysed at the same point in the phonation. The abrupt change in the fundamental frequency (pitch break) was evident in both outputs. However, it was important to note the increased presence of random noise components at corresponding points in both outputs, indicating that in this case, at least some of the noise components originated from an irregular pattern of vocal fold vibration. This need not always be the case, since noise components could also originate from a fricative source higher up in the vocal tract, and these would, therefore, not appear in the corresponding Lx spectrum.

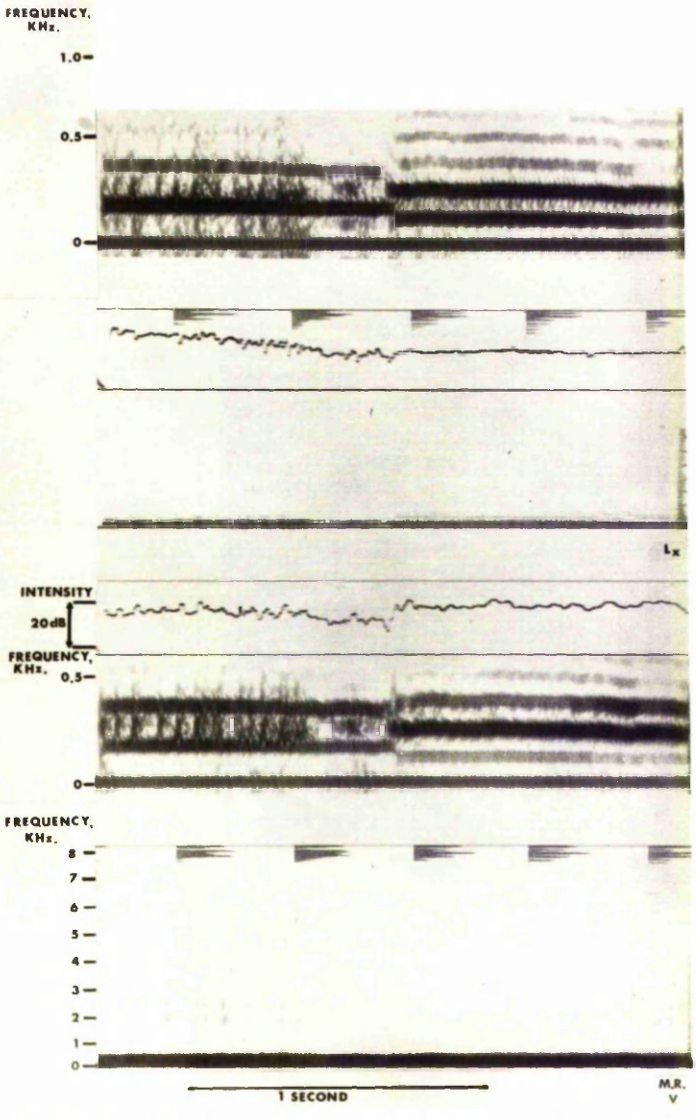


FIG.5.16 EXAMPLES OF SIMULTANEOUS SONAGRAMS OF VOICE AND Lx OUTPUTS DEMONSTRATING PITCH BREAK.

It was demonstrated in CHAPTER 4 that the presence of subharmonic components in the voice spectra of dysphonic subjects was not uncommon. Seymour (1972) noted extra frequency components in the spectra produced by singers, at frequencies around 5.7KHz. for males and 6.8KHz. for females. Some of these components fitted the $\frac{1}{2}$Fo series (Fo = fundamental frequency) and were attributed to the vocal fold vibration, while other extra components in the same frequency range did not fit the $\frac{1}{2}$Fo series and were attributed to selective amplification of noise. In the present study, conclusive evidence was obtained which demonstrated that the origin of subharmonic components was the vocal fold vibration, but that such components would be expected at much lower frequencies than previously reported.

Figures 5.17 and 5.18 show examples of simultaneous voice and Lx spectra produced during the intermittent production of subharmonic frequency components. In each example there were periods of phonation with only normal frequency components present, and in every case, the initiation of subharmonic production in the voice spectrum was coincident with a similar occurrence in the spectrum of the Lx output. Figure 5.19 gives another example of the same phenomenon, with even periods of production of the $\frac{1}{4}$Fo series at point X. It was also noted from the section displays that subharmonic production did not alter the intensities of components of the F_o series, but only involved the addition of components between the original ones. In all cases in which subharmonic components were found in the voice spectra, and Lx outputs were also available, subharmonic components were found at the corresponding points in the Lx spectra.

5.7 MATHEMATICAL DESCRIPTION OF SUBHARMONIC PRODUCTION.

The mathematical description of the changes in the Lx waveform responsible for subharmonic production

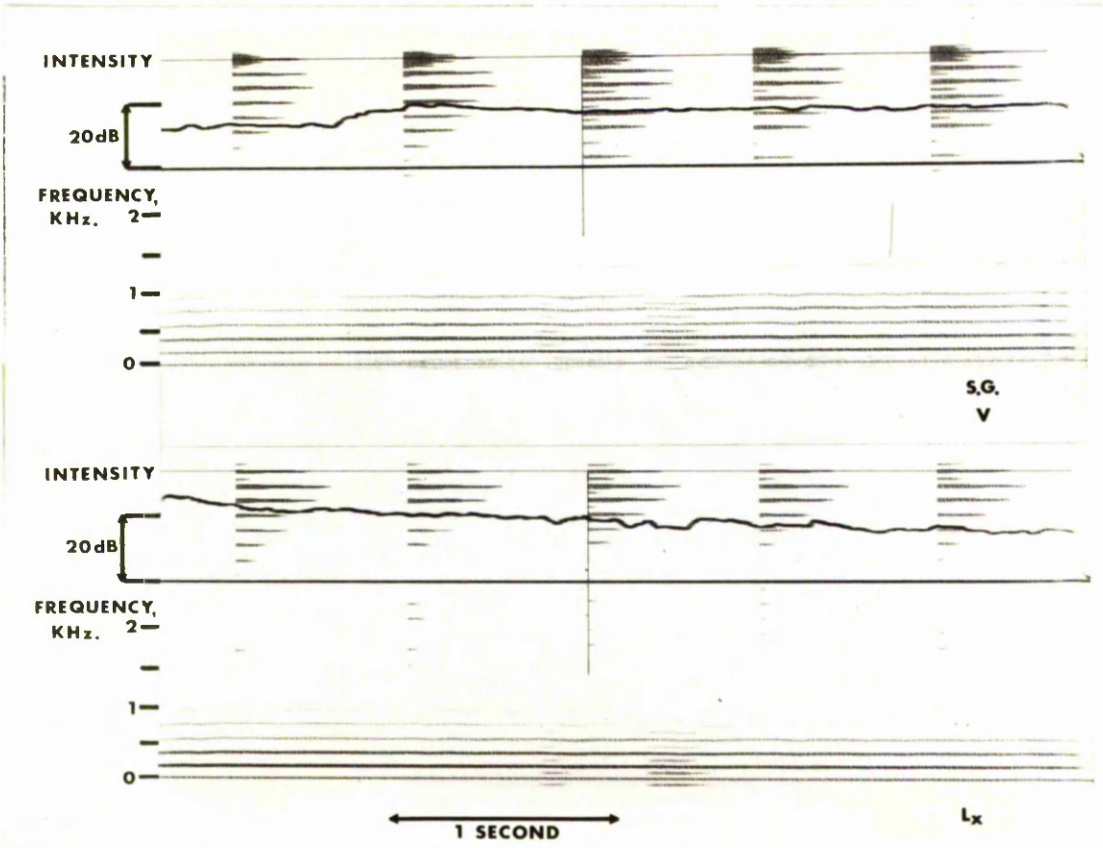


FIG.5.17 SONAGRAMS OF VOICE AND Lx OUTPUTS SHOWING INTERMITTENT SUBHARMONIC PRODUCTION.

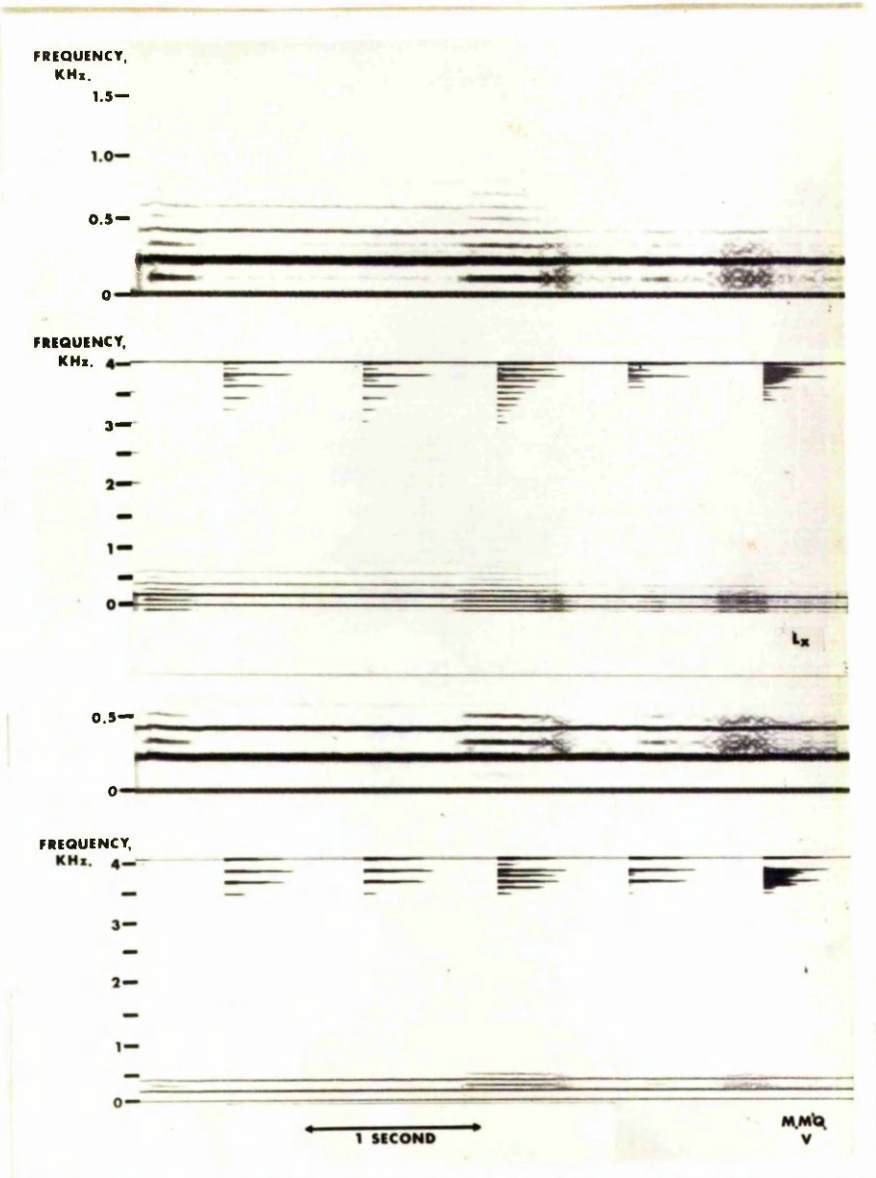


FIG.5.18 SONAGRAMS OF VOICE AND Lx OUTPUTS SHOWING INTERMITTENT SUBHARMONIC PRODUCTION

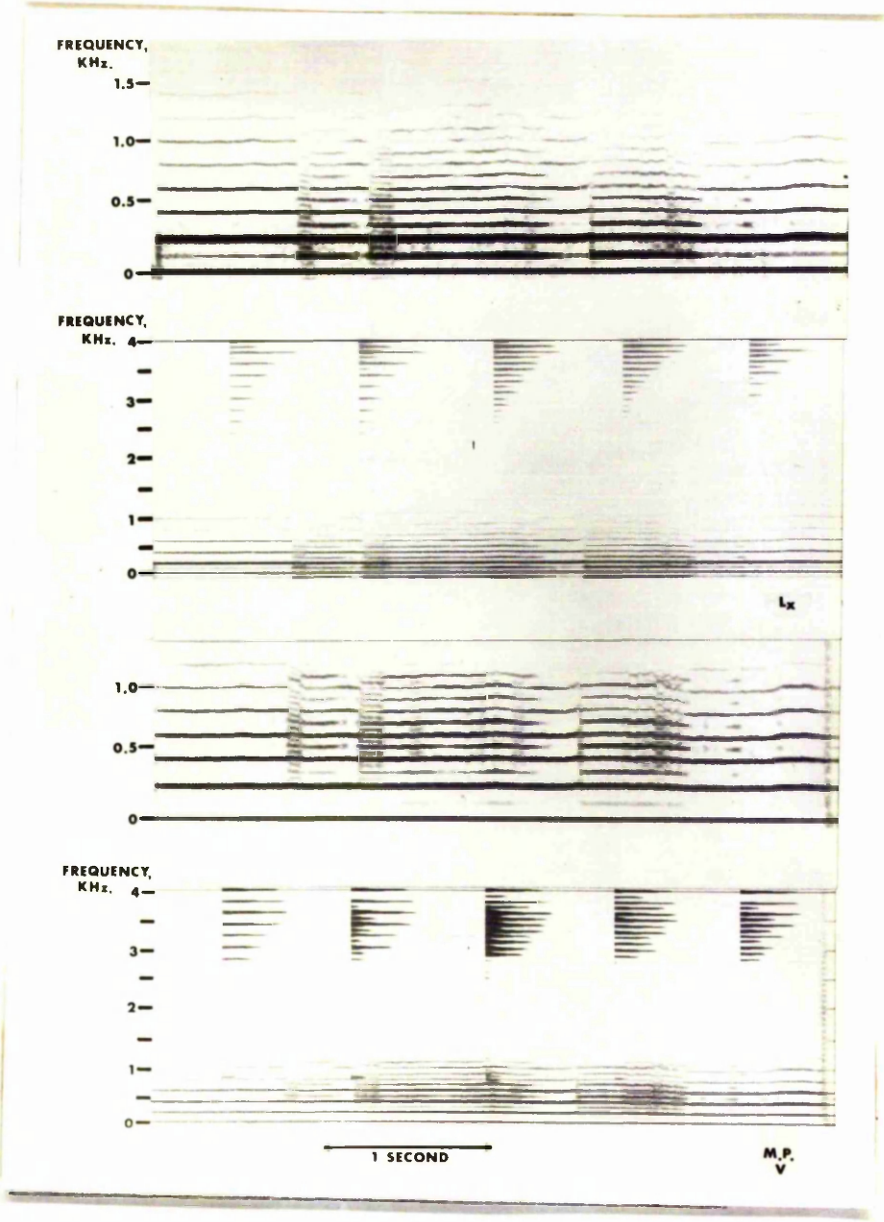


FIG.5.19 SONAGRAMS OF VOICE AND Lx OUTPUTS SHOWING INTERMITTENT SUBHARMONIC PRODUCTION.

and their experimental verification are now to be discussed.

Assume that the vibratory pattern of the vocal folds could be described by a function Q(t), which was periodic with a period of $\frac{T}{2}$ i.e. Q(t) had a fundamental frequency $\frac{2}{T} = 2f$. Let Q(t) satisfy the necessary continuity conditions, so that it was valid to describe Q(t) by a Fourier Series :

$$\text{i.e. } Q(t) = \frac{1}{2} a_o + \sum_{n=1}^{\infty} (a_n \cos 4\pi nft + b_n \sin 4\pi nft) \qquad (5.1)$$

i.e. Q(t) could be described as the summation of a series having components only at the frequencies 2f, 4f, 6f,... .

Let Q(t) now be amplitude modulated at a frequency f, by a function of the form :

$$(1 - p \sin \frac{2\pi t}{T}) = (1 - p \sin 2\pi ft) \qquad 0 \leq p \leq 1$$

and let the resultant function be M(t), where

$$M(t) = Q(t) \cdot (1 - p \sin 2\pi ft) \quad .$$

M(t) would also satisfy the required conditions for Fourier analysis to be valid, and could be described by a series such that :

$$M(t) = \frac{1}{2} c_o + \sum_{m=1}^{\infty} (c_m \cos 2\pi fmt + d_m \sin 2\pi fmt) \qquad (5.2)$$

Thus M(t) could be described as the summation of a series having components at the frequencies f, 2f, 3f, 4f, 5f, .

The components c_m and d_m were calculated in terms of a_n, b_n and p (SECTION 5.9) and the results are summarized in TABLE 5.6.

The components at the frequencies corresponding to the initial function, Q(t) (i.e. components at 2f, 4f, 6f,....) remained unaltered by the modulation function, while the components corresponding to the $\frac{1}{2}F_o$ series (i.e. at frequencies f, 3f, 5f,.....) were highly dependent on the values of the adjacent

FRE-QUENCY	COSINE COMPO-NENT OF Q(t)	SINE COMPO-NENT OF Q(t)	COSINE COMPO-NENT OF M(t)	SINE COMPO-NENT OF M(t)
f	-	-	$-\frac{1}{2}pb_1$	$\frac{1}{2}p(a_1-a_o)$
2f	a_1	b_1	a_1	b_1
3f	-	-	$\frac{1}{2}p(b_1-b_2)$	$\frac{1}{2}p(a_2-a_1)$
4f	a_2	b_2	a_2	b_2
5f	-	-	$\frac{1}{2}p(b_2-b_3)$	$\frac{1}{2}p(a_3-a_2)$
6f	a_3	b_3	a_3	b_3
7f	-	-	$\frac{1}{2}p(b_3-b_4)$	$\frac{1}{2}p(a_4-a_3)$
8f	a_4	b_4	a_4	b_4
'	'	'	'	'
'	'	'	'	'
'	'	'	'	'
'	'	'	'	'

TABLE 5.6

SINE AND COSINE COEFFICIENT FOR THE FUNCTIONS Q(t) AND M(t)

components of the F_o series and also on the depth of modulation, p. The choice of the modulation function was not critical so long as it was periodic with period T.

The effect of such a modulation was to alter the amplitude of every second cycle of the original waveform, thus introducing a periodicity with period double that of the original one, i.e. introducing frequency components of the $\frac{1}{2}F_o$ series.

It was shown that the foregoing analysis was correct during subharmonic production, by obtaining photographic records of the voice and Lx outputs via the U - V recording oscilloscope, and Figure 5.20 shows one example. In Figure 5.20A the Lx and microphone outputs were steady until subharmonic production began at the point X. The double periodicity is evident in both the Lx and microphone outputs and would be evident in frequency spectra as components of the $\frac{1}{2}F_o$ series. Figure 5.20B shows the Lx waveform around the point X using an expanded time scale, and the modulation was evident. The interpretation of such an Lx waveform indicated that the vocal folds were making normal contact only every second cycle. The cause was probably an asymmetry in the adjustment of the folds, perhaps leading to different vibrational frequencies.

Figure 5.21 shows another example of the modulation of the vocal fold vibratory pattern. Initially the Lx waveform was consistent with the production of the $\frac{1}{2}F_o$ series, but at the point Y, the pattern degenerated, and full vocal fold contact occurred at every 3rd, 4th or 5th cycles, consistent with the production of components of the higher subharmonic series. These components would not be resolved by the 45Hz bandwidth filter of the sonagraph.

Another example of an Lx output consistent with subharmonic production of the higher subharmonic series is given in Figure 5.22. This subject was suffering from unilateral vocal fold paralysis, with

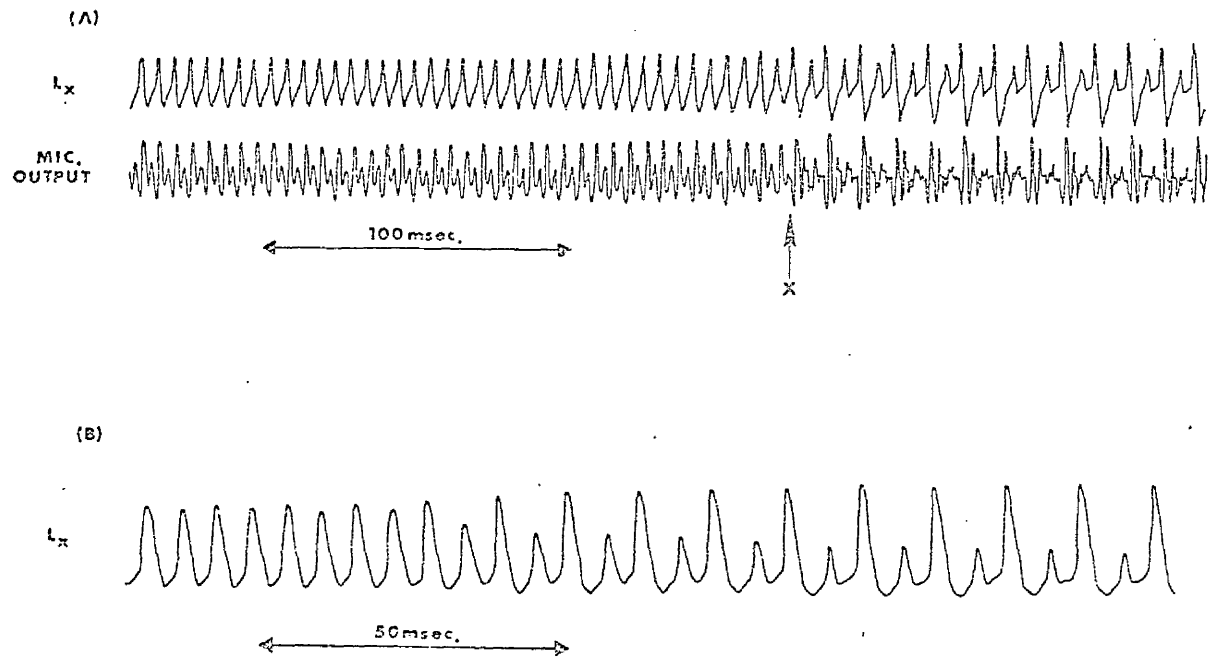


FIG.5.20 VOICE AND Lx WAVEFORM DURING SUBHARMONIC PRODUCTION.

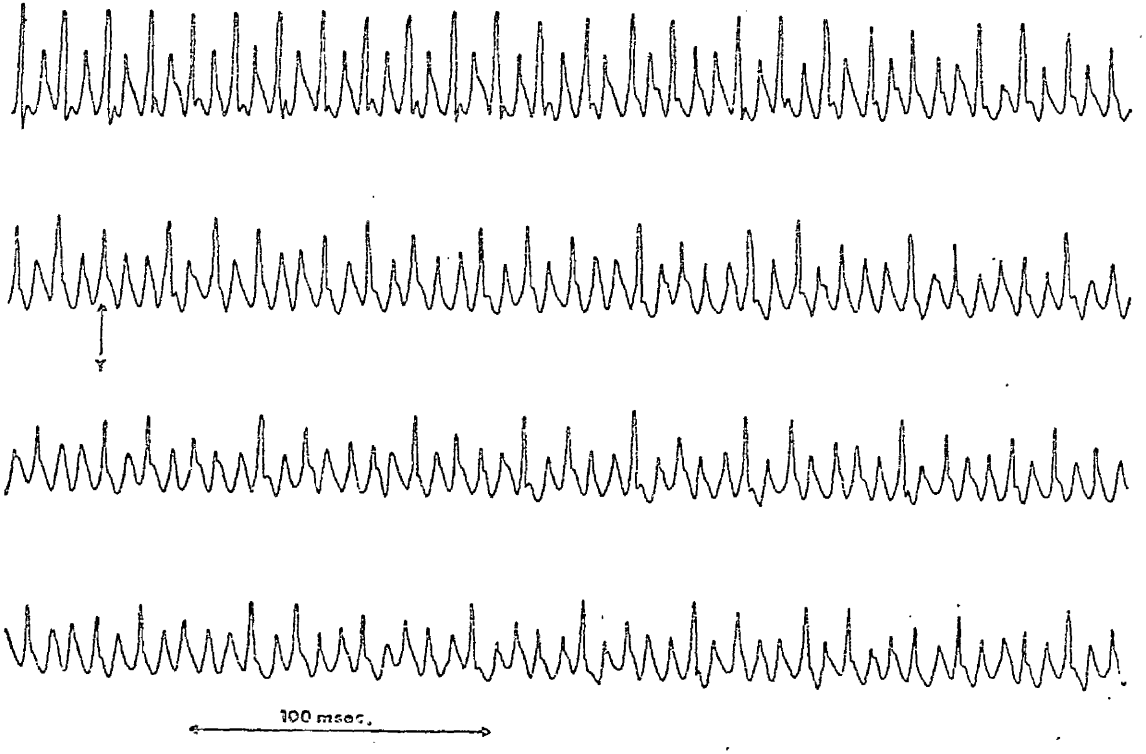


FIG.5.21 CONSECUTIVE PORTIONS OF Lx OUTPUT SHOWING PRODUCTION OF SUBHARMONIC COMPONENTS OF $\frac{1}{2}$Fo AND HIGHER SERIES.

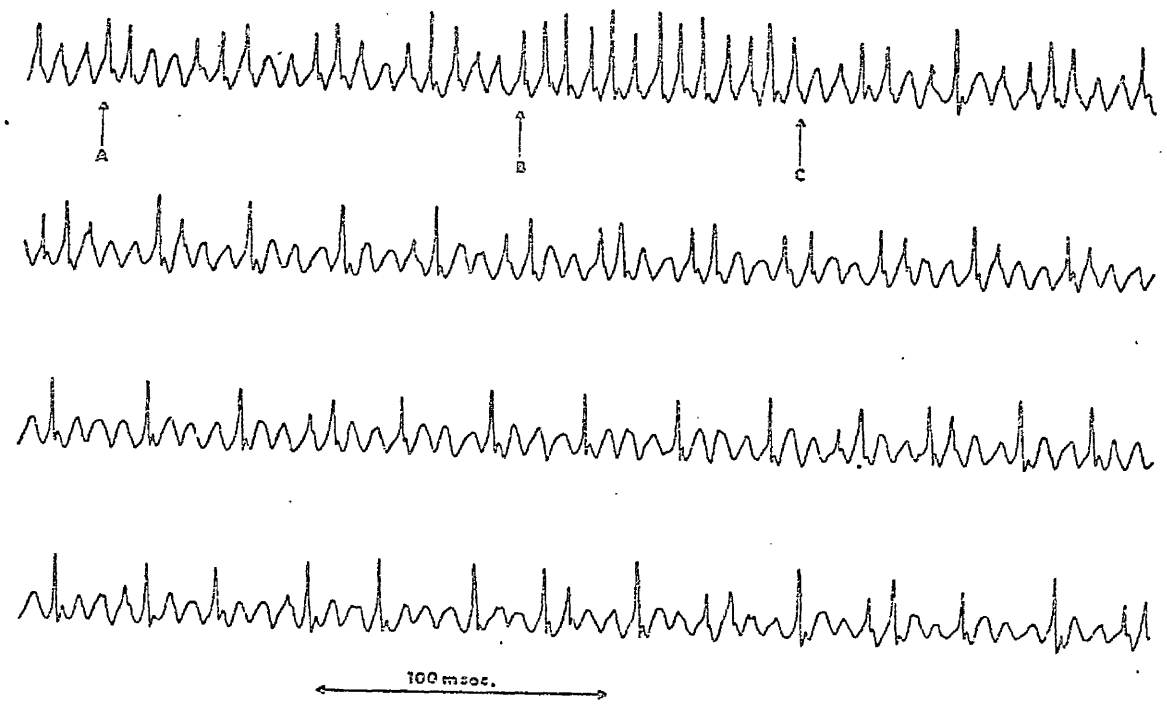


FIG.5.22 CONSECUTIVE PORTIONS OF Lx OUTPUT SHOWING PRODUCTION OF COMPONENTS OF THE HIGHER SUBHARMONIC SERIES.

the paralysed fold being in an almost medial position. Between points A and B, contact was only intermittent, between B and C it was slightly more regular and occurred at almost every cycle, but after C, the contact occurred only every 4th or 5th cycles. In this case, because of the effective de-innervation of the paralysed fold, it was feasible that the two vocal folds had different natural frequencies of vibration, and that this could have produced a beat type phenomenon in the area of contact, and thus in the Lx output.

This could be explained in terms of a simple model of the vocal folds.

5.8 MODEL OF VOCAL FOLDS

The complexity of the structure of the vocal folds was discussed in CHAPTER 1 (SECTION 1.4), and this leads to the folds having an extremely complicated vibratory pattern, with movement in both the horizontal and vertical planes. Thus models used to describe the motion of the folds realistically are necessarily complex, and involve the use of powerful computing facilities (Titze 1973; 1974; Flanagan 1969). However, even simple models can help in the understanding of specific phenomena, and one such will be discussed.

Consider the larynx to have a rectangular cross-section, and each vocal fold to be cuboid in shape, as shown in Figure 5.23A. Let the larynx have width 2C, and when the folds are in the "phonation ready" position, just prior to the onset of phonation, let the glottis have width $2a_o$. (Figure 5.23B). Let the folds be identical, with constant volume V, and vibrate sinusoidally about the "phonation ready" position with amplitude $\underline{a}$, and angular frequency w, while still maintaining a rectangular cross-section. If the origin of the vocal fold vibratory pattern is taken as the "phonation ready" position of one fold, then the equations of motion of the folds are :

$$y_1 = - a_1 \sin wt$$

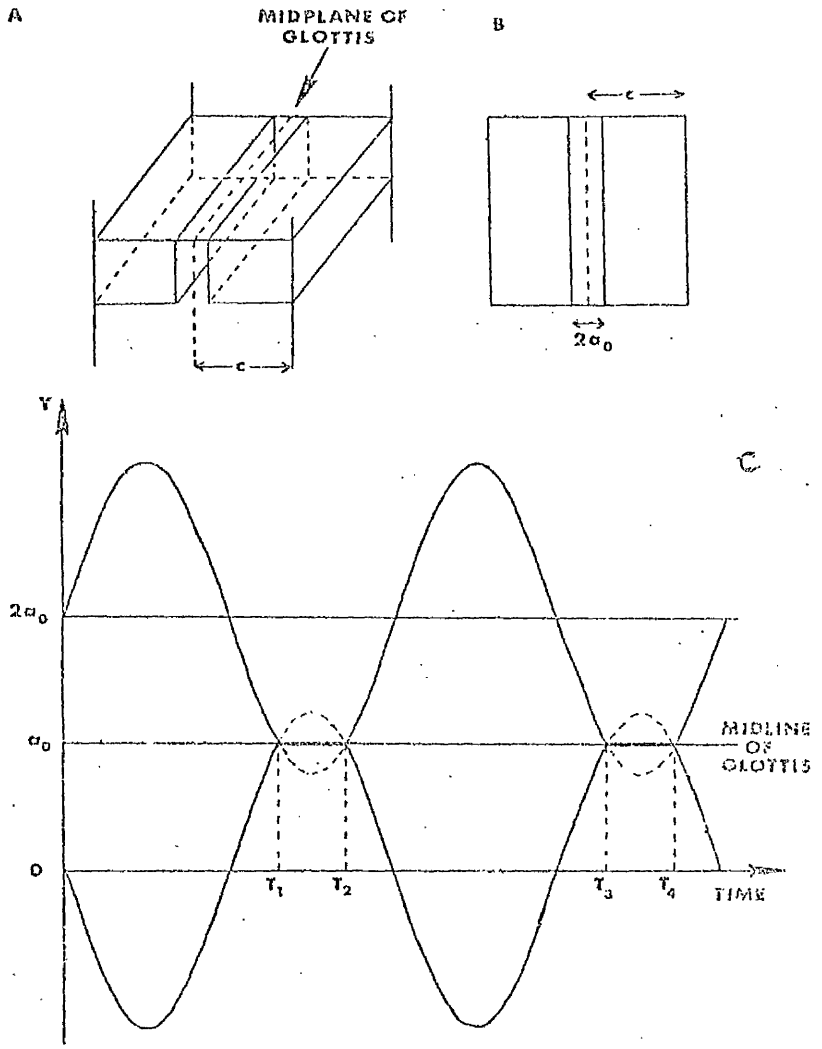


FIG.5.23 SIMPLE MODEL OF VOCAL FOLDS SHOWING (A) THEIR GEOMETRY IN THE LARYNX, (B) THE CROSS-SECTION AT THE LEVEL OF GLOTTIS, AND (C) MOTION OF EACH FOLD.

$$y_2 = 2a_o + a_1 \sin wt \qquad a_1 \geq a_o$$

and this is illustrated in Figure 5.23C.
The folds meet when $y_1 = y_2$ i.e. at times, t_i, such that

$$2a_o + a_1 \sin wt_i = - a_1 \sin wt_i$$

$$\text{i.e. } t_i = \frac{1}{w} \sin^{-1}\left(-\frac{a_o}{a_1}\right) \text{ for } a_1 \geq a$$

These solutions are interpreted to mean that the folds come together at time T_1 and remain in contact until time T_2 (similarly for T_3 and T_4 etc.) and that the area of contact during such times is $A = \frac{V}{C}$.

Now let the folds have different "phonation ready" positions. This can arise in cases of unilateral vocal fold paralysis in which one fold is permanently fixed at a set position in the larynx. Let one fold, the normal fold, vibrate with amplitude a_1, and angular frequency w_1, about a locus a_o from the midline of the larynx. Let the paralysed fold vibrate with amplitude a_2 at angular frequency w_2 about a locus, a, from the midline. (Figure 5.24).

Take the origin as the "phonation ready" position of the normal fold, and the motions of the folds are described by the equations :

$$y_1 = -a_1 \sin w_1 t \quad , \quad a_1 > a_o$$

$$y_2 = a_o + a + a_2 \sin w_2 t \qquad a_2 \leq a_1$$

and the folds come into contact when $y_1 = y_2$, i.e. at times t_i, such that :

$$a + a_o + a_2 \sin w_2 t_i = - a_1 \sin w_1 t_i \quad . \qquad (5.3)$$

i.e. at times T1, T2, T3 etc. (again the second roots are interpreted as the time of the end of contact).

The area of contact is given by the cross sectional area of the more extended fold, and is thus dependent upon the extension. At the time of contact

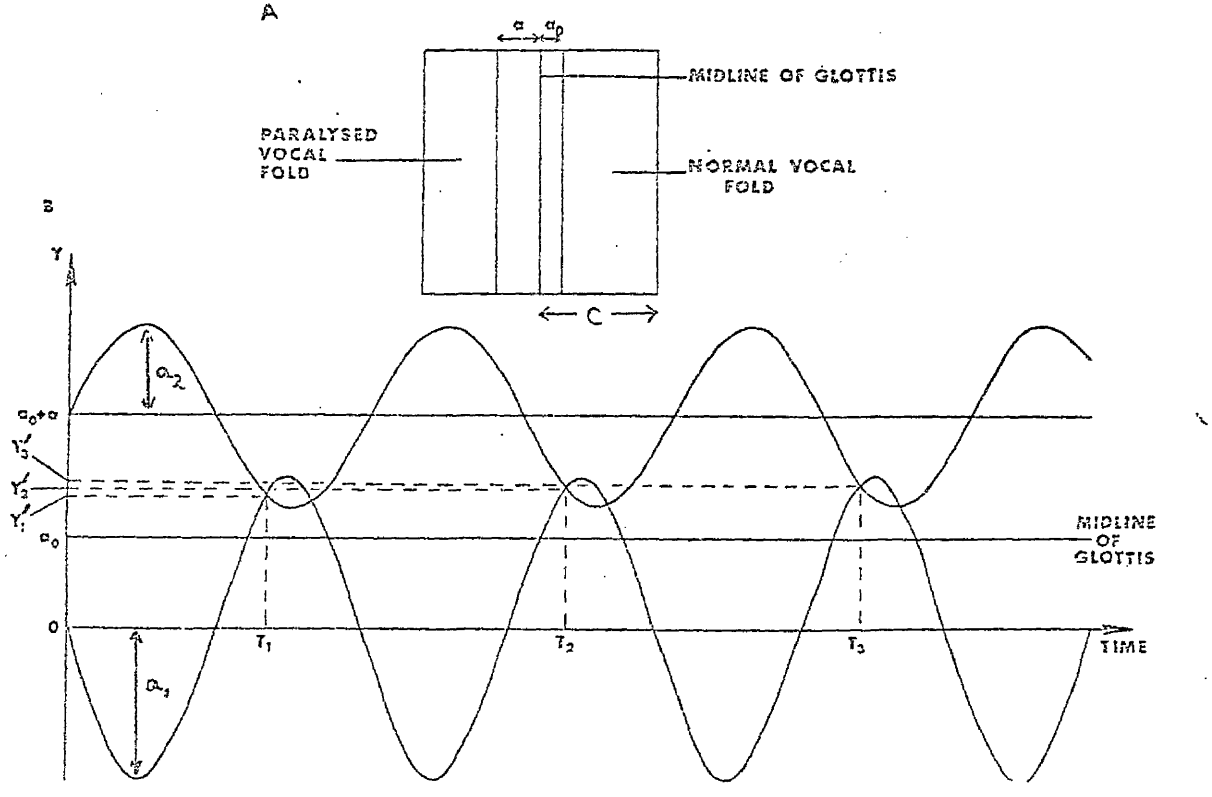


FIG.5.24 MODEL OF FOLDS OF SUBJECT SUFFERING FROM UNILATERAL VOCAL FOLD PARALYSIS.

(A) CROSS-SECTION THROUGH THE GLOTTIS.
(B) MOTION OF EACH FOLD.

the breadth of the normal fold is :

$$C - a_o + y_{1i} = C - a_o - a_1 \sin w_1 t_i = C + a + a_2 \sin w_2 t_i$$

and the area of contact is given by :

$$A_i = \frac{V}{(C + a + a_2 \sin w_2 t_i .)} \qquad (5.4)$$

Equation (5.3) can be written as :

$$- (a + a_o) = a_1 \sin w_1 t_i + a_2 \sin w_2 t_i$$

and so the t_i can be solved for graphically by plotting

$$y = - (a + a_o) \qquad (5.5)$$

and

$$y = a_1 \sin w_1 t + a_2 \sin w_2 t \qquad (5.6)$$

It can be shown (SECTION 5.10) that equation (5.6) simplifies to give :

$$y = a_1 r \sin \theta \qquad (5.7)$$

where $r^2 = 1 + a'^2 + 2a' \cos (w_1 - w_2)t$,

$$\tan \theta = \frac{\sin w_1 t + a' \sin w_2 t}{\cos w_1 t + a' \cos w_2 t} .$$

and $$a' = \frac{a_2}{a_1} \leqslant 1 .$$

Thus, in general, r^2 is periodic with angular frequency $(w_1 - w_2)$.

[In the particular case of $a_1 = a_2$ i.e. $a' = 1$

$$r^2 = 4 \cos^2 \frac{(w_1 - w_2)}{2} t$$

and is thus periodic with angular frequency $\frac{1}{2} (w_1 - w_2)$.]

An example of the form equation (5.7) is plotted in Figure 5.25, and the conditions resulting from equation (5.5) can be interpreted.

In case I, (Figure 5.25) if $a + a_o \leqslant a_1 - a_2$, then the folds come into contact at some point during each cycle, but at different phases in the vibratory pattern.

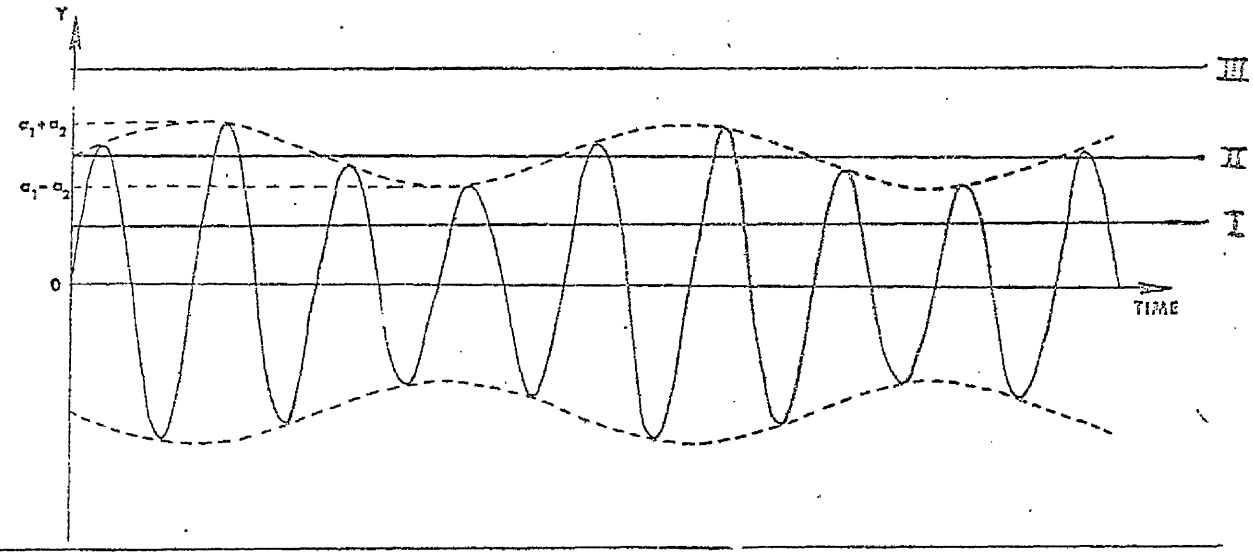


FIG.5.25 GRAPH ILLUSTRATING REGULAR VARIATION IN THE POSITION OF FOLD CONTACT LEADING TO BEATS IN THE AREA OF CONTACT.

This is illustrated in Figure 5.24B, where successive contacts are made at times T_1, T_2, T_3, etc., which occur at different phases in the oscillatory pattern, i.e. $y_1^1 \neq y_2^1 \neq y_3^1 \ldots$. The cyclical nature of this variation leads to a regular variation in the area of contact during successive cycles because of the relationship specified by equation (5.4).

In case II, (Figure 5.25) if $a_1 - a_2 < a + a_o \leq a_1 + a_2$, then the folds make contact for a series of consecutive cycles, followed by a period during which no contact is made, and this pattern is repeated.

In case III, (Figure 5.25) if $a + a_o > a_1 + a_2$, there is no contact between the vocal folds.

This model of the vocal folds was much too simple to describe realistically their complex motion during phonation. The folds are certainly not cuboid in shape; no account was taken of the fact that the upper and lower margins of the folds vibrate slightly out of phase, and no consideration was given to the event of contact, or to what adjustments take place during contact. However, the model did predict the possibility of beats occurring in the area of contact of the folds, and also the possibility that for several cycles there might be no contact at all.

Since the Lx waveform provided some measure of the area of vocal fold contact, the model could be used to try to explain some of the features observed in such outputs, for example, the illustration already described in Figure 5.22. In the portion A-B, the folds made complete contact only two or three times in every five or six cycles and this was consistent with a case II situation with the folds vibrating at frequencies of about 110Hz and 140Hz. The portion B-C was consistent with a case I situation with the folds still vibrating at these frequencies. In the final stages of this phonation, the vibratory pattern was further modified to produce complete contact only every fourth cycle, probably by a further complication

in the oscillatory pattern of the abnormal fold. This output was obtained from a subject suffering from a unilateral vocal fold paralysis, with the paralysed fold remaining in a medial position. Since this fold was still subjected to the aerodynamic forces, it would vibrate during phonation, but because the adjustment of mass and/or tension was impaired, it had a lower natural vibrational frequency than the other fold.

A similar analysis held for the subject whose phonations are illustrated in Figure 5.26. This subject was also suffering from a unilateral vocal fold paralysis with the affected fold maintained in a medial position. The diagram shows the intensity of the Lx signal obtained during four different phonations, and regular oscillations were evident in the intensities, at frequencies between 8 - 12Hz. This was consistent with a case I situation, and the fundamental frequency of the normal fold was about 115Hz, giving a vibrational frequency for the paralysed fold of between about 100 - 110Hz. It was subsequently confirmed under stroboscopic laryngeal examination that the paralysed fold did indeed vibrate during phonation. Thus even a very simple model of the vocal folds could predict some aspects of the vibratory pattern.

The fact that in normal phonation the vocal folds vibrate at the same fundamental frequency may not only reflect the fact that the folds are adjusted symmetrically, but it may be indicative of the strong coupling which exists between the folds. It seems feasible that because of the many combinations of mass and tension which could combine to result in a given natural frequency of oscillation, the possibility of asymmetry in the adjustment of the folds may not be negligible, and that in such cases, the coupling acts to stabilise the vibratory pattern at one fundamental frequency. Thus it seems reasonable, that under conditions of reduced coupling, e.g. the fixation of one fold at a position further than normal from the glottal midline, and extreme asymmetry in the

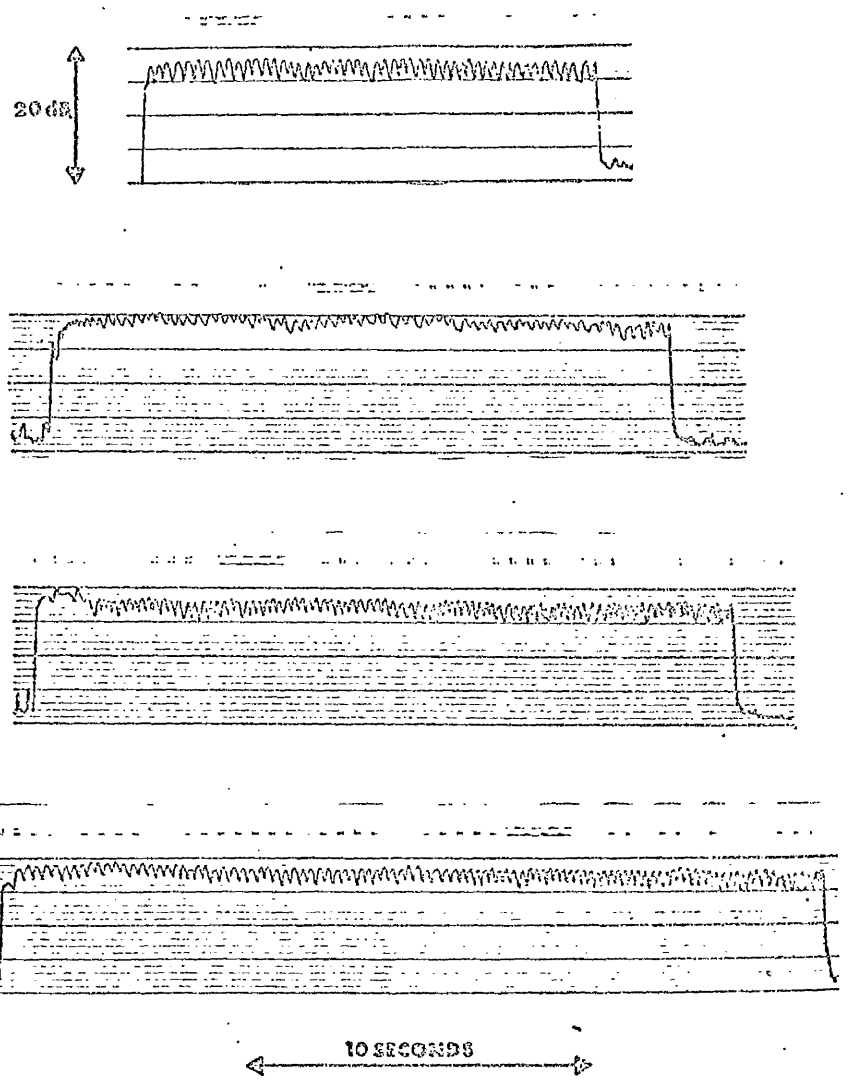


FIG.5.26 EXAMPLES OF RECORDS OF INTENSITY OF Lx OUTPUT DURING PHONATION OBTAINED FROM SUBJECT SUFFERING FROM UNILATERAL VOCAL FOLD PARALYSIS.

adjustment of the folds, the natural frequencies of each fold might be so different as to be beyond compensation.

5.9. DERIVATION OF FOURIER COEFFICIENTS, c_m and d_m.

5.9.1 USEFUL MATHEMATICAL RELATIONSHIPS

a) $$\int_0^{2\pi} \sin nx.dx = \int_0^{2\pi} \cos nx.dx = 0 \qquad (5.8)$$

b) $$\int_0^{2\pi} \sin nx.\cos mx.dx = 0 \qquad (5.9)$$

c) $$\int_0^{2\pi} \sin nx.\sin mx\,dx = \int_0^{2\pi} \cos nx.\cos mx.dx = \delta_{m\,n} \qquad (5.10)$$

where $$\delta_{m\,n} = \begin{cases} 1 & m = n \\ 0 & m \neq n \end{cases}$$

d) $$\int_0^{2\pi} \cos 2nx.\sin x.\cos mx.dx$$

$$= \frac{1}{2}\int_0^{2\pi} \sin (2n+1)x.\cos mx\,dx - \frac{1}{2}\int_0^{2\pi} \sin (2n-1)x.\cos mx.dx$$

$$= 0$$

from (5.9) (5.11)

e) $$\int_0^{2\pi} \sin 2nx.\sin x.\sin mx.dx$$

$$= \frac{1}{2}\int_0^{2\pi} \cos (2n-1)x.\sin mx.dx - \frac{1}{2}\int_0^{2\pi} \cos (2n-1)x.\sin mx.dx$$

$$= 0$$

from (5.9) (5.12)

f) $$\int_0^{2\pi} \sin 2nx.\sin x.\cos mx.dx$$

$$= \frac{1}{2}\int_0^{2\pi} \cos(2n-1)x.\cos mx.dx$$

$$- \frac{1}{2}\int_0^{2\pi} \cos(2n+1)x.\cos mx.dx$$

$$= \frac{\pi}{2}\,\delta_{2n-1,m} - \frac{\pi}{2}\,\delta_{2n+1,m} \qquad (5.13)$$

from (5.10)

g) $$\int_0^{2\pi} \cos 2nx.\sin x.\sin mx.dx$$

$$= \frac{1}{2}\int_0^{2\pi} \sin(2n+1)x \,.\, \sin mx.dx$$

$$- \frac{1}{2}\int_0^{2\pi} \sin(2n-1)x.\sin mx.dx$$

$$= \frac{\pi}{2}\,\delta_{2n+1,m} - \frac{\pi}{2}\,\delta_{2n-1,m} \qquad (5.14)$$

from (5.10)

5.9.2 c_m

From equation (5.1) (SECTION 5.7), Q(t) could be described by a Fourier series such that :

$$Q(t) = \frac{1}{2}a_o + \sum_{n=1}^{\infty} (a_n \cos 4\pi nft + b_n \sin 4\pi nft)$$

The function M(t) was defined as :

$$M(t) = Q(t) \,.\, (1 - p \sin 2\pi ft)$$

and could also be described by a Fourier series :

$$M(t) = \frac{1}{2}c_o + \sum_{m=1}^{\infty} (c_m \cos 2\pi fmt + d_m \sin 2\pi fmt)$$

where $$c_m = 2f\int_0^T M(t)\cos 2\pi fmt.dt \qquad (5.15)$$

and $$d_m = 2f\int_0^T M(t)\sin 2\pi fmt\,dt \qquad (5.16)$$

Make the substitution $x = 2\pi ft$, and consider eqn. (5.15)

$$c_m = \frac{1}{\pi}\int_0^{2\pi} M(x).\ \cos mx.dx$$

$$= \frac{1}{\pi}\int_0^{2\pi}\left[\frac{1}{2}a_o + \sum_n (a_n \cos 2n\,x + b_n \sin 2nx)\right].$$
$$(1-p \sin x)\ .\ \cos mx.dx$$

$$= \frac{1}{\pi}\int_0^{2\pi}\Bigg[\frac{1}{2}a_o.\cos mx. - \frac{1}{2}a_o.p.\sin x.\cos mx +$$
$$\sum_n (a_n.\cos 2nx + b_n \sin 2nx).\cos mx$$
$$-p.\sin x.\cos mx. - \sum_n (a_n \cos 2nx + b_n \sin 2nx)\Bigg]\,dx$$

$$= \frac{1}{\pi}\int_0^{2\pi} \sum_n a_n.\cos 2nx\ .\ \cos mx\,.dx$$
$$- \frac{p}{\pi}\int_0^{2\pi}\sum_n a_n.\cos 2nx \sin x.\ \cos mx.dx$$

$$- \frac{p}{\pi}\int_0^{2\pi}\sum_n b_n.\sin 2nx.\ \sin x.\ \cos mx.dx.$$

(using (5.8) and (5.9)

$$\text{i.e. } c_m = \sum_n a_n\ \delta_{2n,m} - \frac{p}{2}\sum_n b_n\left(\delta_{2n-1,m} - \delta_{2n+1,m}\right) \qquad (5.17)$$

using (5.10),(5.11) and ((5.13))

5.9.3 d_m

Use the same substitution, $x=2\pi ft$, and eqn.(5.16) becomes :

$$d_m = \frac{1}{\pi}\int_0^{2\pi} M(x).\ \sin mx\ dx$$

$$= \frac{1}{\pi} \int_0^{2\pi} \left[\frac{1}{2} a_o + \sum_n (a_n \cos 2nx + b_n \sin 2nx) \right] .$$

$$(1 - p \sin x) \cdot \sin mx \, dx$$

$$= \frac{1}{\pi} \int_0^{2\pi} \left[\frac{1}{2} a_o \sin mx - \frac{1}{2} a_o p \sin x \sin mx \right.$$

$$+ \sum_n (a_n \cos 2nx + b_n \sin 2nx) \sin mx$$

$$\left. - p \sin x \sin mx \sum_n (a_n \cos 2nx + b_n \sin 2nx) \right] dx$$

$$= \frac{1}{\pi} \int_0^{2\pi} \left(- \frac{1}{2} a_o p \right) \sin x \cdot \sin mx \, dx$$

$$+ \frac{1}{\pi} \int_0^{2\pi} \sum_n b_n \sin 2nx \cdot \sin mx \, dx$$

$$- \frac{p}{\pi} \int_0^{2\pi} \sum_n a_n \cos 2nx \sin x \sin nx \, dx$$

$$- \frac{p}{\pi} \int_0^{2\pi} \sum_n b_n \sin 2nx \cdot \sin x \cdot \sin mx \, dx$$

(using (5.8) and (5.9))

$$\text{i.e. } d_m = - \frac{1}{2} a_o p \, \delta_{1,m} - \frac{p}{2} \sum_n a_n \left(\delta_{2n+1,m} - \delta_{2n-1,m} \right)$$

$$+ \sum_n b_n \cdot \delta_{2n,m}$$

(using (5.10), (5.12) and (5.14))

Thus for m even

$$c_m = a_n \qquad m = 2n \ .$$

$$d_m = b_n$$

for m odd

$$c_m = -\frac{1}{2} p \left(b_{\frac{m+1}{2}} - b_{\frac{m-1}{2}} \right)$$

$$d_m = \frac{1}{2} p \left(a_{\frac{m+1}{2}} - a_{\frac{m-1}{2}} \right) .$$

for m = 1

$$c_1 = -\frac{1}{2} p b_1$$

and

$$d_1 = \frac{1}{2} p (a_1 - a_o) .$$

These results were summarized in TABLE 5.6.

5.10. DERIVATION OF BEAT EQUATION.

It was shown in SECTION 5.8 that according to the model of the vocal folds, the folds came into contact when

$$y = -(a + a_o) \qquad (5.5)$$

and $$y = a_1 \sin w_1 t + a_2 \sin w_2 t. \qquad (5.6)$$

Let $a_1 \geqslant a_2$ so that $a' = \frac{a_2}{a_1} \leqslant 1$,

and let $w_1 \geqslant w_2$

Equation (5.6) becomes :

$$y = a_1 (\sin w_1 t + a' \sin w_2 t) = a_1 S.$$

where $S = \text{Im.} (e^{jw_1 t} + a' e^{jw_2 t})$

(where $e^{j.\theta} = \cos \theta + j \sin \theta$)

and $j = \sqrt{-1}$)

Let $e^{jw_1 t} + a' e^{jw_2 t} = r e^{j\theta}$, so that

$$S = r \sin \theta .$$

Since $r e^{j\theta} = e^{jw_1 t} + a' e^{jw_2 t}$

$= (\cos w_1 t + j \sin w_1 t) + a' (\cos w_2 t + j \sin w_2 t)$

then $r^2 = 1 + a'^2 + 2a' (\sin w_1 t \cdot \sin w_2 t + \cos w_1 t \cos w_2 t)$

$$= 1 + a'^2 + 2a' \cos(w_1 - w_2)t \quad .$$

$$\text{and } \tan \Theta = \frac{\sin w_1 t + a' \sin w_2 t}{\cos w_1 t + a' \cos w_2 t}$$

Thus r^2 is periodic with angular frequency $(w_1 - w_2)$,

and when :

$\cos (w_1 - w_2)t = 1$, $r^2 = (1 + a')^2$, $r = \pm (1 + a')$;

$\cos (w_1 - w_2)t = -1$, $r^2 = (1 - a')^2$, $r = \pm (1 - a')$;

$\cos (w_1 - w_2)t = 0$, $r^2 = (1 + a'^2)$, $r = \pm (1 + a'^2)^{\frac{1}{2}}$

giving the limits of the modulation of the beats.

5.11 SUMMARY

In this chapter the laryngograph technique was used to investigate the vibratory pattern of the vocal folds. The interpretation of the Lx waveform was discussed and normative data for the relative lengths of different parts of the Lx cycle were obtained from a small group of normal speaking adult subjects. It was found that the closing time of the vocal fold vibration was of the order of 5-20% of the period, and was always less than half the opening time. The build up of the Lx amplitude to a steady level was used in the study of vocal initiation, and the time taken for the vocal folds to attain a stable vibratory pattern was measured. Female subjects required less time to achieve this, a finding which reflected the greater mass of male vocal folds. The fact that in most cases the required fundamental frequency was immediately used, indicated that this feature of laryngeal adjustment did not depend solely on auditory feedback, but was under the control of the other feedback control loops.

The frequency analysis of the Lx waveform showed that in general the intensities of the spectral harmonics had a logarithmic relationship with frequency. Idealised waveforms were investigated with the aid of a computer program, and the spectral gradient was used

to predict the value of the closing time. There was fine structure in some of the Lx spectra, i.e. series of maxima and minima, and the frequency of the first subsidiary minimum was also used to provide a measure of the closing time. Good agreement was obtained with real Lx waveforms.

Abnormalities in the voice spectra already noted in CHAPTER 4 were found to have correlates in the Lx spectra. This was especially true of subharmonic components, which showed that these components were generated by the vocal fold vibration. The nature of the vocal fold vibratory pattern during subharmonic production was predicted theoretically and subsequently demonstrated experimentally. A simple model of the vocal folds was used to explain some variations found in the Lx pattern. Although the model was too simple to explain the complex pattern of the vocal fold vibration, it made predictions about the possibility of a beat type variation in the amplitude of the Lx pattern under certain conditions, and examples of these were found on several occasions.

CHAPTER 6.

THE CONTROL OF SUSTAINED PHONATION

CHAPTER 6. THE CONTROL OF SUSTAINED PHONATION

6.1 INTRODUCTION

In the previous chapters of this thesis some of the factors relevant to voice production have been studied individually. The aerodynamic characteristics of the production of a sustained vowel, /a/, were discussed in CHAPTER 3, the frequency spectra of such vowels were described in CHAPTER 4, and the characteristics of the vibratory pattern of the vocal folds were investigated in CHAPTER 5. However, the data referred to in these chapters were, by and large, obtained during different recording sessions, and could thus not be directly related. Also, the phonations discussed in earlier chapters were uncontrolled in the sense that each subject was required to phonate at his subjectively "normal" fundamental frequency and comfortable sound level, or about 70dB. This approach was suitable from the point of view of the clinical assessment, and also for the acquisition of normative data, but in order to investigate the relationships between the various factors involved in voice production, a much more controlled experimental approach was necessary.

The aim of the experiments discussed in this chapter was to investigate the control of a sustained vowel phonation, and, in particular, to examine the relationships between the sound intensity of the phonation and the air flow rate, and to relate these to the vibratory pattern of the vocal folds.

Several previous studies had utilized steady sustained phonations, i.e. phonations maintained at constant sound levels and fundamental frequencies, to investigate the variation of flow rate with sound level. Although Isshiki (1964) found that there was no definite relationship between sound intensity and air flow rate, Isshiki (1965), and Yanagihara and

Koike (1967) reported that, in general, the variation of air flow rate as a function of sound intensity was greater at high fundamental frequencies than at low fundamental frequencies. Isshiki (1964) concluded that at a very low fundamental frequency, the vocal intensity was controlled mainly by the glottal resistance, (i.e. by the laryngeal musculature), whereas at a very high fundamental frequency, the vocal intensity was controlled almost solely by the air flow rate, i.e. the muscles controlling expiration.

6.2 EXPERIMENTAL TECHNIQUE

The apparatus was set up as shown in Figure 6.1. In order to maintain the area of lip opening at a constant value, the subject was required to close his lips around the pneumotach flowhead, and to produce an airtight seal. The flowhead had an external diameter of 2.9cms. The air flow rate and volume expired during phonation were recorded, as before, on two channels of the Mingograf ink jet recorder. The condenser microphone was placed at a fixed distance of 20cms. from the end of the flowhead, and was covered by a foam rubber windshield to prevent large fluctuations in the output being caused directly by the airflow. The efficiency of the microphone for acoustic signals was not altered by the windshield. The microphone output was recorded on one channel of the stereo tape recorder, and was also used to devise a voltage proportional to the sound intensity via the level recorder and analogue read-out device (CHAPTER 2). This signal was recorded on a third channel of the Mingograf recorder. The laryngograph output was recorded on the second channel of the tape recorder, and was also utilized to provide a voltage proportional to the fundamental frequency of phonation via the frequency-voltage converter circuit described in CHAPTER 2. This signal was recorded on the fourth channel of the Mingograf, so that, during phonation simultaneous recordings were

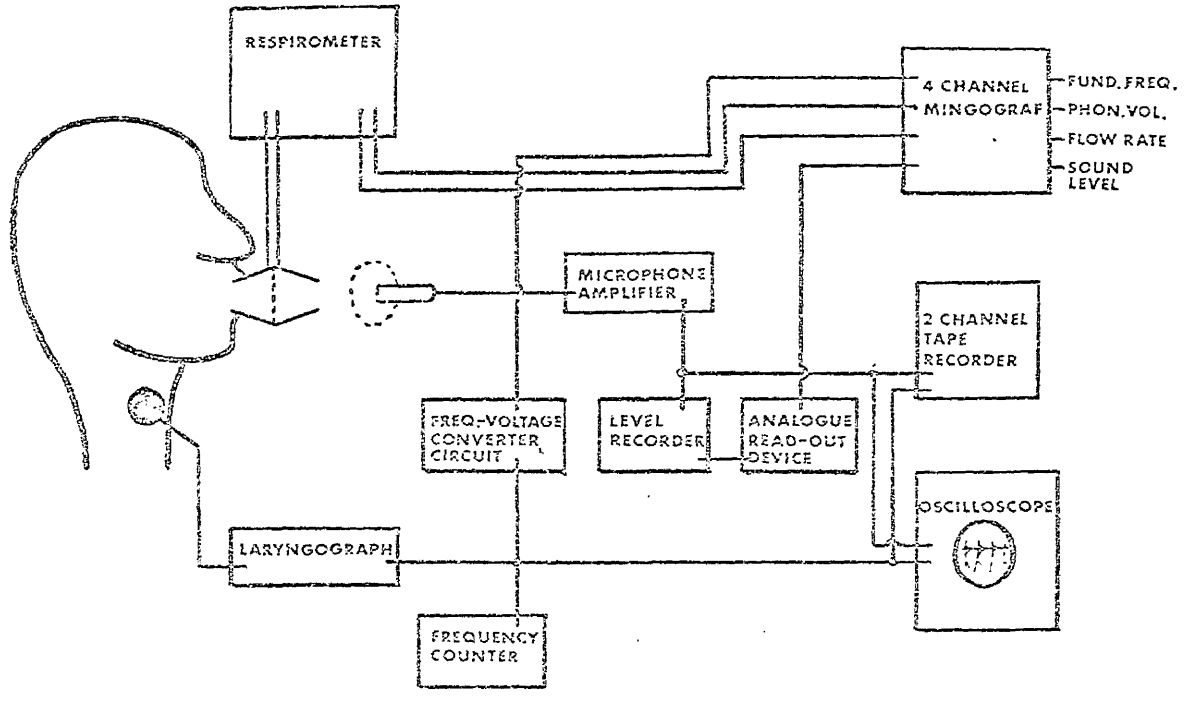


FIGURE 6.1 EXPERIMENTAL ARRANGEMENT FOR THE SIMULTANEOUS RECORDING OF SOUND LEVEL, AIR FLOW RATE, FUNDAMENTAL FREQUENCY, VOICE OUTPUT AND Lx OUTPUT.

made of the fundamental frequency, phonation volume, air flow rate and sound level.

The area of lip opening of the subject was kept constant by the flowhead, but this also added on a hard walled cavity to the vocal tract, and thus modified the subject's voice output. The effect of this addition to the vocal tract was investigated in a preliminary experiment.

A series of recordings was made with the recording system illustrated in Figure 6.1. Initially the subject's nose was clamped off to eliminate nasal air loss, but this proved to be uncomfortable, and the air loss was found to be minimal, so that the nose clamp was dispensed with. The vowel which the subject could most easily maintain was the vowel /ɔ/, and this vowel was used throughout. The phonations were maintained at a constant fundamental frequency and subjective voice effort, and during each, the flowhead was removed from the subject's mouth while he held his lips in the same configuration. The changes in the intensities due to the presence of the flowhead were measured from the Mingograf traces, and it was found that the effect of the flowhead was to add 8dB ($\pm$0.5dB) to the sound level measured at a distance of 20cms. This was due to the resonant characteristics of the hard walled tube, and also to a change in the directional properties. This increase in sound level was found to be constant for a range of fundamental frequencies from 125Hz to 220Hz, and also over the range of sound levels of interest, and so it was not necessary to correct for it unless absolute sound level measurements were required.

Frequency spectra of each phonation were obtained using the sonagraph, and the differences measured before and after the removal of the flowhead from between the subject's lips were directly attributable to the presence of the flowhead. Thus, using the

phonations made at different fundamental frequencies it was possible to obtain a composite frequency spectrum of the transfer characteristics due to the addition of the flowhead to the vocal tract, and this is shown in Figure 6.2. The first four peaks were well defined, but at frequencies above about 2KHz there were few data points due to the lack of high frequency components in the spectra of the vowel sound which could be produced with the particular vocal tract configuration. The transmission spectrum had only to be taken into account when quantitative results were required from the spectra of similar recordings.

6.3 DATA FROM STEADY PHONATIONS

The mean flow rate data obtained from the low pitch, normal pitch and high pitch phonations of the control group of subjects (CHAPTER 3) were plotted as a function of the frequency, and some are shown in Figure 6.3. These results indicated that, in general, an increase in the fundamental frequency of phonation was accompanied by an increase in the air flow rate. However, these phonations had been produced at a comfortable sound level (about 70dB) and at a subjective value for fundamental frequency. Also the sound level could not be measured because of the unknown effects of the rubber face mask and mask microphone.

A series of phonations was recorded using the experimental system shown in Figure 6.1. The subject was a 25 year old, normal speaking untrained male subject, and each phonation was maintained at a constant sound level and fundamental frequency for between 5 and 10 seconds. The subject monitored his fundamental frequency by observing the digital frequency counter. The values of fundamental frequency, sound level and air flow rate were measured for each phonation, and some typical results are shown in Figure 6.4.

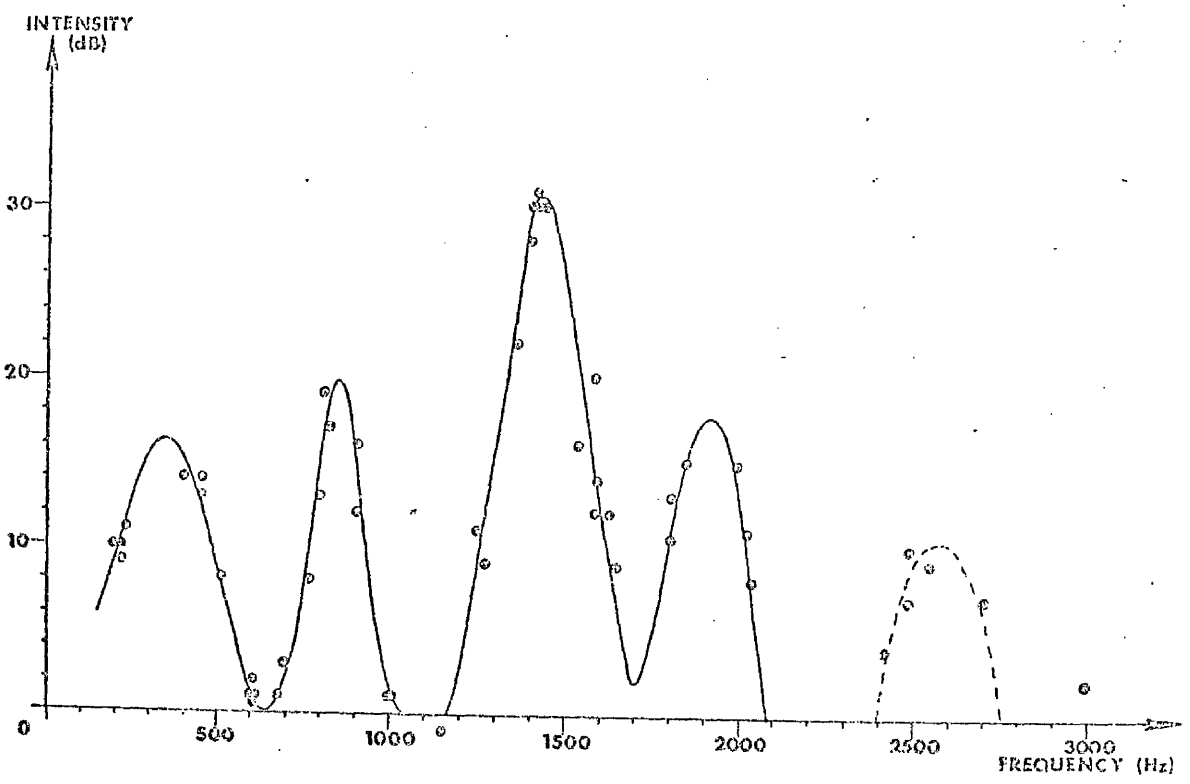


FIGURE 6.2. FREQUENCY SPECTRUM SHOWING TRANSMISSION CHARACTERISTICS CAUSED BY THE ADDITION OF THE PNEUMOTACH FLOWHEAD TO THE VOCAL TRACT.

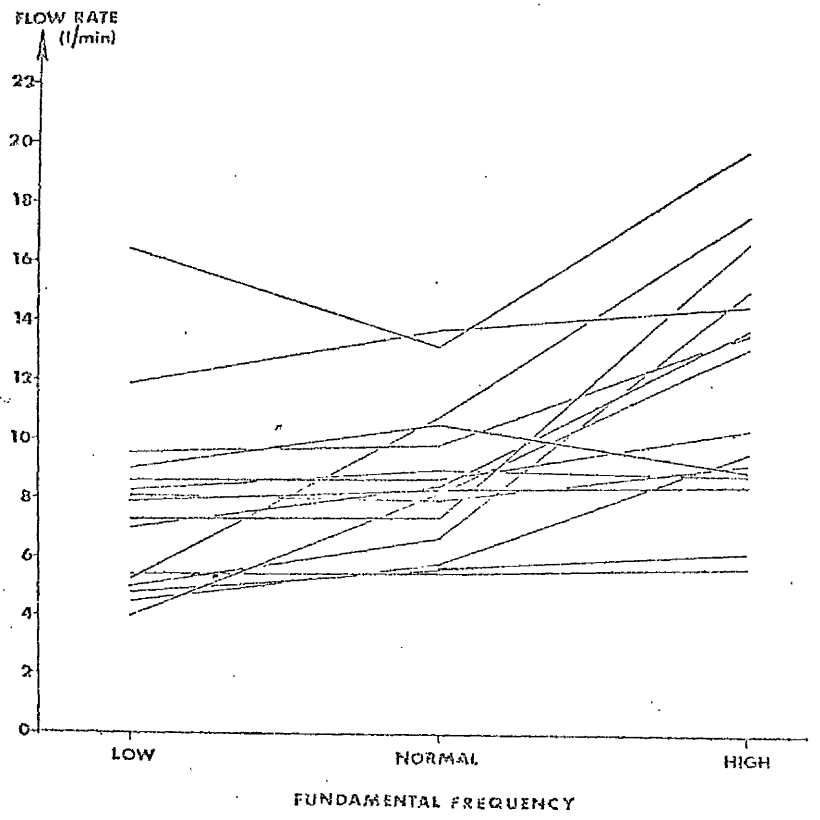
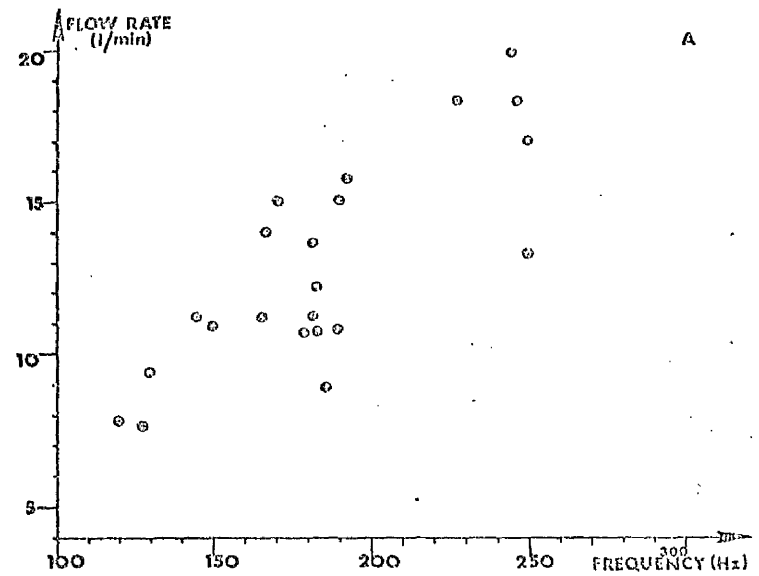


FIGURE 6.3. DATA FROM CONTROL GROUP SHOWING FLOW RATE PLOTTED AS A FUNCTION OF LOW, NORMAL AND HIGH PITCH.



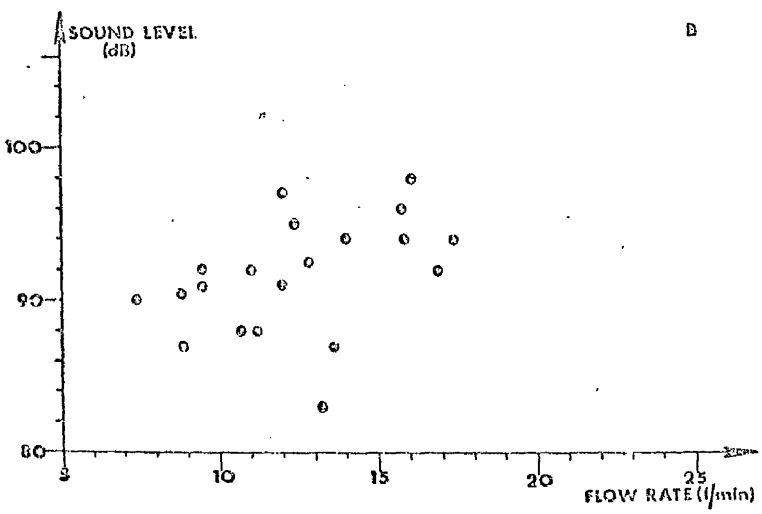


FIGURE 6.4

TYPICAL DATA OBTAINED FROM STEADY, CONTROLLED PHONATIONS.

(A) GRAPH OF FLOW RATE v FUNDAMENTAL FREQUENCY FOR SOUND LEVEL OF 85 - 90dB.

(B) GRAPH OF SOUND LEVEL v FLOW RATE FOR FUNDAMENTAL FREQUENCY IN RANGE 180 - 185Hz.

In Figure 6.4A, the data for the phonations having sound levels in the range 85 - 90dB were plotted as a graph of flow rate v fundamental frequency, and in Figure 6.4B, the data for phonations having fundamental frequencies in the range 180 - 185Hz were plotted as a graph of sound level v air flow rate. The data in the upper graph indicated that, in general, an increase in the fundamental frequency of phonation was accompanied by an increase in the air flow rate, but the spread of the data points was somewhat large. The results in the lower graph indicated that there was no definite relationship between sound level and air flow rate evident from this experiment, and that a particular value of sound level could be maintained over a wide range of air flow rates.

This suggested that the information which could be gained from such an experiment was limited, and that, perhaps, more information could be gained by perturbing the system, i.e. by investigating phonations which were non-steady with respect to sound level and/or fundamental frequency.

6.4 DATA FROM NON-STEADY PHONATIONS - THE RELATIONSHIP BETWEEN SOUND LEVEL AND AIR FLOW RATE.

A series of maximally sustained phonations was now recorded. These were each maintained at a fixed fundamental frequency, which was visually monitored by the subject using the digital frequency counter, but the sound level was deliberately varied throughout the phonation. Two typical examples of the outputs obtained during such phonations are shown in Figure 6.5, and in both cases a definite relationship between sound level and air flow rate was evident.

The example shown in Figure 6.5A was maintained at a fundamental frequency of 176Hz for 16.5 seconds, and measurements of sound level and air flow rate were taken at 0.5sec. intervals. The plot of sound level v air flow rate is shown in Figure 6.6A. The relationship between sound level (SL) and air flow rate (FR) was of the form :

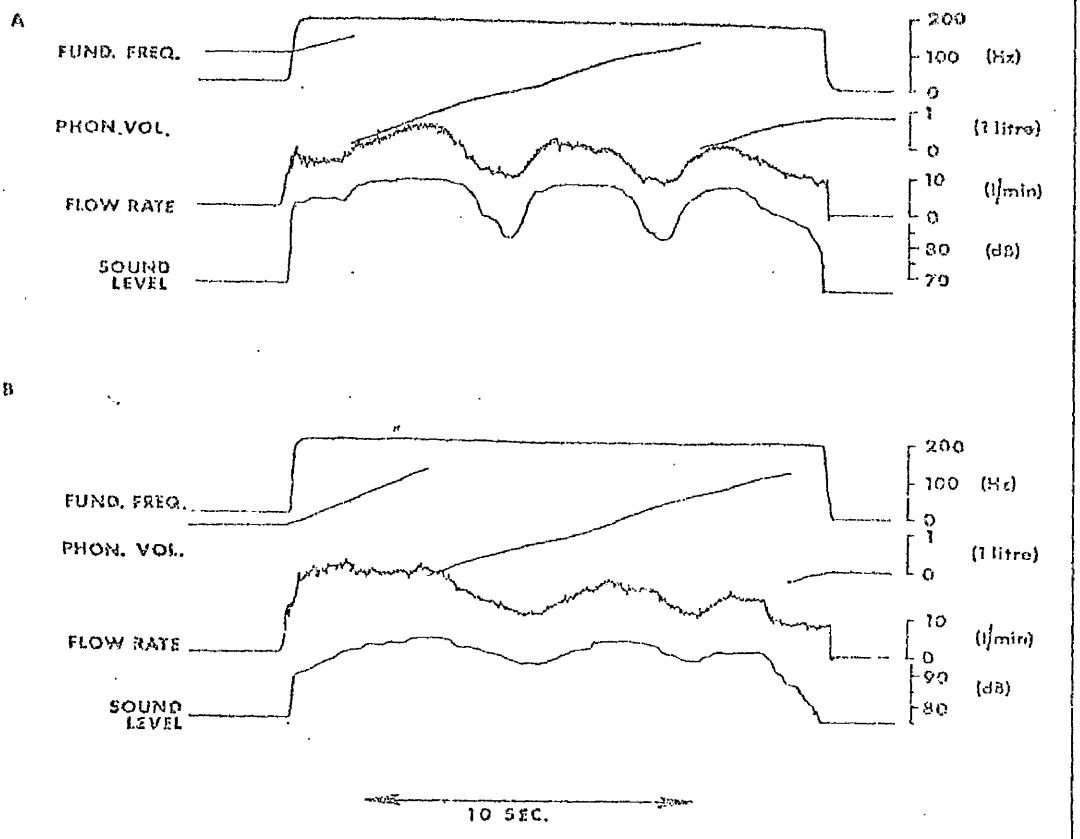


FIGURE 6.5.

EXAMPLES OF TYPICAL OUTPUTS SHOWING SIMULTANEOUS RECORDING OF FUNDAMENTAL FREQUENCY, PHONATION VOLUME, AIR FLOW RATE AND VOICE SOUND LEVEL.

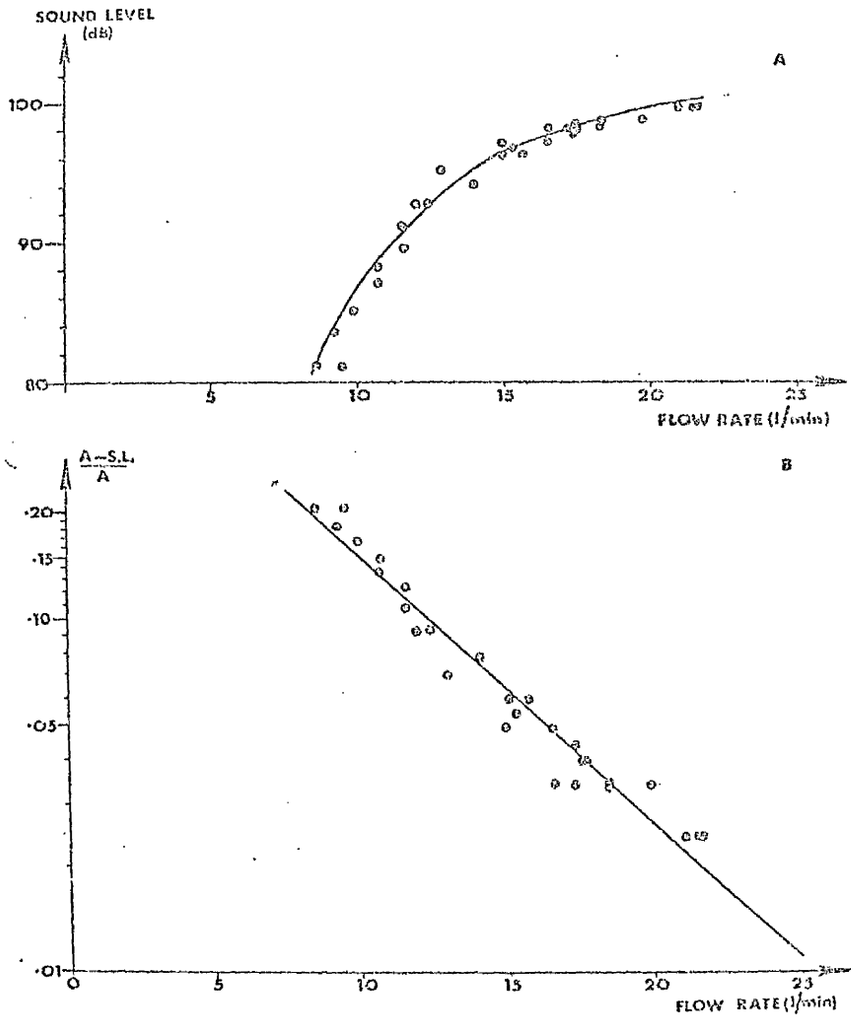


FIGURE 6.6.

SOUND LEVEL v AIR FLOW RATE GRAPHS FOR PHONATION SHOWN IN FIGURE 6.5A.

$$SL = A\,(1 - k e^{-\alpha . FR}) = A(1 - e^{-\alpha . FR + \beta}) \quad \text{—(6.1)}$$

where A, α, β and k were constants for the phonation, and

$$k = e^{\beta}$$

Equation (6.1) was linearized to give the form :

$$\ln\left(\frac{A-SL}{A}\right) = -\alpha . FR + \ln k \quad \text{———(6.2)}$$

The values of the parameter A, α and k were obtained by using a least squares fit computer program, and were found to have the values A = 100.0, α = 0.26, k = 1.878, with a highly significant regression coefficient, r = 0.980, ($p < 0.001$). Thus this form of equation proved to fit the data satisfactorily, and the graph of (A-SL)/ A v FR is shown in Figure 6.6B, together with the line of best fit.

However, the plot of sound level v air flow rate resulting from the phonation illustrated in Figure 6.5B, had a slightly different form, and is shown in Figure 6.7A. The phonation, which lasted for 16.2 secs, had a fundamental frequency of 200Hz and was initiated at the point marked X. The sound level increased rapidly with only a small change in flow rate until the point Y was reached, and thereafter the phonation followed the same pattern as already described, with the data points showing an exponential relationship of the form given in equation (6.1). The parameter A was estimated to have the value 103.4dB, and the graph of (A-SL)/A v FR is shown in Figure 6.7B, with (A-SL)/A again being plotted on a logarithmic scale. The values of α and k were calculated as before, using only the data points obtained from the phonation after the point Y, and the regression line is shown in the diagram. The values obtained for the parameters were α = 0.108 and k = 0.322 with the correlation coefficient being r = 0.973 ($p < 0.001$).

A series of such recordings was made, and in each case a regression line of the form of equation (6.1) was

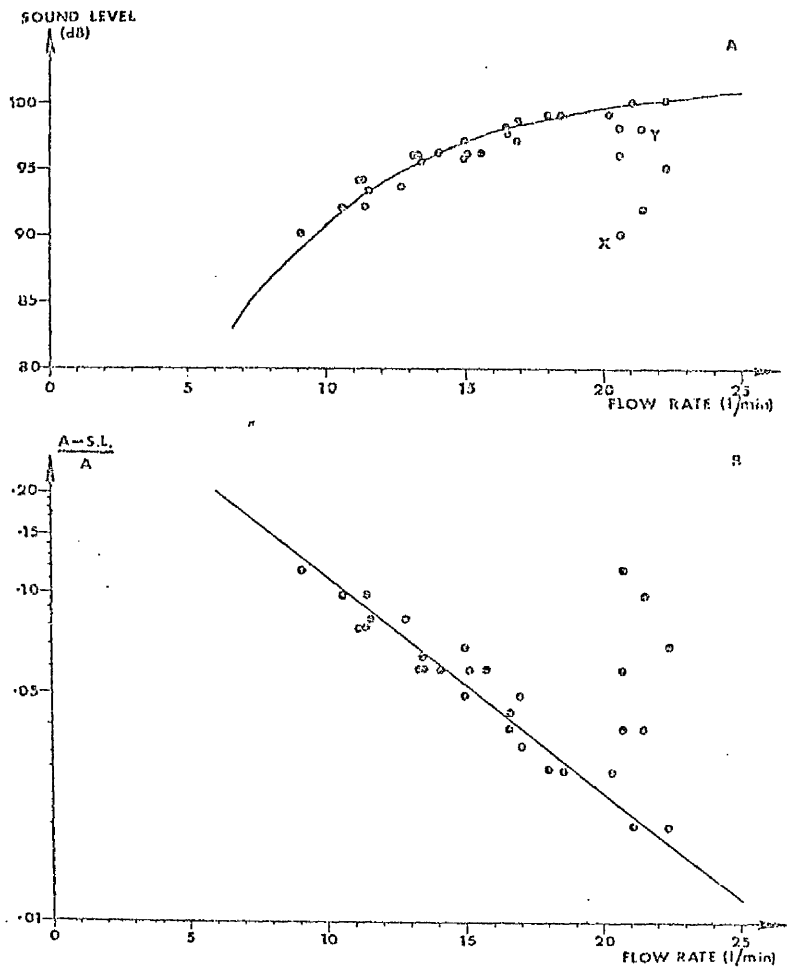


FIGURE 6.7.

SOUND LEVEL v AIR FLOW RATE GRAPHS FOR PHONATION SHOWN IN FIGURE 6.5B.

fitted. The regression coefficients indicated that these lines represented satisfactory descriptions of the relationship between SL and FR ($p < 0.001$), and the values of the parameters are given in TABLE 6.1. The fundamental frequencies for the phonations varied from 150Hz to 287Hz, but there was no relationship between either fundamental frequency and α, or fundamental frequency and k, or indeed between α and k. It seemed that, in general, there existed an exponential relationship between the SL and FR, with the α coefficient being the order of 0.15. The mean values of the parameters were :

$$\bar{\alpha} = 0.154 \text{ /L/MIN.} \quad (\text{S.D.} = 0.072 \text{ /L/MIN.})$$

$$\bar{k} = 0.718 \quad (\text{S.D.} = 0.575)$$

and

$$\bar{A} = 100.2 \text{ dB} \quad (\text{S.D.} = 3.4\text{dB})$$

The parameter A represented an upper limit to the sound level which could be produced with the particular vocal tract configuration, and this value was asymptotically approached at large values of flow rate. The equation of the regression line showed that a particular change in flow rate produced a greater change in sound level at low flow rates than at high flow rates, i.e.

$$\Delta(\text{SL}) = A.\alpha.k.e^{-\alpha.FR}.\ (\Delta\ FR)$$

The locus given by the regression line for each phonation represented a state of equalibrium or maximum efficiency for this type of phonation. (SL,FR) states lying to the left of the locus were never found in phonations in which the fundamental frequency was maintained constant, although on several occasions such states did exist during periods when the fundamental frequency was varying rapidly. States lying to the right of the locus did exist, and represented states in which the sound level was not maximal for

PHON.	FUNDA-MENTAL FREQ. (Hz)	A	α	k	CORR. COEFF. r	SIGNIF. LEVEL.
1	150	93.6	0.225	0.579	0.921	< 0.001
2	167	91.7	0.271	1.157	0.920	< 0.001
3	175	95.8	0.174	0.500	0.939	< 0.001
4	176	100.0	0.260	1.878	0.980	< 0.001
5	179	99.5	0.110	0.361	0.897	< 0.001
6	191	102.3	0.194	1.092	0.971	< 0.001
7	193	103.3	0.179	0.568	0.987	< 0.001
8	200	99.9	0.259	1.809	0.890	< 0.001
9	200	103.4	0.108	0.322	0.973	< 0.001
10	200	98.2	0.166	0.471	0.936	< 0.001
11	200	102.7	0.121	0.365	0.968	< 0.001
12	203	102.7	0.089	0.366	0.934	< 0.001
13	204	102.2	0.145	0.508	0.980	< 0.001
14	204	100.9	0.158	0.499	0.972	< 0.001
15	212	101.2	0.059	0.158	0.942	< 0.001
16	213	101.0	0.294	1.893	0.961	< 0.001
17	214	101.9	0.106	0.385	0.989	< 0.001
18	216	97.8	0.098	0.312	0.971	< 0.001
19	217	101.0	0.079	0.247	0.924	< 0.001
20	219	106.7	0.060	0.322	0.853	< 0.001
21	222	101.8	0.292	2.08	0.930	< 0.001
22	230	97.4	0.156	0.719	0.944	< 0.001
23	233	97.3	0.169	1.415	0.971	< 0.001
24	250	95.1	0.188	0.303	0.944	< 0.001
25	250	97.4	0.073	0.140	0.897	< 0.001
26	257	103.6	0.054	0.338	0.973	< 0.001
27	262	102.1	0.151	0.692	0.982	< 0.001
28	275	101.3	0.112	0.303	0.956	< 0.001
29	287	104.7	0.102	1.038	0.912	< 0.001

TABLE 6.1

VALUES OF A, α AND k CALCULATED USING "LEAST SQUARES" METHOD.

the particular air flow rate. The locus represented the limit in which the maximum sound level was produced for each value of flow rate, and as such could be regarded as the locus of maximum efficiency, if only flow rate and sound level were considered. This phenomenon explained the spread of results illustrated in Figure 6.4 which consisted of data points randomly spread throughout the permissible portion of the SL v FR space. Such phonations could be sustained by the maintenance of suitable adjustment of the respiratory and laryngeal musculature. Phonations such as that illustrated in Figures 6.5B and 6.7 occurred in about one third of the recorded series, and the trend was always to move towards and then along the equilibrium locus. It was possible that these phonations were initiated at maximal lung volume, and that the initial phases of the SL v FR curves reflected a different adjustment of the relevant musculature, e.g. the initial activity of the diaphragm at full lung volumes referred to by Draper et al (1959) and Proctor (1974). An analysis of the volumes expired before and after the equilibrium locus had been attained in each phonation, revealed no significant relationship.

A similar series of experiments was attempted in which the sound level was to be maintained at a constant value, while the fundamental frequency was varied continuously. However, it proved impossible for the subject to maintain the sound level constant, and no meaningful results were obtained.

If the 8dB correction was made to the sound level measurements to allow for the effect of the flowhead, there was no effect on the gradient of the (A-SL)/A v FR graph, i.e. the value of $\propto$ remained unchanged. However, the parameter k would be replaced by k', where $k' = \frac{A}{A-8} .k$. Since A had a fairly constant value of about 100dB, k' was about 8% greater than k for all curves, and so made little difference to the analysis.

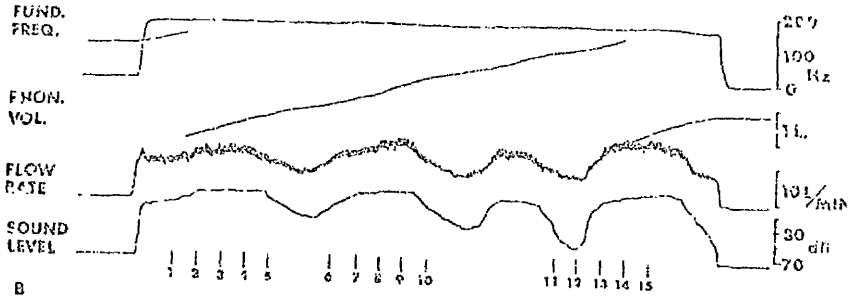
6.5 DATA FROM NON-STEADY PHONATIONS - THE DEPENDENCE ON THE VOCAL FOLD VIBRATORY PATTERN.

The variation of sound level with changes in air flow rate during the sustained phonation of the vowel /ɔ/ was discussed in the previous section. However, in terms of a transmission line model of the vocal tract, one method of controlling the sound level would be to either relatively increase or decrease the intensities of the higher frequency harmonic components of the source function, in relation to the lower frequency components. This would necessitate an alteration in the vibratory pattern of the vocal folds, especially during the closing phase, since this was shown to correspond to the period of maximum excitation of the vocal tract. The aims of the experiments described in this section were to relate the control of sound level to the vibratory pattern of the vocal folds as measured by the Lx waveform, both in terms of the frequency content of the Lx spectrum and also in terms of the closing time of the vocal fold vibration.

Several of the phonations discussed in SECTION 6.4 were analysed in terms of frequency using a two channel sonagraph which allowed the harmonic components of both the microphone output and the Lx output to be obtained at the same points in the recording, and two typical examples will be discussed in detail.

The phonation, whose respirometry outputs are shown in Figure 6.8A, lasted for 19.5 secs., and was maintained at a fundamental frequency of 167Hz. The sound level v flow rate graph is shown in Figure 6.8B, and the equilibrium locus had parameters A = 91.7dB, α = 0.271, k = 1.157, with the correlation coefficient, r = 0.920 ($p < 0.001$). The recordings were analysed using the two channel sonagraph, and these outputs were accurately aligned with the respirometry trace using the discontinuities in the sound level. Section displays of both the microphone and Lx outputs were obtained at the points marked 1 - 15 in Figure 6.8A.

A



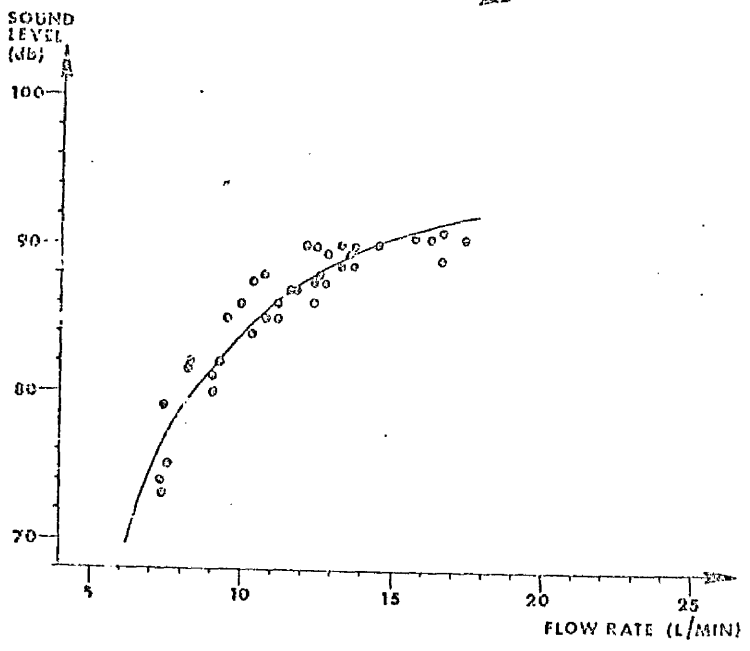


FIGURE 6.8

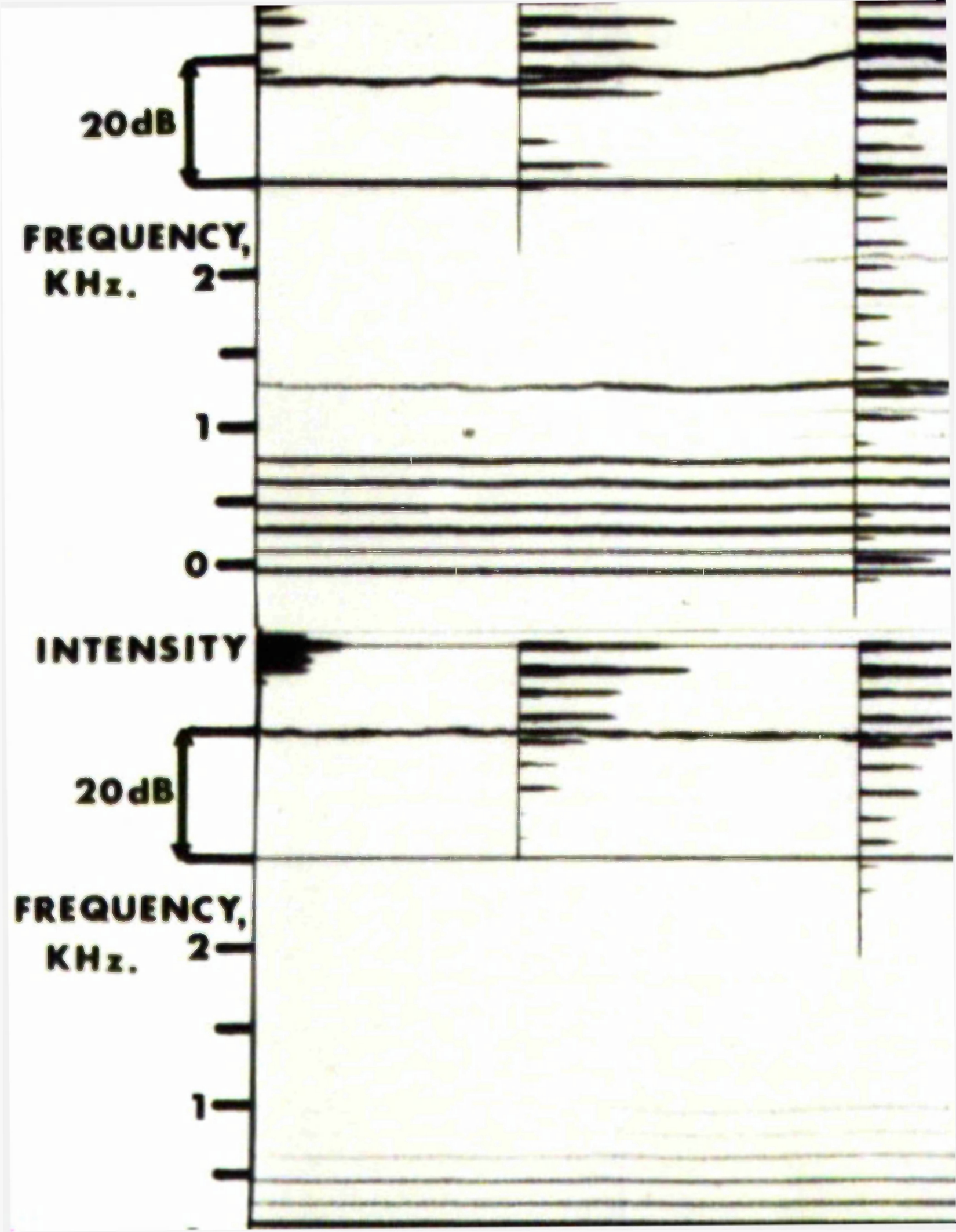
(A) SIMULTANEOUS RECORDINGS OF A NON-STEADY PHONATION.

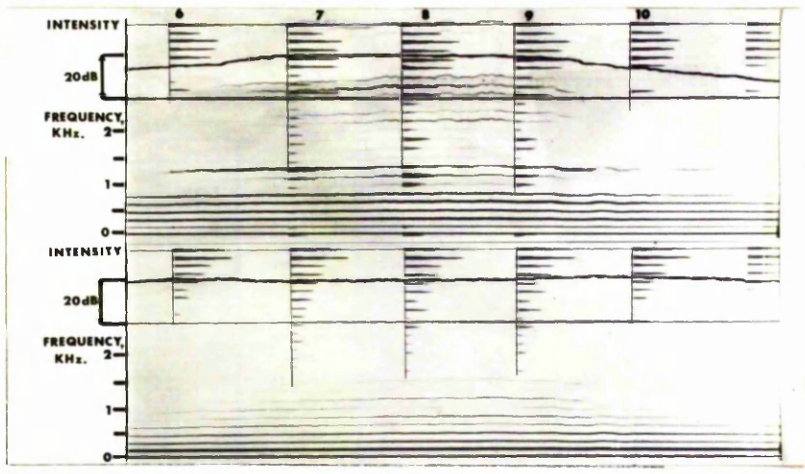
(B) SOUND LEVEL v FLOW RATE GRAPH FOR THE ABOVE PHONATION.

The sonagraph displays are shown successively in Figure 6.9, with, in each case, the upper outputs being the frequency analysis of the microphone recording, and the lower being those of the Lx signal. It was evident from the displays that increases in the sound level of the voice signal were largely caused by differential increases in the intensities of the higher harmonic components, e.g. comparing sections 1 and 2, the 2nd harmonic increased in intensity from 32 to 34dB, a change of 2dB, whereas the 8th harmonic increased by 13dB, and the 17th by at least 20dB. Similar changes were observed by comparing other sections, e.g. 6 and 7, 13 and 14 etc. The same trend was evident in the equivalent spectra of the Lx signal, i.e. an increase in the sound level of the voice output corresponded to an increase in the intensities of the higher harmonic components in the Lx spectra, and could thus be attributed to changes in the vibratory pattern of the vocal folds. The overall intensity of the Lx signal remained fairly constant throughout the phonation, apart from a decrease near Section 12, when the voice sound level decreased to about 73dB.

The differential changes in the intensities of the higher harmonic components of the Lx spectra were measured by changes in the spectral gradient, as described in CHAPTER 5. The results for Lx spectral gradient, voice sound level and air flow rate are shown for each point at which simultaneous readings were available (points 1-15) in Figure 6.10. The relationship between sound level and air flow rate has already been discussed. However, there was also a marked relationship between the Lx spectral gradient and the voice sound level, such that an increase in the voice sound level corresponded to an increase in the Lx spectral gradient (i.e. becoming less negative). Regression analysis showed that a linear relationship of the form

$$SL = B+C.\ (\text{Lx gradient}), \qquad (6.3)$$


20dB
FREQUENCY, KHz.
2
1
0
INTENSITY
20dB
FREQUENCY, KHz.
2
1



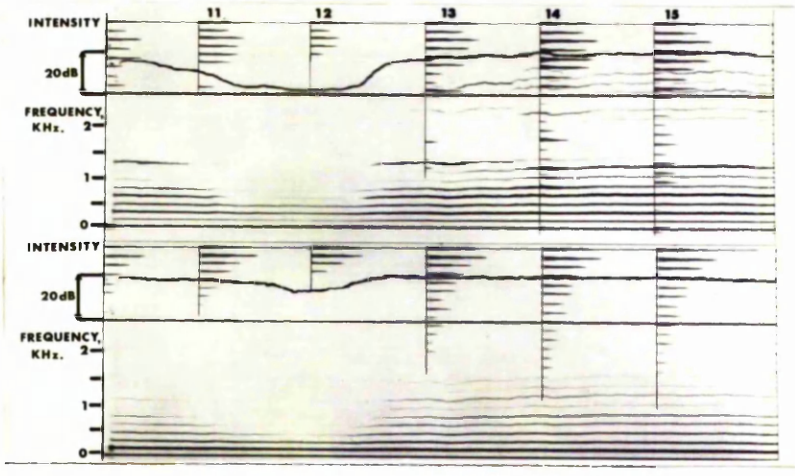


FIG.6.9 SONAGRAPH OUTPUTS FOR PHONATION SHOWN IN FIGURE 6.8, SHOWING SIMULTANEOUS FREQUENCY ANALYSIS OF VOICE OUTPUT AND Lx SIGNAL.

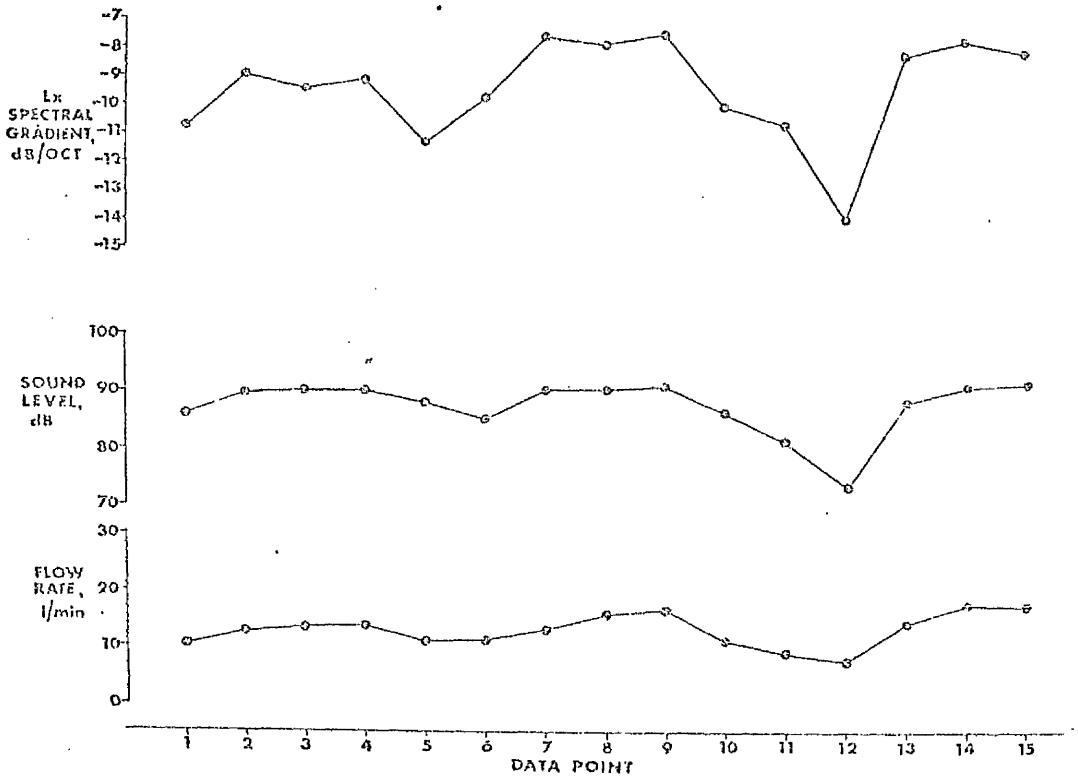


FIGURE 6.10.

VARIATION OF Lx SPECTRAL GRADIENT, SOUND LEVEL AND FLOW RATE FOR PHONATION SHOWN IN FIG.6.8.

was a satisfactory representation of the data, with a correlation coefficient, $r = 0.872$ ($p < 0.001$). The relationship between sound level and air flow rate, and sound level and Lx spectral gradient in turn implied that there was a direct relationship between Lx spectral gradient, or some measure of the vocal fold vibratory pattern, and the air flow rate.

A similar phonation is illustrated in Figure 6.11. The phonation lasted for 21.0secs. and had a fundamental frequency of 204Hz, and the SL v FR locus had parameter values of $A = 100.9$dB, $\alpha = 0.158$, $k = 0.499$, $r = 0.972$ ($p < 0.001$). Sonagraph outputs were also available for this phonation, with section displays being made at points 1-15, and these are shown in Figure 6.12. The general results were the same as those reported for the previous case, with the intensity of the Lx signal remaining fairly constant throughout the phonation, apart from at the very lowest voice intensity (about section 11), and an increase in the voice intensity corresponding to a differential increase in the intensities of the higher harmonic components in both the voice and Lx spectra. This was again reflected in the spectral gradient of the Lx waveform, and the Lx spectral gradient, voice sound level and air flow rate results obtained simultaneously for the 15 data points are plotted in Figure 6.13. Regression analysis indicated that in this case, too, a linear relationship existed between voice sound level and Lx spectral gradient, with a correlation coefficient,

$$r = 0.809 \quad (p < 0.001).$$

Figure 6.14 illustrates a phonation which was not initiated on the equilibrium locus in terms of the sound level v flow rate relationship, this locus being reached about data point 4 (arrowed). Analysis showed that subsequent to this point, there again existed a linear relationship between sound level and spectral gradient ($r = 0.714$, $p < 0.01$), but that prior to reaching the locus, the sound level did not follow the Lx spectral gradient. Thus the initial phase of this

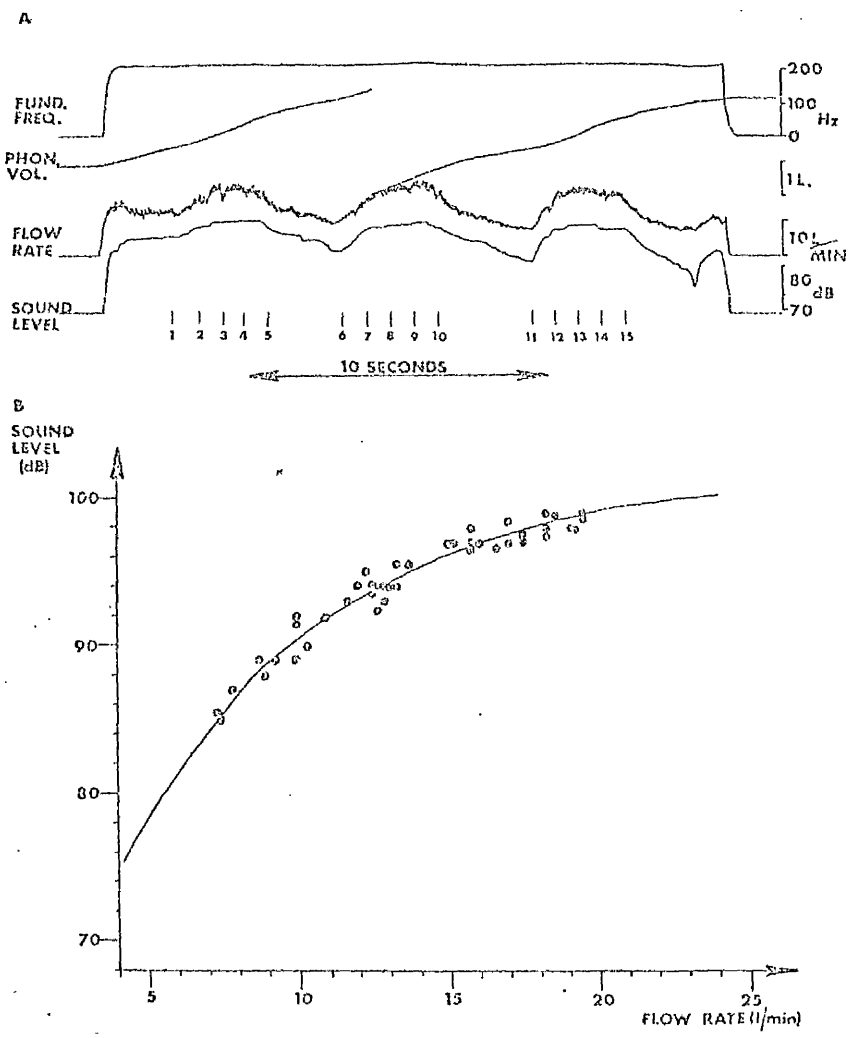
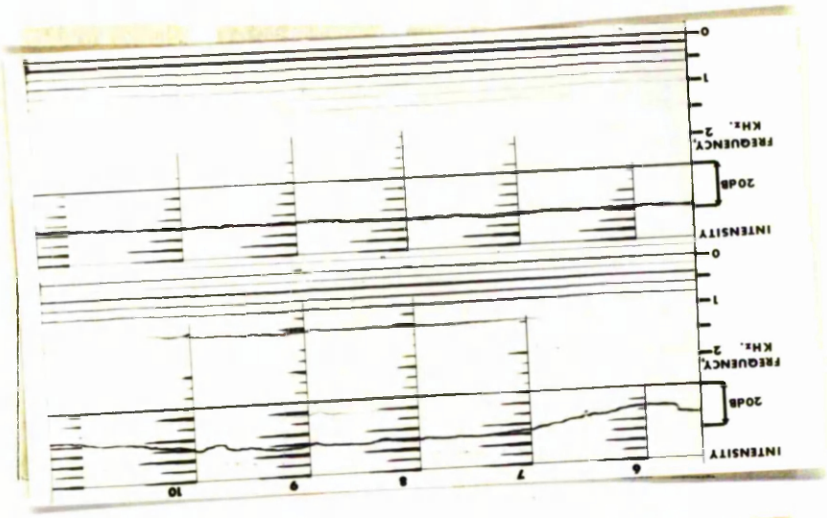
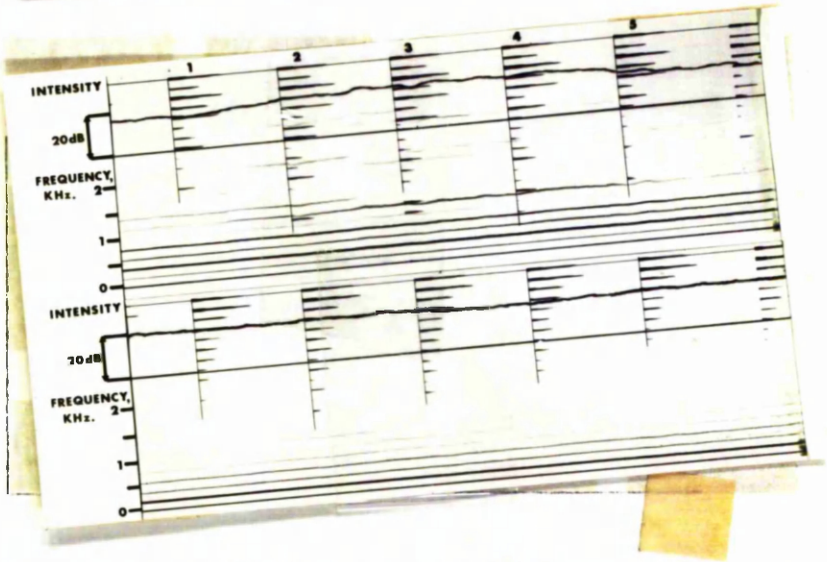


FIGURE 6.11.

(A) SIMULTANEOUS RECORDINGS MADE DURING A NON-STEADY PHONATION.

(B) GRAPH OF SOUND LEVEL v AIR FLOW RATE FOR THE SAME PHONATION.





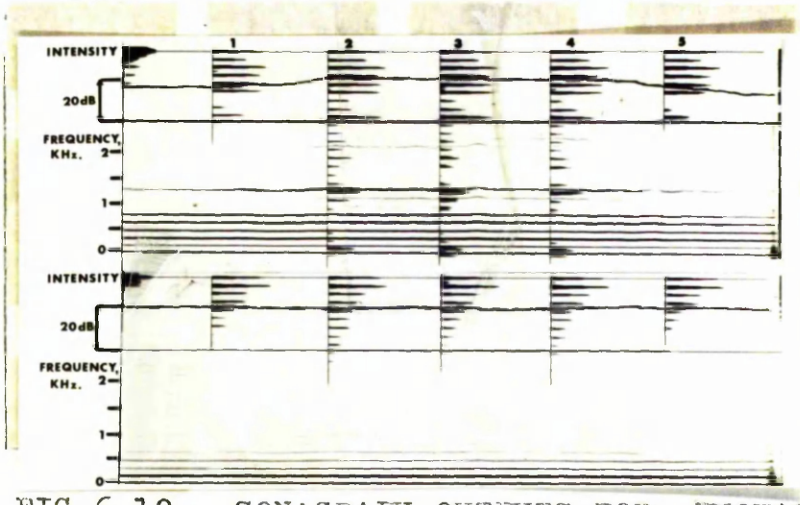


FIG.6.12 SONAGRAPH OUTPUTS FOR PHONATION SHOWN IN FIGURE 6.11 SHOWING SIMULTANEOUS FREQUENCY ANALYSIS OF VOICE OUTPUT AND Lx SIGNAL.

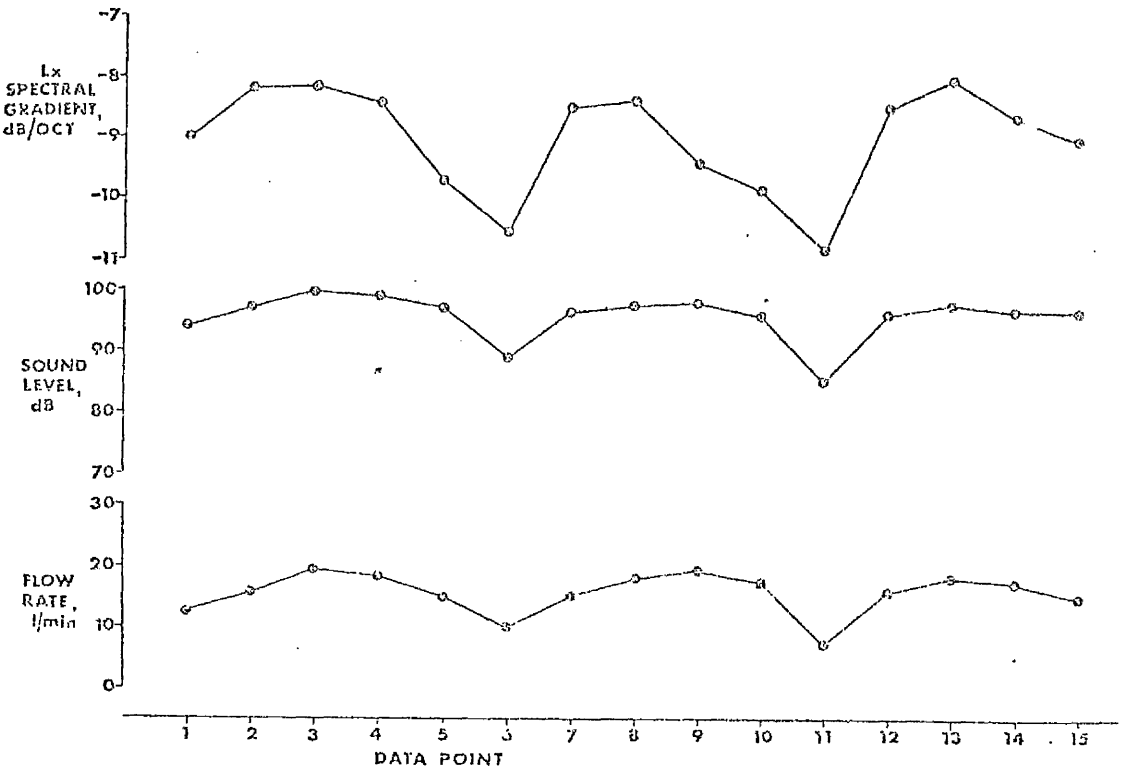


FIGURE 6.13.

VARIATION OF Lx SPECTRAL GRADIENTS, SOUND LEVEL AND AIR FLOW RATE FOR PHONATION SHOWN IN FIG.6.11.

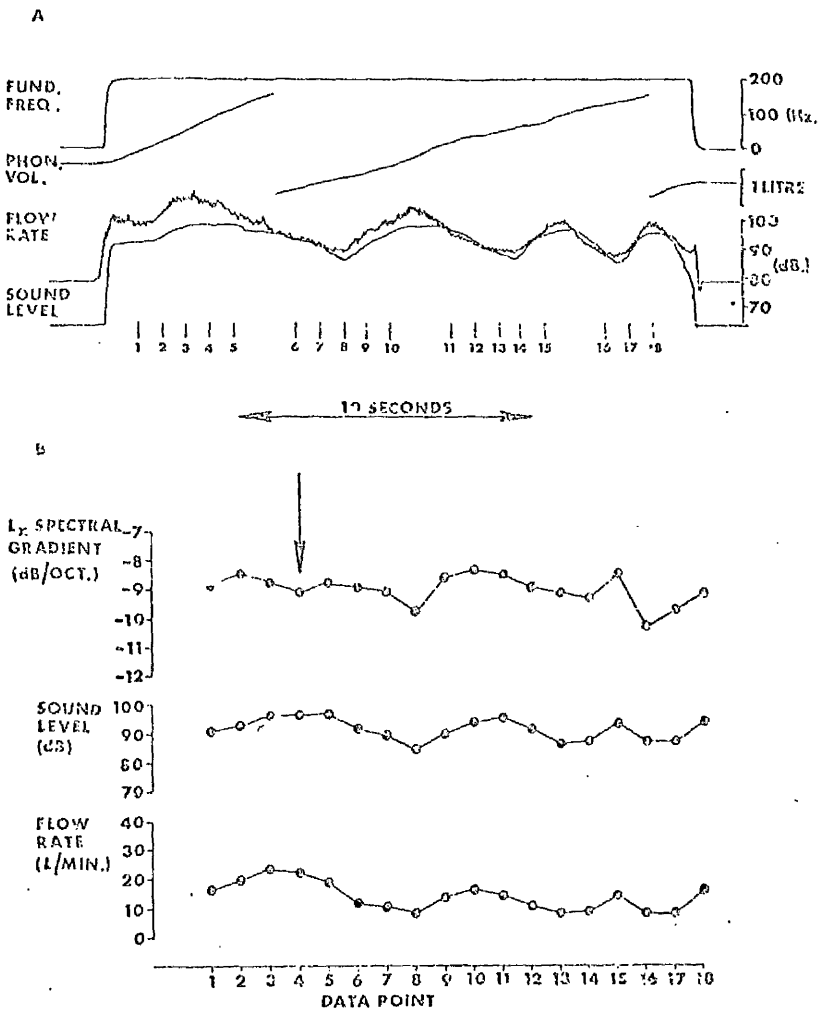


FIGURE 6.14.

(A) SIMULTANEOUS RECORDINGS MADE DURING A NON-STEADY PHONATION WHICH WAS NOT INITIATED ON THE EQUILIBRIUM LOCUS.

(B) VARIATION OF Lx SPECTRAL GRADIENT, SOUND LEVEL AND AIR FLOW RATE FOR THE SAME PHONATION.

phonation represented a period of vocal fold adjustment, and that if there was an explicit relationship between the parameters used to define the vibratory pattern, and either sound level or air flow rate, it was not as simple as when the equilibrium locus had been attained.

In general, in the phonations investigated, there was a direct correlation between the Lx spectral gradient and the sound level of the voice, and the results are tabulated in TABLE 6.2. The Lx waveform again seemed to behave in a fashion similar to that of the laryngeal source function in that changes in the harmonic components in the Lx spectra were directly reflected in the voice spectra. Eliminating SL from equations 6.1 and 6.3 gives :

$$A - Ak\, e^{-\alpha . FR} = B + C.M \text{ , where M=Lx gradient}$$

$$\text{i.e.} \quad M = P - Q\, e^{-\alpha . FR} \qquad \text{——————(6.4)}$$

$$\text{and} \quad FR = -\frac{1}{\alpha} \cdot \ln \frac{(P-M)}{Q} \qquad \text{——————(6.5)}$$

$$\text{where} \quad P = \left(\frac{A-B}{C}\right) \quad \text{and} \quad Q = \left(\frac{Ak}{C}\right)$$

The values for P and Q are given in TABLE 6.3 for some of the phonations recorded, and examples of curves derived from equation (6.4) are plotted in Figure 6.15. The general form of these curves indicated that at low flow rates, an increase in the flow rate produced a relatively large increase in the Lx spectral gradient, (i.e. becoming less negative) caused by the more rapid closing of the folds due to increase in the Bernoulli force. However, at high flow rates a further increase in air flow produced little or no change in the spectral gradient, indicating that there existed a lower limit for the closing time, i.e. the vocal folds could not close at an infinitely fast speed, and so the increase in air flow must be accompanied by a decrease in the opening

PHON.	FUNDA-MENTAL FREQ.	B	C	CORR. COEFF. r	SIGNIF. LEVEL.
2	167	109.4	2.33	-0.872	< 0.001
4	176	146.8	5.95	-0.860	< 0.001
-	190	133.4	4.48	-0.918	< 0.001
10	200	132.7	4.62	-0.714	< 0.01
11	200	115.0	2.20	-0.671	< 0.01
-	204	126.1	3.56	-0.531	< 0.05
14	204	127.3	3.49	-0.809	< 0.001
18	216	113.1	2.49	-0.700	< 0.01
24	250	127.6	3.85	-0.776	< 0.01
26	257	138.1	4.97	-0.790	< 0.01

TABLE 6.2

DATA FROM REGRESSION ANALYSIS FOR THE EQUATION OF THE FORM :

$$SL = B + C. \text{ (Lx gradient)}$$

PHON.	FUND.FREQ.	P	Q	α
2	167	-7.60	45.54	0.271
4	176	-7.87	31.56	0.260
10	200	-7.47	10.01	0.166
11	200	-5.59	17.04	0.365
14	204	-7.56	14.43	0.158
18	216	-6.14	12.25	0.312
24	250	-8.44	7.48	0.188
26	257	-6.94	7.05	0.054

TABLE 6.3

CALCULATED VALUES FOR P and Q, where

$$P = \left(\frac{A - B}{C}\right) \text{ and } Q = \left(\frac{Ak}{C}\right).$$

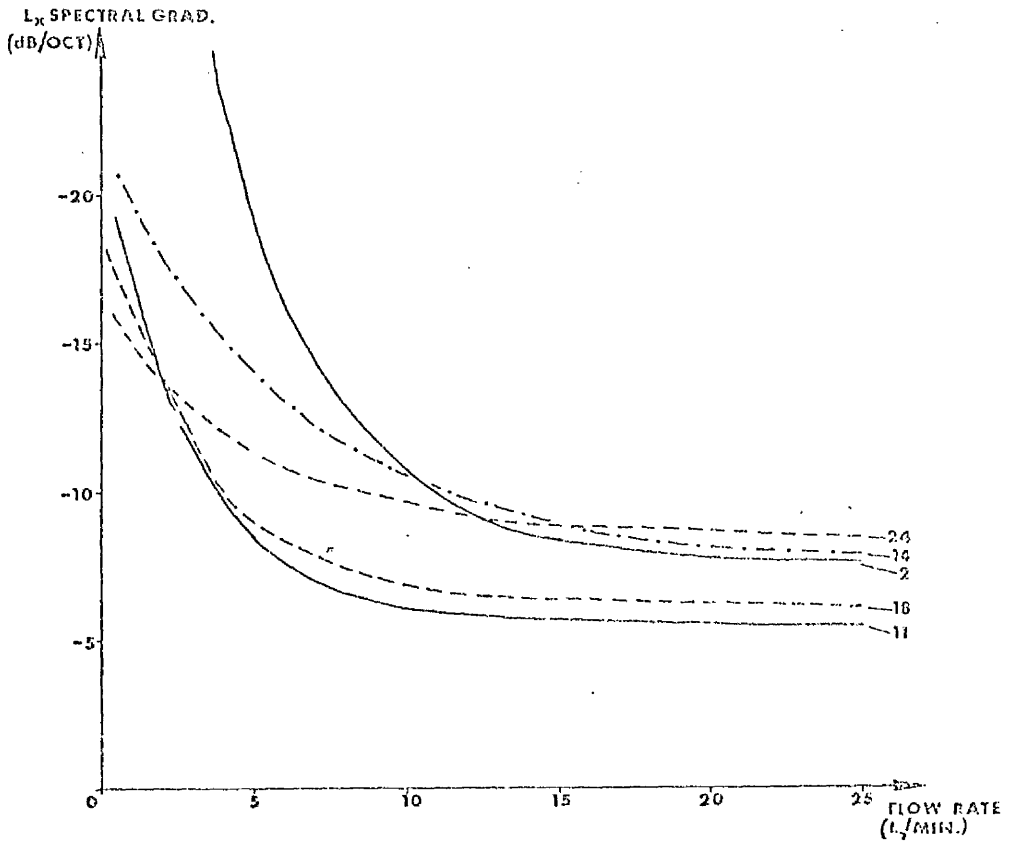


FIGURE 6.15.

EXAMPLES OF GRAPHS OF FORMS

$$L_x \text{ grad.} = P - Q e^{-\alpha . FR}$$

OBTAINED FROM EXPERIMENTAL PHONATIONS.

time and consequent increase in the open time of the vocal fold vibratory pattern. Thus the increase in air flow rate was achieved by a relative decrease in the glottal impedance. The existance of a lower limit for the closing time suggested that a similar limit would exist for the opening time, these being dependent upon the mass of the folds. Since the open time would still have to be long enough to allow sufficient air flow, an upper limit was placed on the fundamental frequency which could be sustained under a particular vocal fold adjustment. Thus to reach high fundamental frequencies, the vibrating portions of the vocal folds must be adjusted to be less massive than that for lower frequencies.

Although some of the direct parameters used to define the vibratory pattern of the vocal folds could be inferred from the Lx spectral gradient (CHAPTER 5), a limited number of phonations was recorded to investigate directly the relationship between voice sound level and the closing time of the vocal fold vibration. During these phonations, the sound level output obtained from the analogue read-out device was recorded on one channel of an FM tape recorder, while the simultaneous Lx output was recorded on the other channel. This tape was then replayed via the U-V recording oscilloscope to obtain the Lx waveform and the voice sound level at several parts of the phonation. The ratio (closing time/period%) was measured from the Lx traces, and these values, together with the voice sound level are plotted for corresponding points in each phonation in Figure 6.16. In each case, a decrease in the relative closing time (resulting in more intense higher harmonic components) corresponded to an increase in the voice sound level, and linear regression analysis was carried out between the two variables. There was a significant linear relationship ($p < 0.001$) for each phonation, and the results are given in TABLE 6.4.

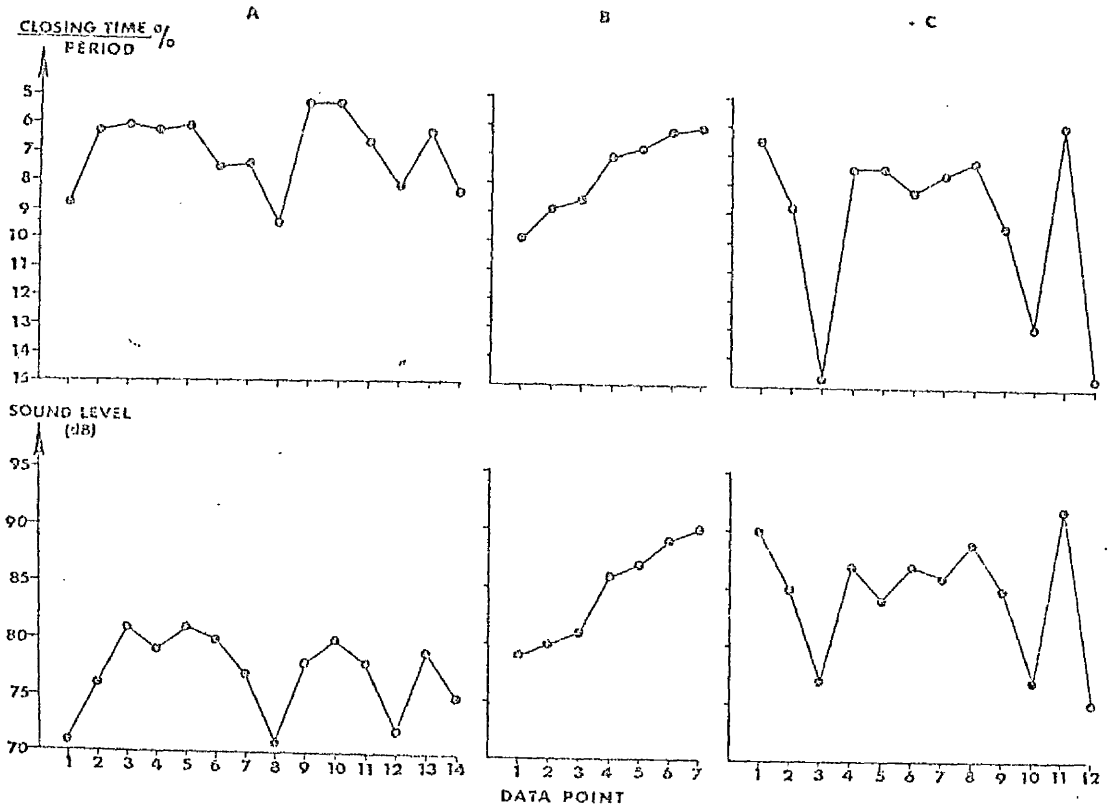


FIGURE 6.16.

GRAPHS SHOWING THE VARIATION OF RELATIVE CLOSING TIME OF THE VOCAL FOLD VIBRATORY PATTERN AND ALSO THE VOICE SOUND LEVEL FOR THREE DIFFERENT NON-STEADY PHONATIONS.

PHONATION	K_1	K_2	r	LEVEL OF SIGNIFICANCE
A	92.4	-2.20	-0.817	< 0.001
B	99.9	-1.67	-0.959	< 0.001
C	107.6	-3.01	-0.987	< 0.001

TABLE 6.4

DATA FROM LINEAR REGRESSION ANALYSIS FOR THE EQUATION OF THE FORM :

$$SL = K_1 + K_2 \cdot \left(\frac{CL.T}{T} \cdot \% \right)$$

The variation in the shape of the Lx waveform for phonation C (Figure 6.16) is shown in Figure 6.17, with the voice sound levels being given below each output. It was observed that in the waveforms corresponding to the higher sound levels, the closing phases of the vibratory pattern were steeper than those in the lower sound level outputs, e.g. compare the waveforms A, B and C. However, it also appeared that in the high sound level outputs, the peak of the Lx signal became more rounded, and there was also a tendency to shorten the opening phase and thus lengthen the period of no contact, agreeing with results obtained by other methods by Timcke et al.(1958) and Flanagan (1958), and also those inferred from the previous experiment.

6.6 DISCUSSION

The data obtained from previous recordings agreed with those reported by Isshiki (1965) and Perkins and Yanagihara (1968) viz. that, in general, there was an increase in air flow rate with both fundamental frequency and sound level, these data being recorded during steady sustained phonations. The results reported in this chapter for non-steady sustained phonations strictly applied only for the test subject, and then only for phonations near and above his normal fundamental frequency in chest register (about 150Hz). However, it was found from phonations in which the fundamental frequency was maintained at a constant value, while the sound level was varied, that there existed a preferred area or locus in the SL v FR space, along which each phonation progressed. The loci for the phonations did not coincide, due to the many possible adjustments of respiratory and laryngeal musculature, but they all had the same form of equation :

$$SL = A\,(1 - ke^{-\alpha . FR}) .$$

Phonations lying to the right of this locus in the

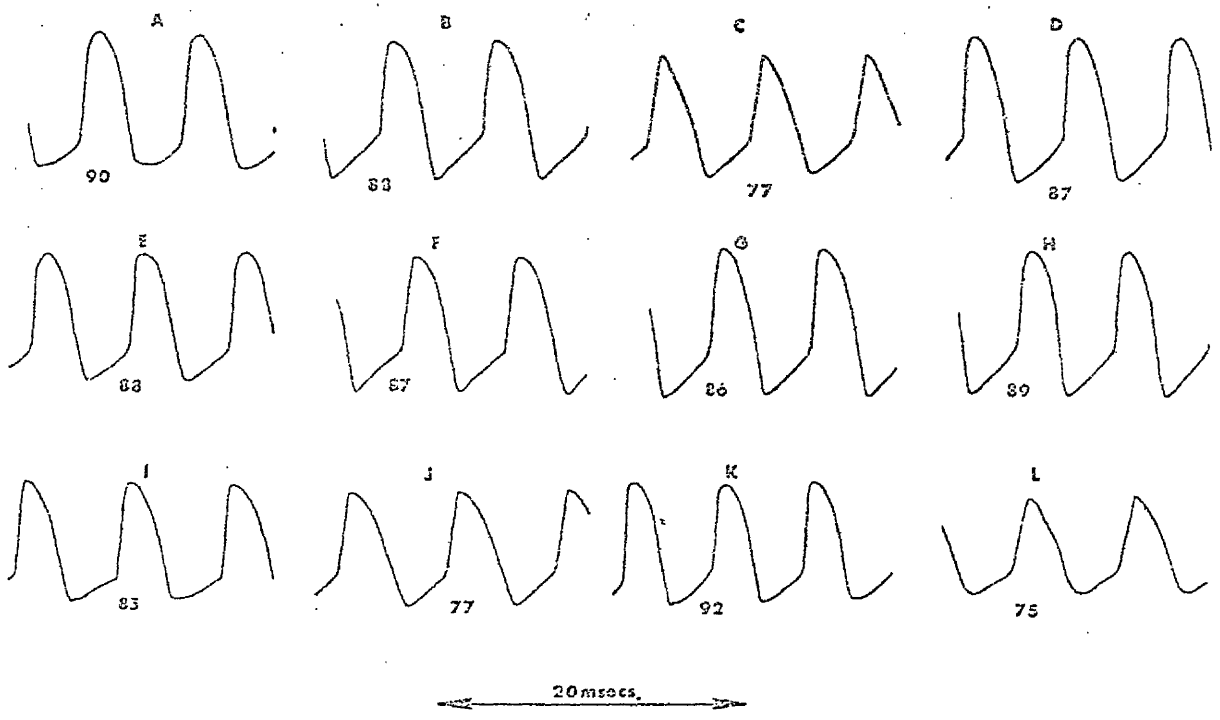


FIGURE 6.17.

EXAMPLES OF Lx WAVEFORMS OBTAINED DURING THE PHONATIONS SHOWN IN FIGURE 6.15.C.

SL v FR space, i.e. corresponding to greater air flow rates for a given voice sound level, did exist, although the tendency during the non-steady phonations was always to move towards the locus. Thus the locus represented a condition of maximum efficiency in terms of sound level and air flow rate. Less efficient states could be maintained during steady sustained phonation by suitable muscular adjustment, and accounted for the earlier data in which there was no relationship between sound level and flow rate, the data points representing a random selection of possible adjustments. States corresponding to higher sound levels than the locus for a given flow rate did not exist in phonations in which the fundamental frequency remained constant, but appeared to be permissible when the fundamental frequency was rapidly varying. This point could not be investigated in a similar manner, as the test subject could not maintain a particular sound level while varying his fundamental frequency.

The variation in voice sound level was also related to the vibratory pattern of the vocal folds in terms, firstly, of the Lx spectral gradient, and then in terms of the relative closing time of the vibratory cycle. A highly significant linear relationship was found between the voice sound level and the Lx spectral gradient, although it was likely that a much more complex relationship existed. In general, a decrease in the spectral gradient, corresponding to a relative increase in the intensities of the higher harmonic components, resulted in a similar trend in the voice spectra and an increase in the overall voice sound level. In this respect, the Lx acted in a manner similar to the laryngeal source function. The relationship between voice sound level and the relative closing time of the vocal fold vibration was also found to be linear, with the shorter closing times corresponding to higher voice sound levels. Such a

relationship was qualitatively expected, since it had been shown that the closing time corresponded to the period of maximum excitation of the vocal tract. However, other changes also took place in the Lx waveform. At increased sound level, the closing time became shorter, but so too did the opening time, thus giving a relatively longer period of no contact. Thus the mean glottal impedance could either be increased or decreased by varying the proportions of the different phases of the vocal fold vibration. The only satisfactory method of investigating such changes was to measure the subglottal pressure and flow rate, cycle by cycle. However, the techniques of measuring subglottic pressure involve either the insertion of a pressure sensing balloon into the oesophagus, and after the correction for lung volume and thoracic pressure effects, the laryngeal pressure can be estimated, or the insertion of a hypodermic needle between the cartilage rings and into the subglottal space of the larynx, and measuring the subglottic pressure directly. Both techniques are highly invasive to the subject and were beyond the scope of this thesis. Thus no correlations could be made between the present data and subglottic pressure. However, the present results did provide valuable information about the relationships governing the control of sound level by both air flow rate and the vibration of the vocal folds.

6.7 SUMMARY

The relationships governing the control of sound level were discussed in this chapter. When this was investigated by using random steady sustained phonations, there was found to be no significant relationship between the voice sound level and the air flow rate. However, when phonations in which the fundamental frequency was maintained constant while the sound level was allowed to vary, were investigated, a limiting condition was found, which could be considered to consist of states of maximum efficiency for

these variables. Although states of lower efficiency were permitted, states corresponding to a greater sound level at a particular air flow rate than that of this locus, were not found in this type of phonation. The equations of these loci all took the form :

$$SL = A(1 - ke^{-\alpha . FR}) .$$

The relationship between voice sound level and the vibratory pattern of the vocal folds was also investigated. Linear relationships were found between the voice sound level and the Lx spectral gradient, and also between the voice sound level and the relative closing time of the vocal fold vibration. Both of these relationships reflected that a relative increase in the intensities of the higher harmonic components in the Lx spectra resulted in a similar trend in the voice output spectra, and also an increase in voice sound level. Thus the Lx was again acting in a fashion similar to the laryngeal source function.

CHAPTER 7.

A STUDY OF THE ELECTROMYOGRAPHIC ACTIVITY OF THE LIP MUSCULATURE.

CHAPTER 7. A STUDY OF THE ELECTROMYOGRAPHIC ACTIVITY OF THE LIP MUSCULATURE.

7.1 INTRODUCTION.

In the previous chapters of this thesis, the role of air flow rate and the vibratory pattern of the vocal folds has been discussed and related to the final voice output in terms of frequency components and also overall sound intensity. The phonations investigated consisted entirely of sustained vowel sounds, although the experiments described in CHAPTER 6 utilized non-steady phonations with a view to quantifying the relationships between air flow rate, voice sound level and the vocal fold vibratory pattern. However, continuous speech consists of rapid transitions between vowels and consonants, involving rapid changes in the adjustment of the musculature of the larynx and articulators, thus necessitating a fine degree of neuro-muscular control. The purpose of the experiments described in this chapter was to investigate some of the aspects of muscle activity during the production of a limited series of simple speech utterances. The technique involved measuring the electromyographic activity at particular sites using non-invasive, surface electrodes, and interpreting the results with regard to the muscle structure at those sites.

Interest in the electromyographic (EMG) activity of the musculature involved in speech production originated from the phonetic theory that each phoneme (smallest indivisible sound in a language) had some corresponding invariant physiological characteristic. This invariant characteristic was patently not the waveform of the acoustic signal, which was extremely variable, and it was suggested that each phoneme resulted from a particular series of neuro-muscular events. The most accessible muscles directly involved in speech production are those of the face and, in particular, those surrounding the lips. Several studies had been carried

out to investigate the EMG activities of these muscles, associated with specific speech gestures, but the results obtained were conflicting both in terms of the magnitudes of the activities measured, and also in their general forms. However, the techniques employed in these studies were investigated, and it was found that different electrode sites and electrode configurations (i.e. either monopolar or bipolar) had been utilized. Fromkin (1966) used an electrode situated on the medial line of the lower lip, whereas Tatham and Morton (1969) had used a bipolar pair of electrodes placed on the upper lip, but off midline in one investigation, but had used a monopolar electrode situated near midline and referenced to an electrode placed on the nose in another series of experiments (Tatham and Morton, 1973). Lubker and Parris (1970) had used a bipolar electrode configuration on the lower lip, and to the left of midline, while leanderson et al (1967,1971,1972) utilized arrays of needle electrodes to record simultaneously from several specific muscle sites.

In monopolar recordings, the absolute EMG activity at one site is measured with reference to the activity at a site which is not involved in the process under test. In bipolar recordings, the differential activity between two electrode sites is measured, with both electrodes recording from muscles involved in the test process. Thus a bipolar electrode pair measures the vector sume of the monopolar outputs of each electrode.

Clinical applications resulting from this type of study include the possible diagnosis of organic diseases of the nervous system which give rise to alterations in the pattern of EMG activity (Leanderson et al 1970). In another application, the EMG output is employed to produce an audio feedback signal in the treatment of stuttering, and the patient is encouraged to reduce the level of EMG activity (Guitar, 1975).

The aim of the present investigation was to record the EMG activity associated with a limited number of speech utterances, using different electrode sites and

configurations, and to interpret the results in terms of the underlying muscle structure. Surface electrodes were used throughout, as the invasive technique of using needle electrodes was beyond the scope of this thesis. However, the results obtained were compared with those of other investigators who had employed needle electrode recording.

7.2. RECORDING TECHNIQUE

The EMG activity was obtained from the sites under investigation by the use of standard silver/silver chloride cup electrodes, each having an active diameter of 5mm. The skin of the subject was suitably prepared, and the electrodes were attached using adhesive discs and filled with standard electrode jelly. The active electrodes were situated at the required sites on the vermilion border. In the case of monopolar recordings, the reference electrode was situated on the boney point of the mastoid process, this being remote from the muscles involved in speech production, while the ground electrode was always place on the forehead.

Four biological amplifiers having identical frequency characteristics, namely flat in the range 1.5 - 700Hz, were used to amplify the EMG signals, which were then recorded on four channels of an FM tape recorder (Ampex SP300). The tape recorder was run at a tape speed of 3.75ips (9.5 cps), and thus had a flat frequency response in the range DC.-625Hz. The channel gains were set equal by using a calibration signal, and good electrode contact was ensured by monitoring each channel in turn using an oscilloscope. The voice output of the subject was obtained via a throat contact microphone and recorded on another channel of the tape recorder, used in the "direct" record mode.

Since the musculature under investigation was that surrounding the lips (principally the muscle orbicularis oris, m. orb. oris), the speech utterances chosen were consonant - vowel - consonant (CVC) syllables containing the plosives /p/ and /b/. The main syllable used was /pIp/, while the syllables /pʌp/, /pIb/, /bIb/, /bIp/

and /bʌb/ were also recorded for comparative purposes. The subject was required to repeat the required CVC utterance at 5 second intervals, using his normal voice.

7.3. ANALYSIS TECHNIQUE

The analysis system is shown schematically in Figure 7.1. In order to obtain a measure of the magnitude of the EMG activity, the EMG signal was full-wave rectified and integrated, and in order to remove the random fluctuations present between even consecutive utterances of the same CVC syllable, a process of signal averaging was employed, with the averager trigger signal being derived from the voice output.

The voice signal was full-wave rectified and integrated, the time constant of integration being 50msec., and this signal was used as the input to a Schmitt trigger circuit. This circuit produced square output pulses which were used to provide the trigger for the signal averager. The Schmitt trigger circuit also incorporated a variable dead-time facility during which it could not be retriggered. The dead-time was set to 4 sec., and this served as a noise rejection device, preventing the averager from being triggered by spurious noise between CVC utterances.

Since the EMG activity was initiated prior to the onset of the audio signal, and hence the trigger pulse, it was necessary to delay the EMG signal before presenting it to the averager. This was achieved in the playback mode by re-recording the EMG signal on another FM tape recorder (based on a Ferrograph tape deck, and having a flat frequency response in the range DC-625 Hz.) and using the output of the replay head of this recorder as the input to the integrator. A delay of 0.45sec. was introduced when this recorder was run at a tape speed of 3.75ips. The EMG signal was then rectified and integrated (the integration time constant being 50msec.) and presented to the signal averager

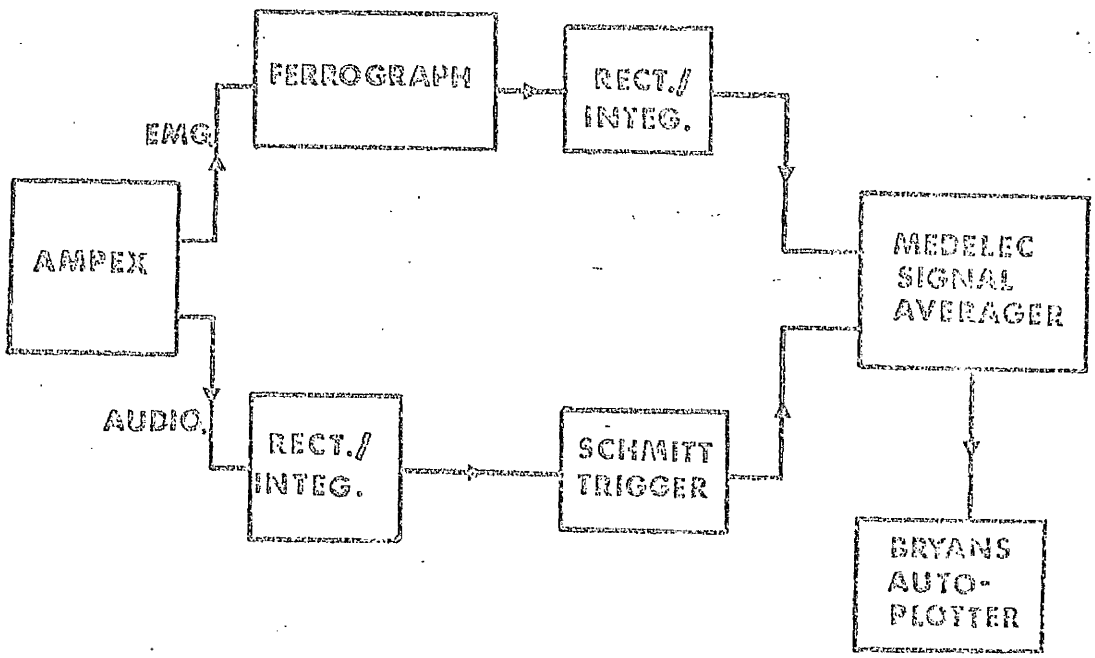


FIGURE 7.1.

SCHEMATIC DIAGRAM OF EMG ANALYSIS SYSTEM.

(Medelec AVM 62). Figure 7.2 shows four examples of the unprocessed EMG signals and integrated EMG and audio signals for the same CVC utterance, prior to averaging. The averager had a 200 point store and a resolution of 5msec. with a 1sec. averaging time. During analysis, the Medelec system was used in the "superposition" mode, so that the outputs obtained from each utterance of the same CVC syllable were superimposed and gave an indication of the variability present. The final averager output was recorded using a Bryans Auto Plotter (22000 series), which allowed averaged results obtained from different electrode sites to be superimposed.

In a preliminary series of experiments, designed to elicit a suitable sample size, recordings were obtained during which 80 utterances were made. Averaged outputs were obtained using different samples of 30 tokens, and Figure 7.3 shows the results obtained for different samples from one such recording. There were no significant differences between such averages and those of larger sample size, and so a sample size of 30 was chosen as producing representative results. During the subsequent recording sessions, the subject was required to perform 40 utterances of the required syllable, with the initial and final 5 being discarded to avoid possible artifacts.

In order to obtain a measure of the variability present in the data which provided a single averaged output, several sets of rectified and integrated signals were displayed individually by using the Mingograf ink jet recorder (as illustrated in Figure 7.2), and the amplitudes of the main peaks were measured. The coefficients of variations (std.dev./mean.%) were calculated, and found to be in the range 13.3 - 22.1%, with a mean value of 18.3%

The experimental subjects were two normal speaking male subjects, aged 26 years. Although the main features of the results were similar for both subjects, there

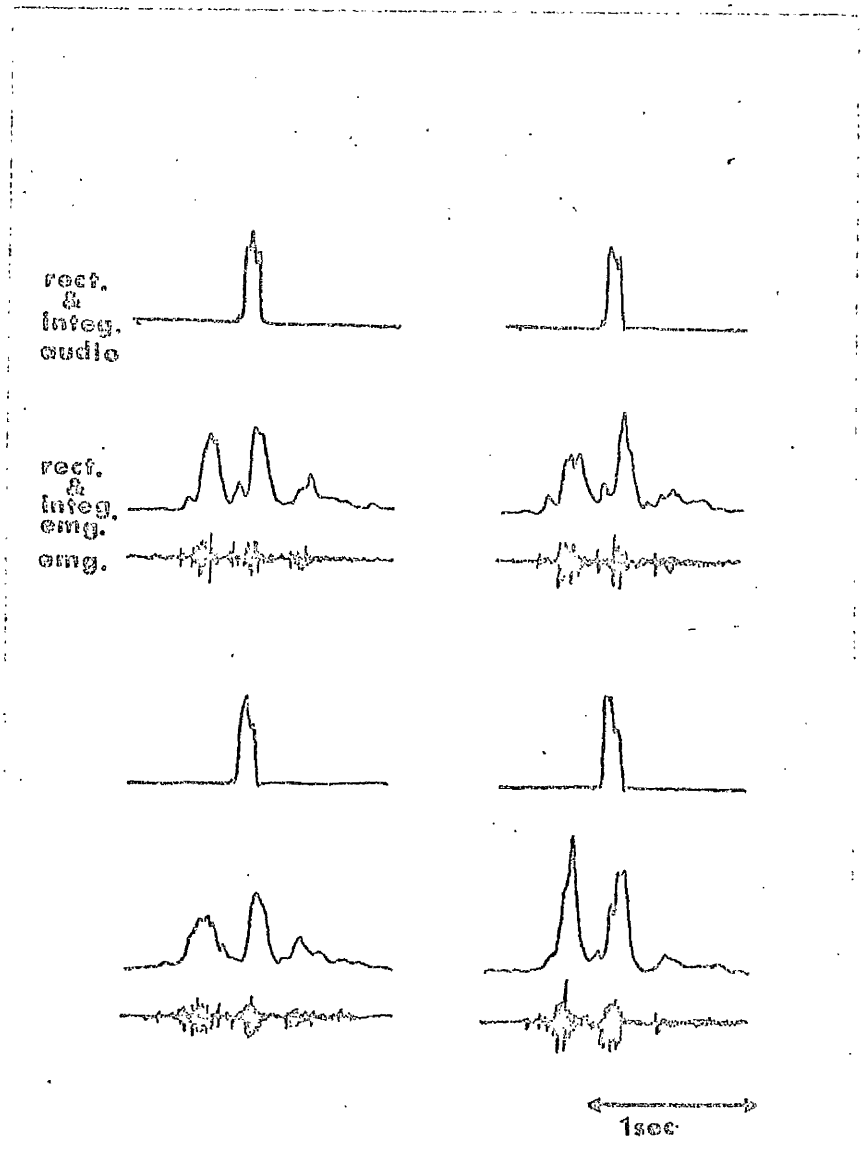


FIGURE 7.2.

FOUR EXAMPLES OF UNPROCESSED EMG SIGNAL TOGETHER WITH RECTIFIED AND INTEGRATED EMG AND AUDIO SIGNALS. (UTTERANCE /pIp/).

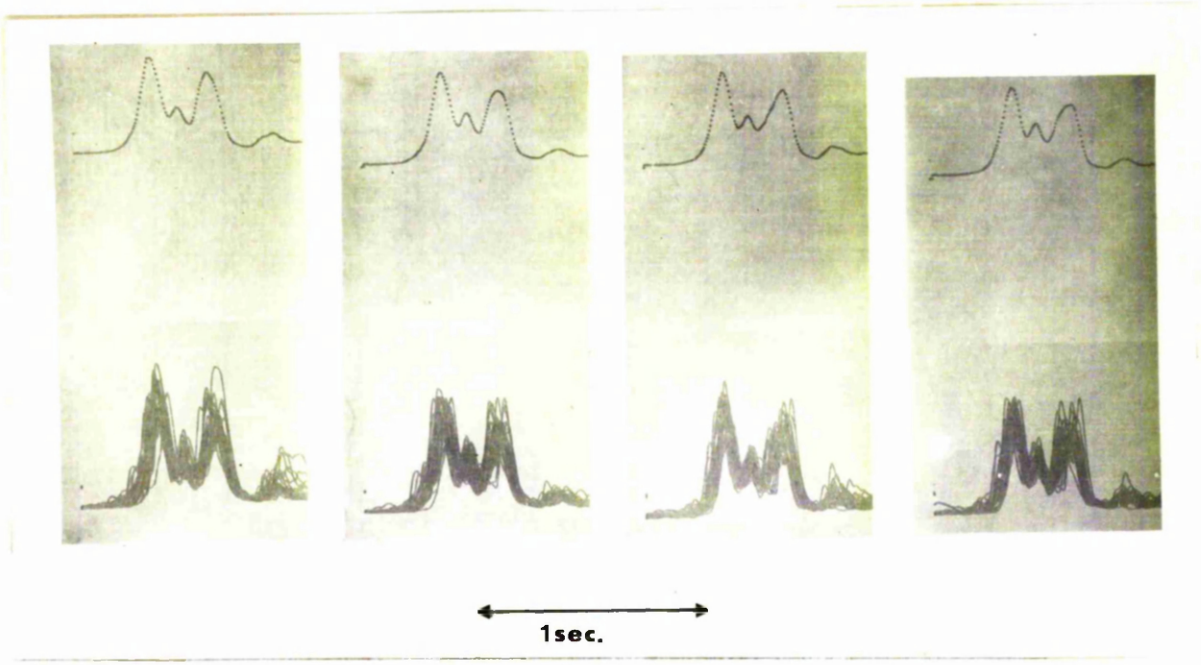


FIGURE 7.3

EXAMPLES OF SUPERPOSITION AND AVERAGED OUTPUTS OF DIFFERENT SAMPLES OF 30 TOKENS OF UTTERANCE /pIp/.

were some significant differences, and so the results for each subject are presented separately in each section.

In the initial parts of the study, the aims were to compare the activities measured from left and right electrode sites, i.e. to investigate if there was lateral symmetry, and also to compare the activities measured from superior and inferior electrode sites. These recordings were made with the electrodes connected firstly in the bipolar configuration (SECTION 7.4) and then in the monopolar one (SECTION 7.5), and the results obtained using the different configurations were also compared (SECTION 7.6). In the later parts of the study, the EMG activity associated with the initial and final consonant was investigated, and the muscle groups giving rise to the activity were identified.

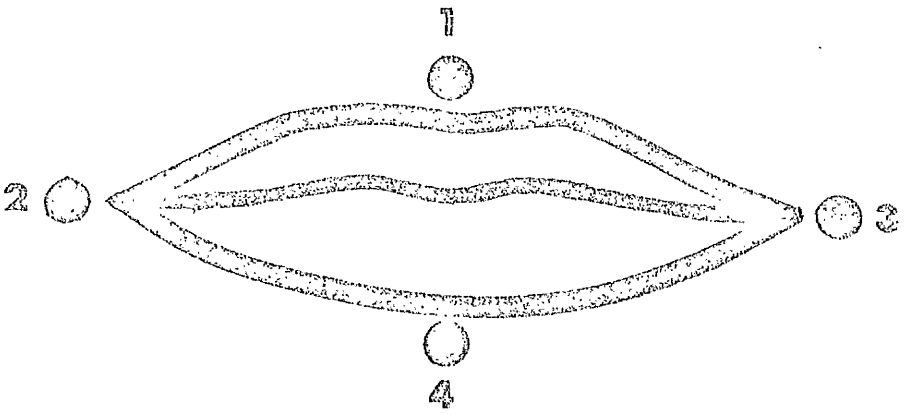
7.4. BIPOLAR RECORDINGS DURING THE PRODUCTION OF THE SYLLABLE /pIp/.

The surface electrodes were placed as close as possible to the vermilion border at the sites shown in Figure 7.4, and connected in the bipolar pairs 1-2, 1-3, 2-4, and 3-4. The test syllable was /pIp/, and the EMG activity was recorded and analysed as detailed in SECTIONS 7.2 and 7.3.

A typical averaged result obtained from subject A is shown in Figure 7.5. There were two main peaks, labelled P1 and P2, in the output obtained from the superior electrode pairs, and also a much smaller peak, P3, lying between them. (It was later demonstrated that the peaks P1 and P2 were directly related to the initial and final consonants in the utterance (SECTION 7.7).) The output from the inferior electrode sites showed the same three peaks, but P1 and P2 were much smaller in magnitude, and P3 was the largest. A peak occurring after P2 was also present and was labelled P4, but the peak which followed P4 by about 150msec. was caused by muscular readjustment well after the speech gesture had been completed.

Left

Right



1 = Superior electrode

2,3 = Lateral electrodes

4 = Inferior electrode

FIGURE 7.4.

POSITION OF ELECTRODES ON VERMILION BORDER.

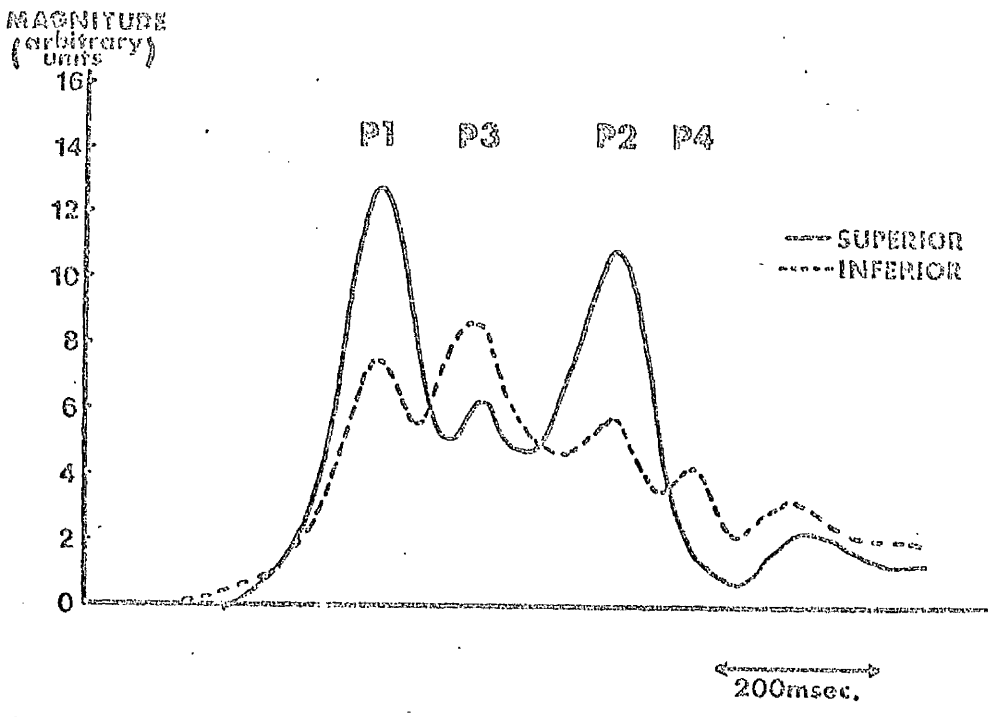


FIGURE 7.5.

EXAMPLE OF TYPICAL RESULTS OBTAINED FROM SUBJECT A WITH BIPOLAR ELECTRODE CONFIGURATION.

Several recording sessions were carried out, and the general forms of the results were as described above. The magnitude of each peak was measured for each electrode pair, and the results obtained for the series of recordings are shown in Figure 7.6. The range of magnitudes for each peak obtained from left side electrode sites was compared to the range for the corresponding peak obtained from right side electrodes, using Student's t-test, and the data are presented in TABLE 7.1A. The data were also tested using the non-parametric Wilcoxon rank test, and the results were similar at the 0.05 level of significance. It was noted that there were no significant differences measured for any peak, and that, as expected, there was a lateral symmetry in the EMG activity.

However, the data were also compared using a paired t-test, and in this case there were significant differences in the magnitudes of the activities attributed to peaks P1 and P2 from both superior and inferior electrode sites (TABLE 7.1B). This indicated that for these peaks there was consistently greater activity measured from the right electrode pairs, although the results from the unpaired test showed that the variation resulting from different recording sessions, and perhaps slightly different electrode sites, nullified any differences in the peak magnitude ranges.

The magnitude of the peaks obtained from the superior electrodes were compared to those measured by the inferior electrodes (TABLE 7.1C), and for all peaks which were common, i.e. P1, P2 and P3, there were statistically significant differences between the ranges.

The general form of the outputs obtained from subject B are shown in Figure 7.7. Similarly to subject A, for the superior electrode sites, P1 and P2 were the dominant features, with P3 being evident, but rarely well enough defined to be measured. There was also a late peak present from the superior sites. The inferior

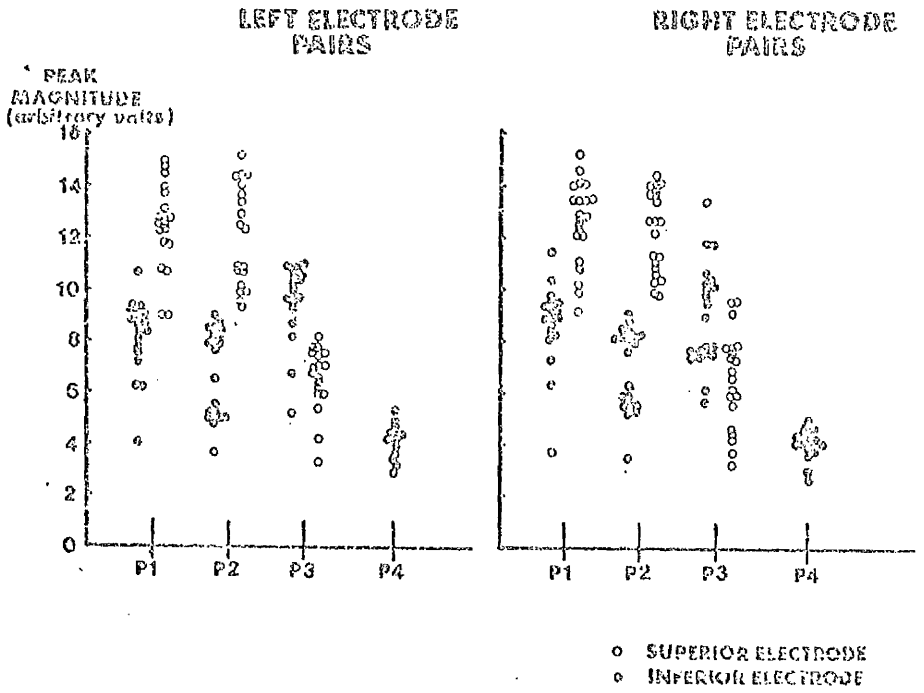


FIGURE 7.6.

DISTRIBUTION OF PEAK MAGNITUDES FOR SUBJECT A.

A.

PEAK	ELECTRODE PAIRS	n	t-VALUE	SIGNIFICANCE LEVEL	WILCOXON COEFF.	SIGNIFICANCE AT 0.05 LEVEL
P1	1-3/1-2	38	0.50	NS	453.5	NS
P1	3-4/2-4	37	1.30	NS	469.0	NS
P2	1-3/1-2	38	0.38	NS	444.0	NS
P2	3-4/2-4	37	0.51	NS	430.0	NS
P3	1-3/1-2	38	0.29	NS	447.5	NS
P3	3-4/2-4	37	0.76	NS	365.5	NS
P4	3-4/2-4	32	0.05	NS	286.5	NS

B.

PEAK	ELECTRODE PAIRS	n	t-VALUE	SIGNIFICANCE LEVEL
P1	1-3/1-2	18	4.23	< 0.001
P1	3-4/2-4	18	4.36	< 0.001
P2	1-3/1-2	18	2.95	< 0.01
P2	3-4/2-4	18	3.75	< 0.001
P3	1-3/1-2	18	0.93	NS
P3	3-4/2-4	18	1.15	NS
P4	3-4/2-4	14	0.39	NS

C.

PEAK	ELECTRODE PAIRS	n	t-VALUE	SIGNIFICANCE LEVEL	WILCOXON COEFF.	SIGNIFICANCE AT 0.05 LEVEL
P1	1-3/3-4	39	7.75	< 0.001	635.5	S
P1	1-2/2-4	36	8.28	< 0.001	558.5	S
P2	1-3/3-4	39	10.55	< 0.001	610.0	S
P2	1-2/2-4	36	8.88	< 0.001	610.0	S
P3	1-3/3-4	39	3.79	< 0.001	321.0	S
P3	1-2/2-4	36	6.68	< 0.001	197.0	S

TABLE 7.1

STATISTICAL COMPARISONS OF PEAK MAGNITUDES FOR SUBJECT A. (BIPOLAR CONFIGURATION)

A - LATERAL COMPARISON, B - LATERAL COMPARISON (PAIRED DATA)

C - SUPERIOR - INFERIOR COMPARISON.

ELECTRODES - 1 - SUPERIOR; 2, 3 - LATERAL; 4 - INFERIOR

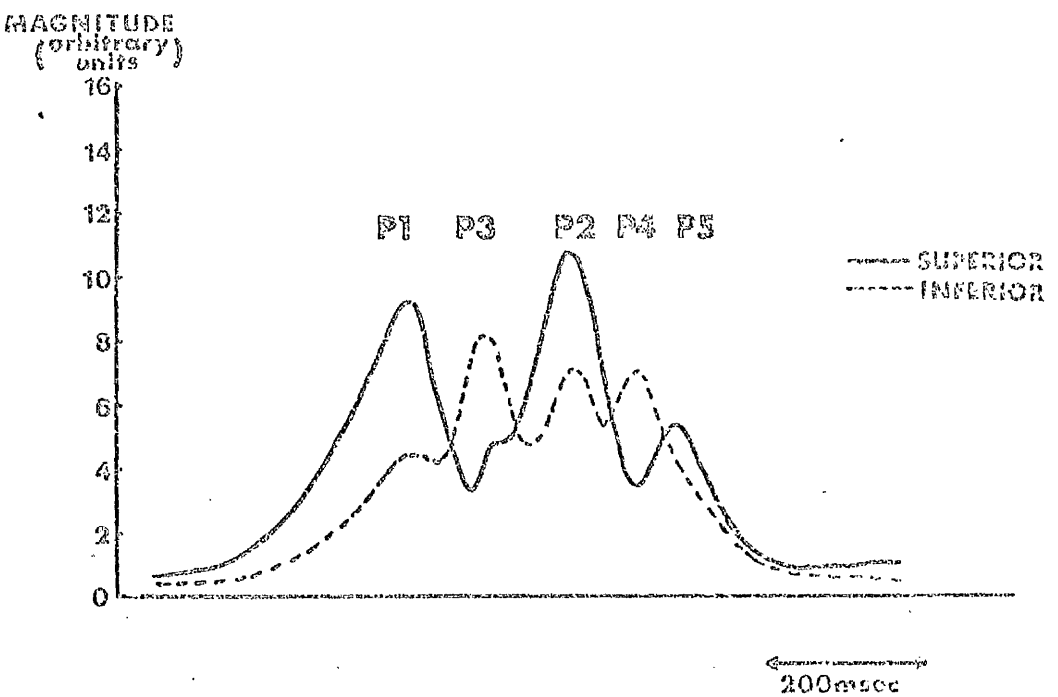


FIGURE 7.7.

EXAMPLE OF TYPICAL RESULTS OBTAINED FROM SUBJECT B WITH BIPOLAR ELECTRODE CONFIGURATION.

electrodes also gave outputs having the same form to those of subject A, with peaks P1 and P2 being much reduced in magnitude, and P3 being the largest. Peak P4 was also evident from the inferior sites, the nomenclature being the same as for subject A. However, the late peak observed from the superior electrode pairs did not correspond in time with the peak P4 from the inferior sites, and this peak was labelled P5.

The data obtained from a series of recording sessions are shown in Figure 7.8, and the results of the statistical comparisons between the ranges of the magnitudes of corresponding peaks are given in TABLE 7.2. In each case, there was found to be no statistically significant difference between the ranges in peak magnitudes resulting from left and right electrode pairs (TABLE 7.2A), and when the paired t-test was carried out, only for P2 was there any significant difference, with the activity measured from the left side electrodes being greater than that from the right side electrode pairs (TABLE 7.2B). The results of the superior-inferior comparisons showed that for P1 and P2 there was significantly greater activity measured from the superior electrodes. Peak P3 could not be reliably measured from the superior electrode results, and P5 was only evident from the superior recordings, while P4 was only present in the recordings obtained using inferior electrode pairs.

Thus the results from both subjects showed, as expected, lateral symmetry in the ranges of the magnitudes of the corresponding peaks, although the paired t-test did indicate a consistently greater level of activity recorded from the right side electrode pairs for peaks P1 and P2 from subject A. However, for all peaks from both subjects, the magnitudes of the corresponding peaks obtained from superior and inferior electrode pairs were statistically significantly different.

7.5. MONOPOLAR RECORDINGS DURING THE PRODUCTION OF THE SYLLABLE /ɔIp/.

In the previous section, there were found to be

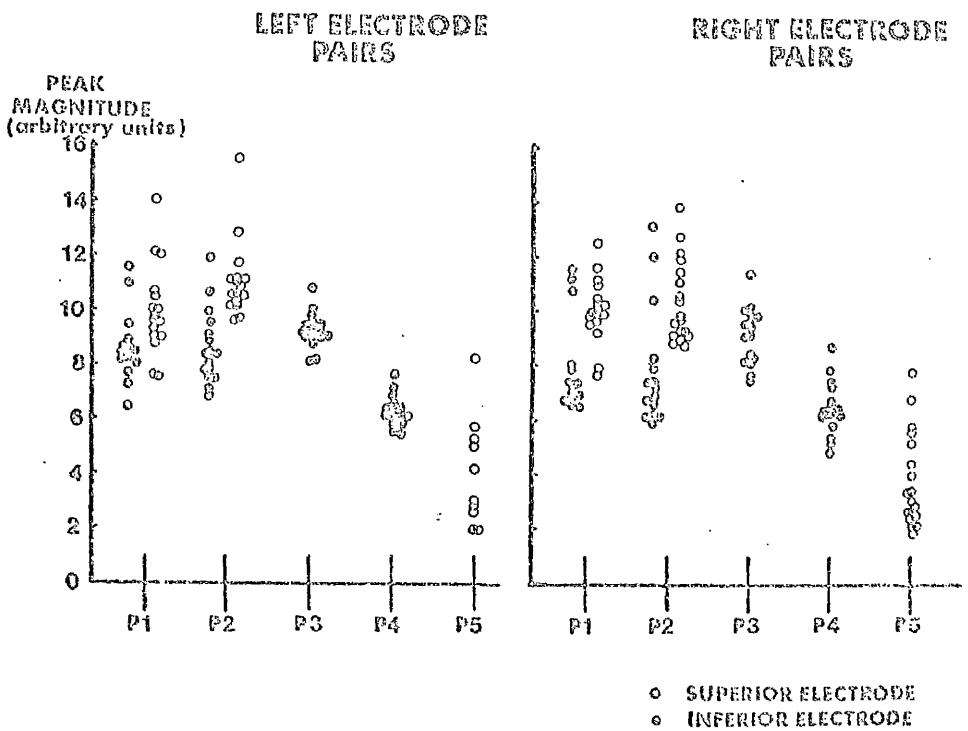


FIGURE 7.8.

DISTRIBUTION OF PEAK MAGNITUDES FOR SUBJECT B.

A.

PEAK	ELECTRODE PAIRS	n	t-VALUE	SIGNIFICANCE LEVEL	WILCOXON COEFF.	SIGNIFICANCE AT 0.05 LEVEL
P1	1-3/1-2	32	0.18	NS	339.0	NS
P1	3-4/2-4	25	0.62	NS	152.5	NS
P2	1-3/1-2	32	1.42	NS	265.0	NS
P2	3-4/2-4	30	1.25	NS	200.0	NS
P3	3-4/2-4	30	0.45	NS	261.0	NS
P4	3-4/2-4	29	1.10	NS	265.0	NS
P5	1-3/1-2	26	0.80	NS	256.0	NS

B.

PEAK	ELECTRODE PAIRS	n	t-VALUE	SIGNIFICANCE LEVEL
P1	1-3/1-2	15	0.73	NS
P1	3-4/2-4	12	1.31	NS
P2	1-3/1-2	15	3.43	< 0.005
P2	3-4/2-4	14	2.50	< 0.05
P3	3-4/2-4	14	0.67	NS
P4	3-4/2-4	14	1.11	NS
P5	1-3/1-2	9	1.58	NS

C.

PEAK	ELECTRODE PAIRS	n	t-VALUE	SIGNIFICANCE LEVEL	WILCOXON COEFF.	SIGNIFICANCE AT 0.05 LEVEL
P1	1-3/3-4	29	3.44	< 0.005	351.5	S
P1	1-2/2-4	28	2.43	< 0.025	310.0	S
P2	1-3/3-4	32	4.05	< 0.001	396.0	S
P2	1-2/2-4	30	5.04	< 0.001	358.5	S

TABLE 7.2

STATISTICAL COMPARISONS OF PEAK MAGNITUDES FOR SUBJECT B. (BIPOLAR CONFIGURATION)

B - LATERAL COMPARISON, B - LATERAL COMPARISON (PAIRED DATA)

C - SUPERIOR - INFERIOR COMPARISON

ELECTRODES - 1 - SUPERIOR; 2, 3 - LATERAL; 4 - INFERIOR

significant differences in the EMG activities measured between electrode pairs placed on the superior and inferior vermilion border. However, the activity measured by such pairs of electrodes resulted from the vector summation of the absolute activities measured at each electrode, and the differences noted in the previous section reflected, to a large extent, differences in the EMG activities measured by electrodes 1 and 4 (Figure 7.4), since either electrode 2 or 3 was common to each pair. In order to investigate the origin of these differences, the experiment described in SECTION 7.4 was repeated, but the electrodes were now connected in the monopolar mode with the reference electrode being situated on the mastoid process. The utterance used was the syllable /pIp/.

A typical result obtained from subject A is shown in Figure 7.9, with labelling of the peaks being consistent with the bipolar case. The activity measured from the superior electrode showed peaks P1, P2 and P3, with the magnitudes of P1 and P2 being similar, and very much greater than that of P3. The same peaks were evident in the output from the lateral electrode, but P1 and P2 had much reduced magnitude. The output from the inferior electrode showed the presence of P1, P2, P3 and P4, with P3 being the greatest in magnitude. The results (mean $\pm$ 1 S.D.) obtained for the peak magnitudes from a series of recording sessions, are plotted in Figure 7.10, and the results of the statistical comparisons between the ranges are given in TABLE 7.3A. The superior electrode indicated significantly greater P1 and P2 activity and significantly less P3 activity than that measured at the other sites. The P1 and P2 magnitudes measured by the lateral and inferior electrodes were not significantly different, but the P3 activity at the inferior site was significantly greater than at the other sites.

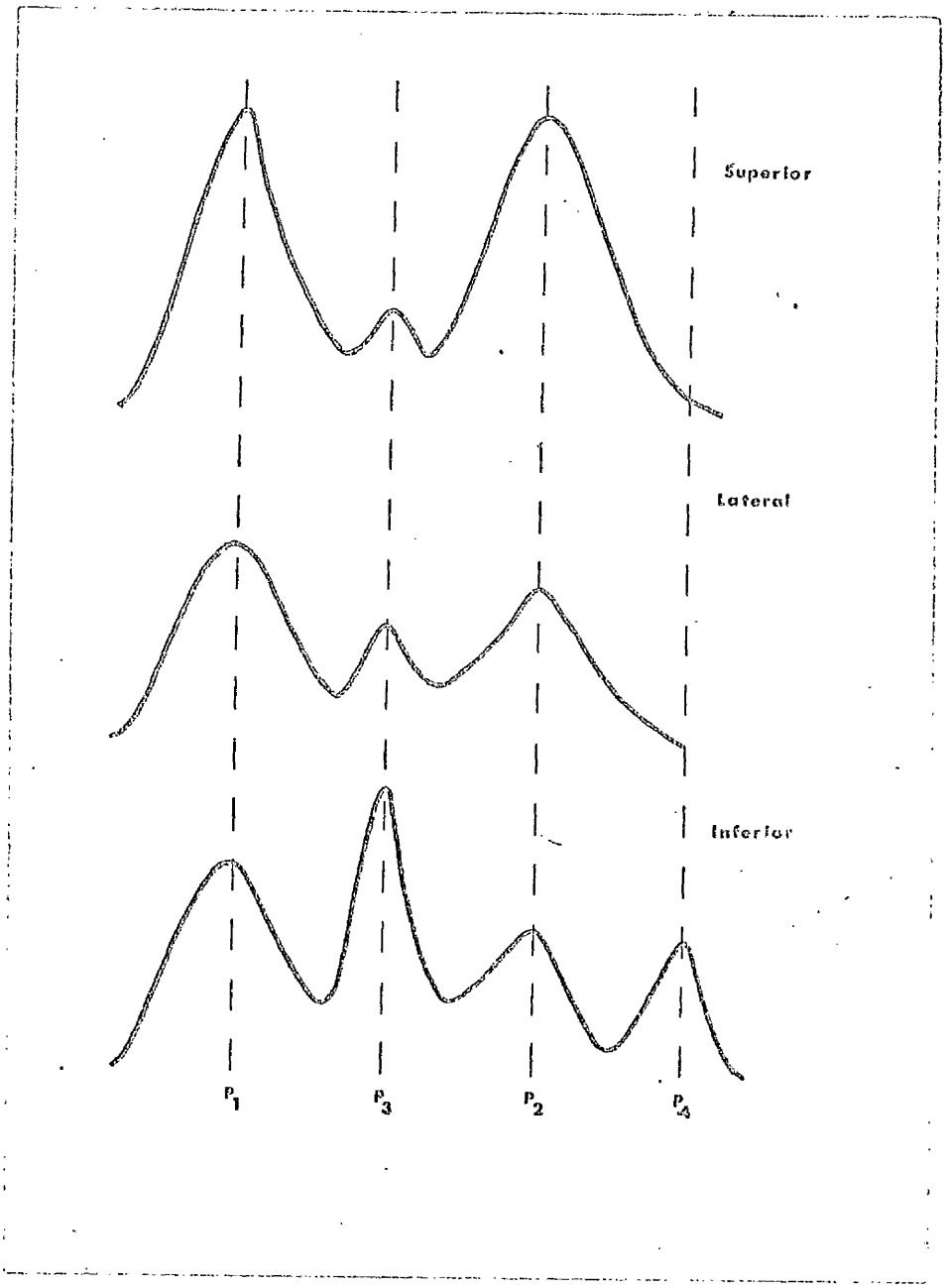


FIGURE 7.9.

EXAMPLE OF TYPICAL RESULTS OBTAINED FROM SUBJECT A WITH MONOPOLAR ELECTRODE CONFIGURATION.

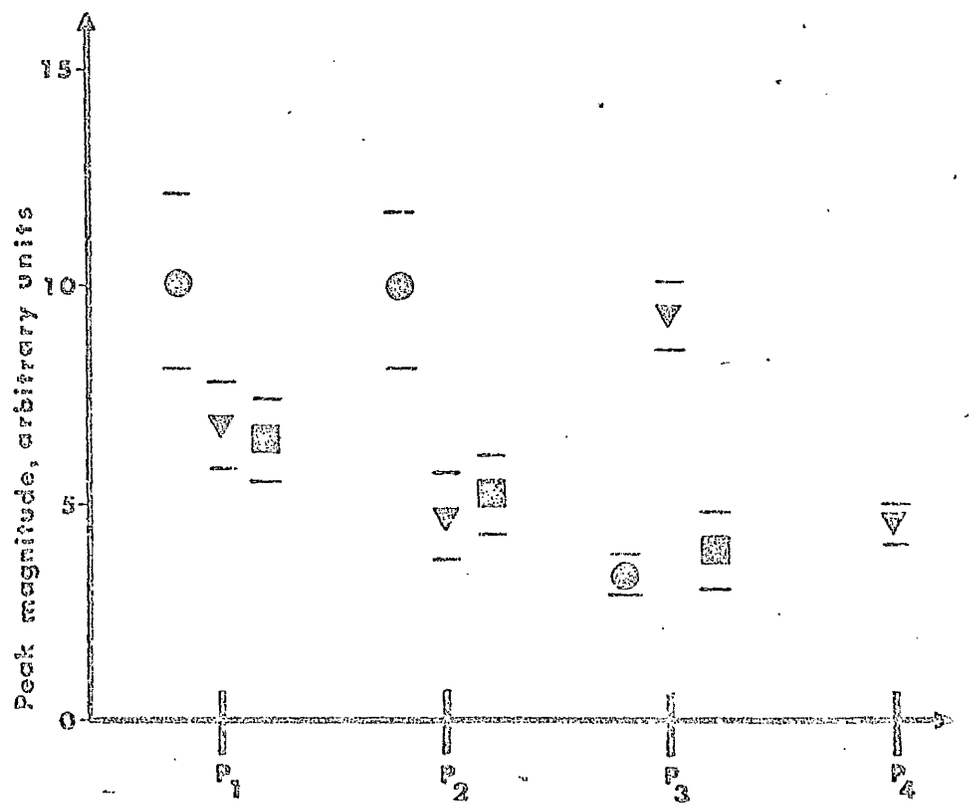


FIGURE 7.10.

MEAN PEAK MAGNITUDES OBTAINED FROM SUBJECT A

◎ - SUPERIOR ELECTRODE, ▽ - INFERIOR ELECTRODE,

▣ - LATERAL ELECTRODE.

A. SUBJECT - A

PEAK	ELECTRODES	n	t-VALUE	SIGNIFICANCE LEVEL	WILCOXON COEFF.	SIGNIFICANCE AT 0.05 LEVEL
P1	1/4	22	4.96	< 0.001	70.0	S
P2	1/4	22	8.54	< 0.001	66.0	S
P3	1/4	22	22.94	< 0.001	66.0	S
P1	1/3	32	7.08	< 0.001	238.0	S
P2	1/3	32	9.87	< 0.001	231.0	S
P3	1/3	32	2.26	< 0.05	130.0	S
P1	4/3	32	0.82	NS	175.5	NS
P2	4/3	32	1.40	NS	146.5	NS
P3	4/3	32	17.60	< 0.001	231.0	S

B. SUBJECT - B

P1	1/4	15	7.41	< 0.001	10.0	S
P2	1/4	22	8.86	< 0.001	66.0	S
P1	1/3	32	7.43	< 0.001	66.0	S
P2	1/3	32	7.70	< 0.001	316.0	S
P5	1/3	21	9.24	< 0.001	55.0	S
P1	4/3	24	0.92	NS	63.0	NS
P2	4/3	32	1.54	NS	139.0	NS

TABLE 7.3

STATISTICAL COMPARISON OF PEAK MAGNITUDES OBTAINED FROM DIFFERENT ELECTRODE SITES (MONOPOLAR CONFIGURATION)

ELECTRODES - 1 - SUPERIOR; 3 - LATERAL; 4 - INFERIOR

Figure 7.11 shows a typical output obtained from subject B. Peaks P1, P2 and P5 were present from the superior and lateral electrodes, with P1 and P2 having the greatest magnitudes in each case. The inflexion present in the output from the superior electrode may have corresponded to a small amount of activity consistent with peak P3. The recordings from the inferior electrode showed peaks P1, P2, P3 and P4, with P3 having the largest magnitude. The results for the peak magnitudes (mean $\pm$ 1S.D.) calculated for the series of recordings are shown in Figure 7.12 and the statistical comparisons are given in TABLE 7.3B. Peaks P1 and P2 were significantly larger from the superior electrode than from the other two, and P5 was significantly larger from the superior electrode than from the lateral. There were no significant differences between the magnitudes of either P1 or P2 obtained from the lateral and inferior sites. P3 and P4 were only evident from the inferior electrode.

Thus, the results obtained were similar for both subjects, and explained the general differences which had been found from the bipolar recordings.

7.6. COMPARISON BETWEEN BIPOLAR AND MONOPOLAR RESULTS.

The results so far have indicated the form of the EMG activity measured at different electrode sites and using different electrode configurations. A direct comparison was carried out, where possible, between the ranges of corresponding peaks obtained with the electrodes in either bipolar or monopolar mode, and the results are given in TABLE 7.4. Comparisons were made between the outputs obtained from the superior electrode in monopolar mode and that from the superior-lateral bipolar pair, and also between the inferior electrode in monopolar mode and the bipolar pair, inferior-lateral.

There were significant differences in the magnitudes of P1 and P2 in all cases for both subjects, and also for P3 (superior/superior-lateral) for subject A.

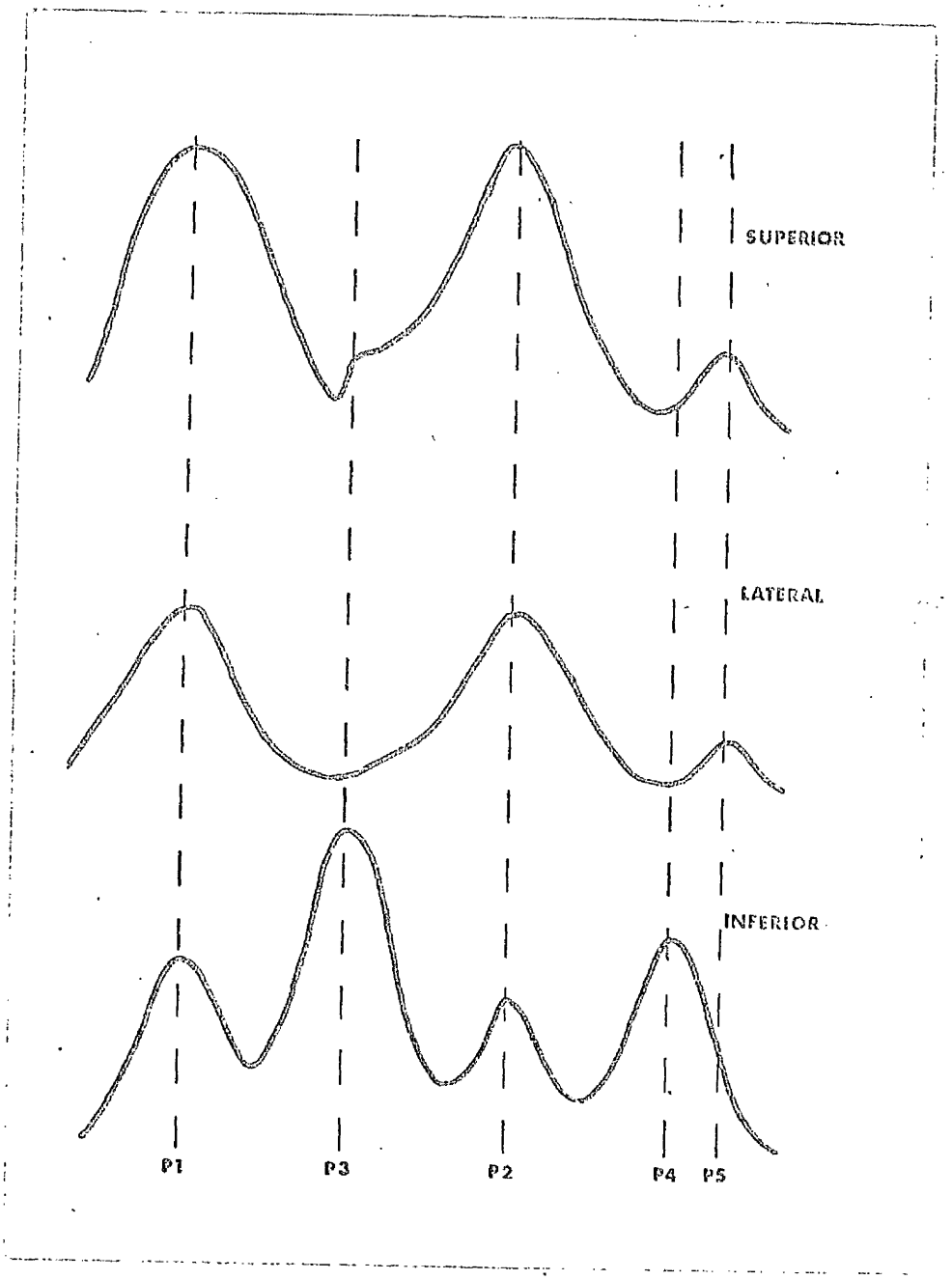


FIGURE 7.11.

EXAMPLE OF TYPICAL RESULTS OBTAINED FROM SUBJECT B WITH MONOPOLAR ELECTRODE CONFIGURATION.

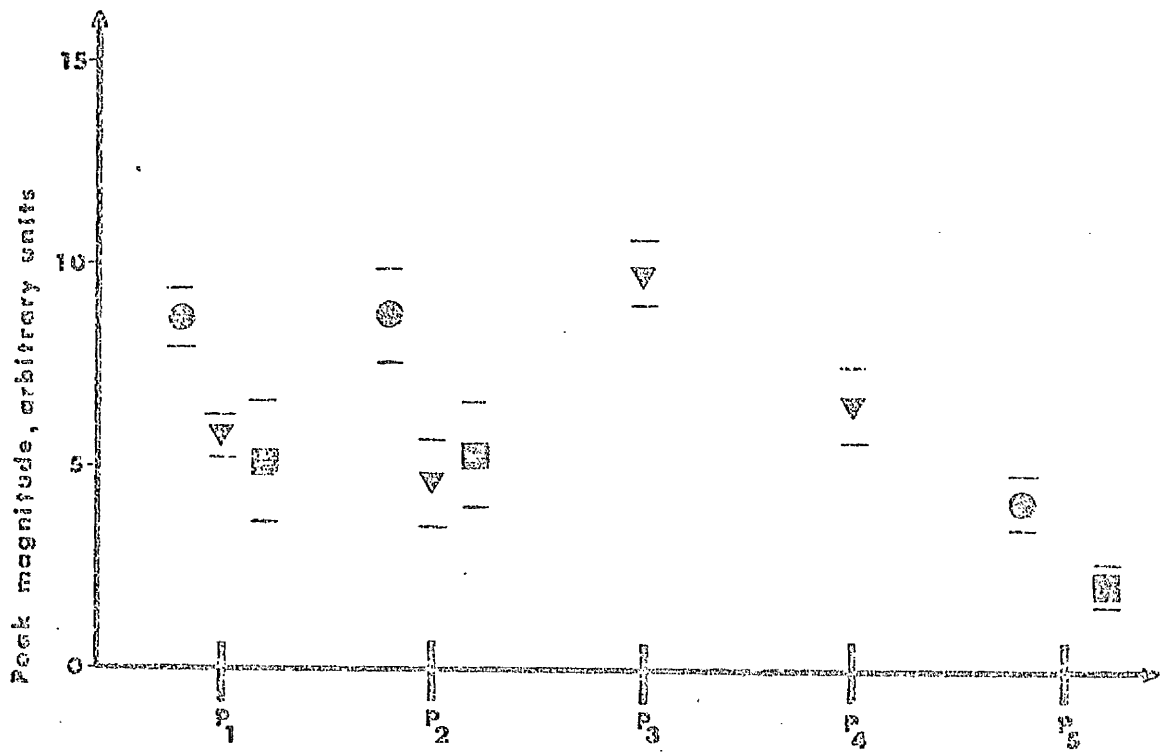


FIGURE 7.12.

MEAN PEAK MAGNITUDES OBTAINED FROM SUBJECT B

◍ - SUPERIOR ELECTRODE, ▽ - INFERIOR ELECTRODE,

▩ - LATERAL ELECTRODE.

A. SUBJECT - A

PEAK	ELECTRODES	n	t-VALUE	SIGNIFICANCE LEVEL	WILCOXON COEFF.	SIGNIFICANCE AT 0.05 LEVEL
P1	1-3/1	51	3.96	< 0.001	141.0	S
P1	4-3/4	50	3.32	< 0.005	145.0	S
P2	1-3/1	51	3.71	< 0.001	154.5	S
P2	4-3/4	50	4.22	< 0.001	136.5	S
P3	1-3/1	51	6.79	< 0.001	82.0	S
P3	4-3/4	50	0.05	NS	262.0	NS

B. SUBJECT - B

P1	1-3/1	45	3.04	< 0.005	131.0	S
P1	4-3/4	31	3.29	< 0.005	10.0	S
P2	1-3/1	45	3.88	< 0.001	118.0	S
P2	4-3/4	43	6.16	< 0.001	71.5	S
P3	4-3/4	43	1.87	NS	296.5	NS
P4	4-3/4	42	0.88	NS	229.5	NS
P5	1-3/1	39	0.45	NS	258.5	NS

TABLE 7.4

STATISTICAL COMPARISONS OF PEAK MAGNITUDES OBTAINED WITH BIPOLAR AND MONOPOLAR CONFIGURATIONS.

ELECTRODES - 1 - SUPERIOR; 3 - LATERAL; 4 - INFERIOR

P3 (inferior/inferior/lateral) for both subjects and also P4 and P5 for subject B showed no statistically significant differences. These results were also expected on the basis of vector addition, e.g. in general if a peak had a contribution from both electrodes connected in the bipolar mode, then the peak magnitude would be different to that measured from either electrode connected in monopolar mode. Similarly, if a peak had a contribution from only one electrode of a bipolar pair, then the other electrode acted as the reference electrode, and the configuration was basically monopolar for the activity corresponding to that peak. Thus great care must be taken, when comparing data from different studies, that the electrode configurations are similar.

7.7. ACTIVITY ASSOCIATED WITH INITIAL AND FINAL CONSONANT.

The forms of the magnitude of EMG activity associated with different electrode sites, superior and inferior midline and also lateral, during the utterance of the syllable /pIp/ have been discussed. In order to associate with particular parts of the utterance the EMG activity giving rise to the various peaks observed, several series of recordings were made in which the utterances used were /pI/ and /Ip/, i.e. in order to confine the labial gestures to either the initial or final consonant. Recordings were made in both bipolar and monopolar modes, and the results are illustrated by one example for each subject.

Figure 7.13 shows typical outputs obtained from subject A, with the electrodes connected in the bipolar mode. The EMG activity associated with the utterance /pI/ consisted mainly of P1 from the superior-lateral electrode pair, with a small contribution to P3, while from the inferior-lateral electrodes, P1 and P3 were of similar magnitude, but smaller than that of P1

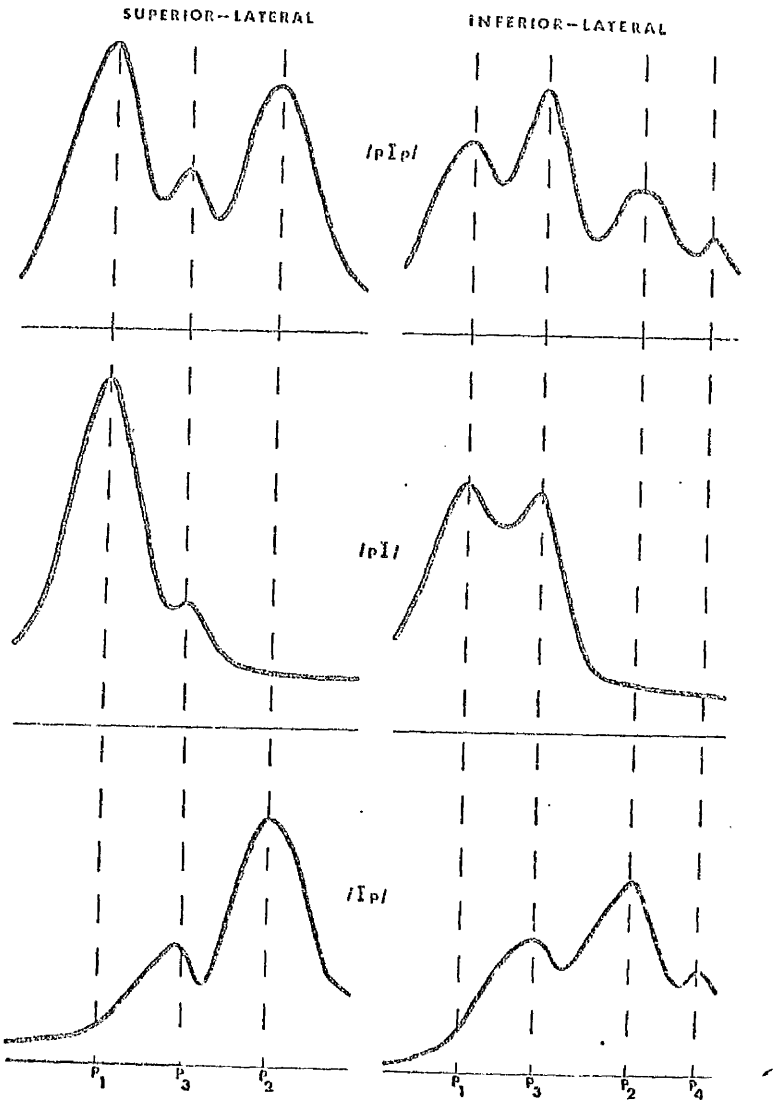


FIGURE 7.13.

EXAMPLES OF OUTPUTS OBTAINED FROM SUBJECT A WITH BIPOLAR ELECTRODE CONFIGURATION FOR UTTERANCES /pIp/, /pI/ AND /Ip/.

from the previous sites. The utterance /Ip/ produced the large peak P2 and a smaller P3 from the superior-lateral electrodes, and peaks P2, P3 and P4 from the inferior-lateral pair, with P2 having the greatest magnitude. The magnitudes of the peaks measured from the utterances were very similar to those obtained from the composite utterance, /pIp/.

Figure 7.14 illustrates the outputs obtained from subject B with the electrodes connected in monopolar mode. The utterance, /pI/, resulted in a large peak, P1, from the superior electrode, with a small contribution to P2. The output from the lateral electrode showed peak P1 to be smaller than in the previous case, with peak P3 being evident and larger than P2, which had approximately the same magnitude as previously. Peaks P1 and P3 were obtained from the inferior electrode site, with P3 being the larger and of similar magnitude to the same peak arising from the utterance /pIp/. The utterance /Ip/ produced a large peak corresponding to P2 from the superior electrode, with the peak, P5, also being present. The lateral electrode site demonstrated peak P2, although with a reduced magnitude, and also small contributions corresponding to P3 and perhaps P5. P2 was the largest peak present in the output from the inferior electrodes, but P3 was greatly increased in magnitude, and P4 had replaced P5 as the peak occurring after P2.

These results indicated that peaks P1 and P2 were associated with the lip closure of the initial and final /p/'s respectively. P3 was present in both utterances /pI/ and /Ip/ and could not, therefore, be directly correlated to either consonant. Peaks P4, and also P5 in subject B, were only observed during the production of /Ip/. It was probable that the peaks P3, P4 and P5 resulted from the muscular readjustments necessary for the transitions from consonant to vowel, vowel to consonant and also those occurring after the cessation of the utterance.

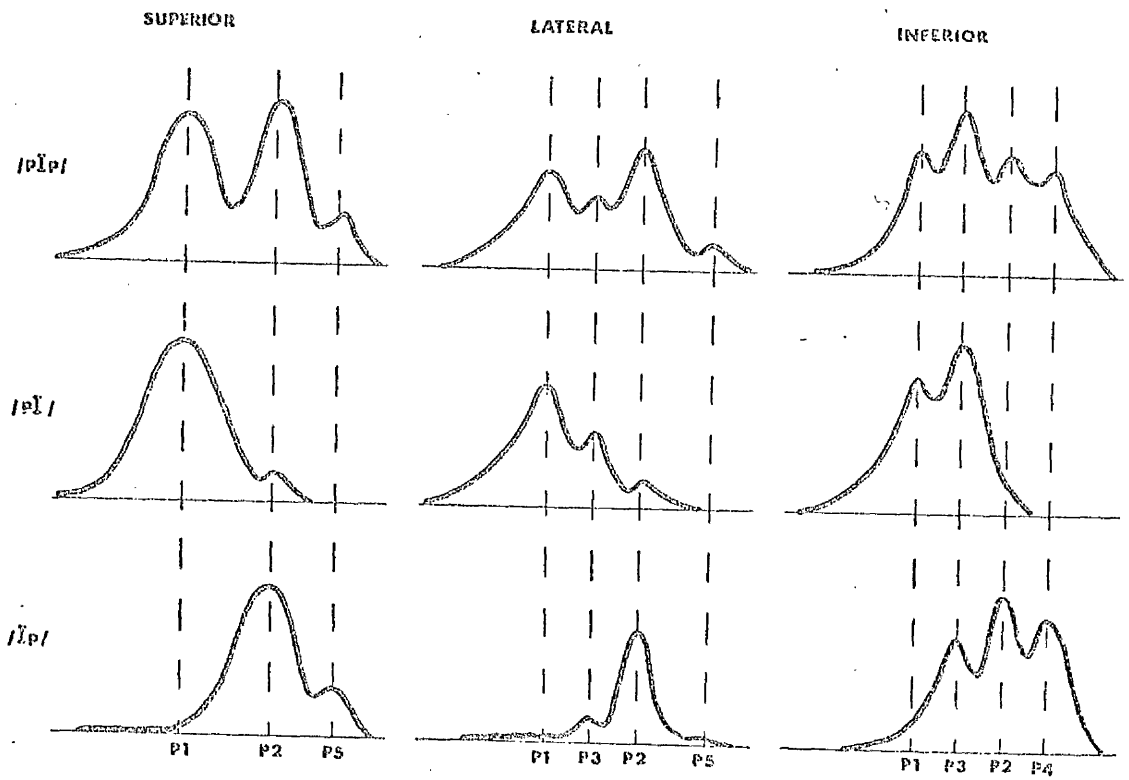


FIGURE 7.14.

EXAMPLES OF OUTPUTS OBTAINED FROM SUBJECT B WITH MONOPOLAR ELECTRODE CONFIGURATION FOR UTTERANCES /pIp/, /pI/ AND /Ip/.

7.8. THE EFFECT OF PHONETIC CONTEXT ON THE MEASURED EMG ACTIVITY.

The results described in the previous section permitted the peaks in the EMG activity to be associated with specific events, in particular, P1 and P2 were associated with the muscular adjustments during the initial and final consonant, /p/, respectively. In order to investigate the dependence of the magnitudes of the various peaks upon the particular consonant or vowel sound, a limited experiment was carried out in which a series of utterances were used, i.e. /pIp/, /pʌp/, /pIb/, /bIp/, /bIb/ and /bʌb/. This allowed comparisons to be made between the bi-labial plosives /p/ and /b/, in various combinations with the vowels /I/ and /ʌ/. The EMG activities were recorded in both bipolar and monopolar modes, and the peak magnitudes (mean $\pm$ 1S.D.) obtained from subject A are shown in figures 7.15 and 7.16. The forms of the outputs were similar in each case, and so the peaks P1 and P2 were interpreted as resulting from the initial and final consonants, whether these were /p/ or /b/. The results from the bipolar recordings from subject A suggested that the magnitude of P1 was greater in the syllables in which the initial consonant was /p/, but the final consonant and central vowel sound appeared to have no consistent effect upon the magnitudes of any of the other peaks. The data for subject B (Figures 7.17 and 7.18) demonstrated that none of the phonetic contexts investigated produced consistent differences in the measured EMG activity.

The results obtained agreed, in general, with those reported by Fromkin (1966), and MacNeilage and De Clerk (1969) i.e. that the central vowel in a CVC utterance had little or no effect on the peak heights of the measured EMG activity. Fromkin and Ladefoged (1966) and Lubker and Parris (1970) also found that the EMG signal recorded from the lip musculature did not produce a reliable discrimination between the labial consonants

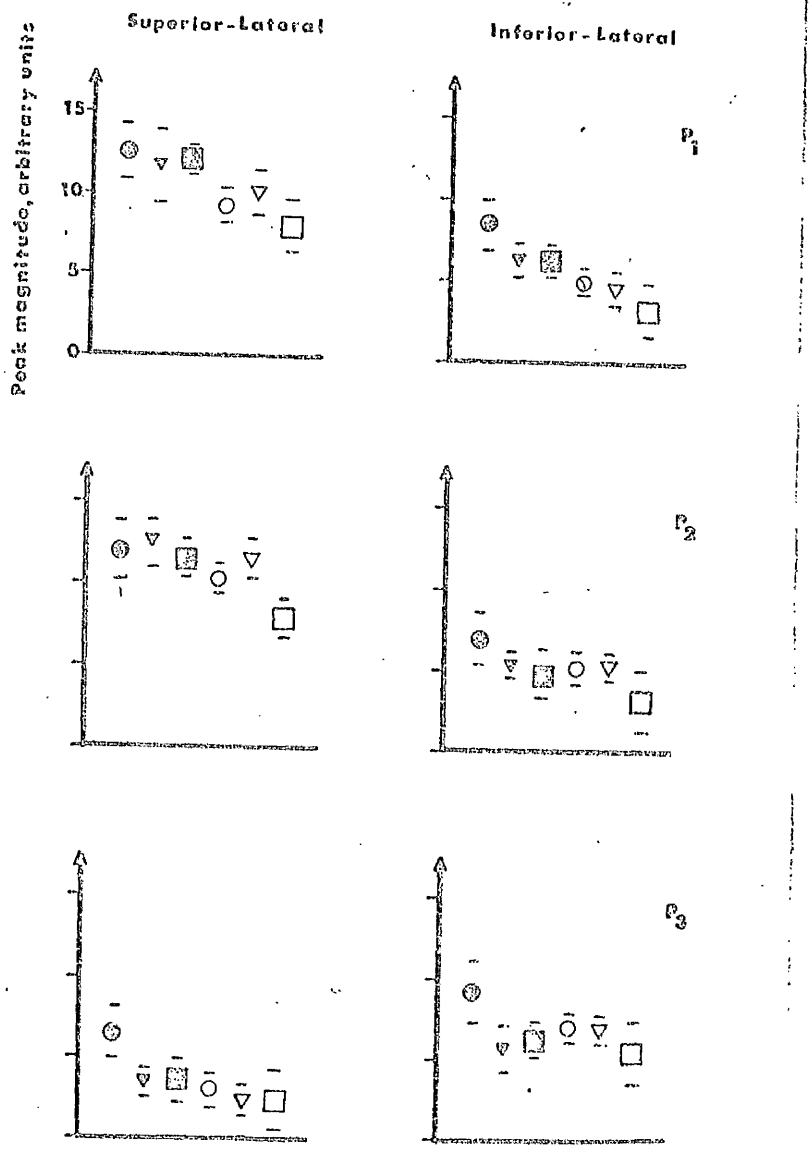


FIGURE 7.15.

MEAN PEAK MAGNITUDES OBTAINED FROM SUBJECT A (BIPOLAR CONFIGURATION).

◍ - /pIp/, ▽ - /pʌp/, ▣ - /pIb/,
○ - /bIp/, ▽ - /bIb/, □ - /bʌb/.

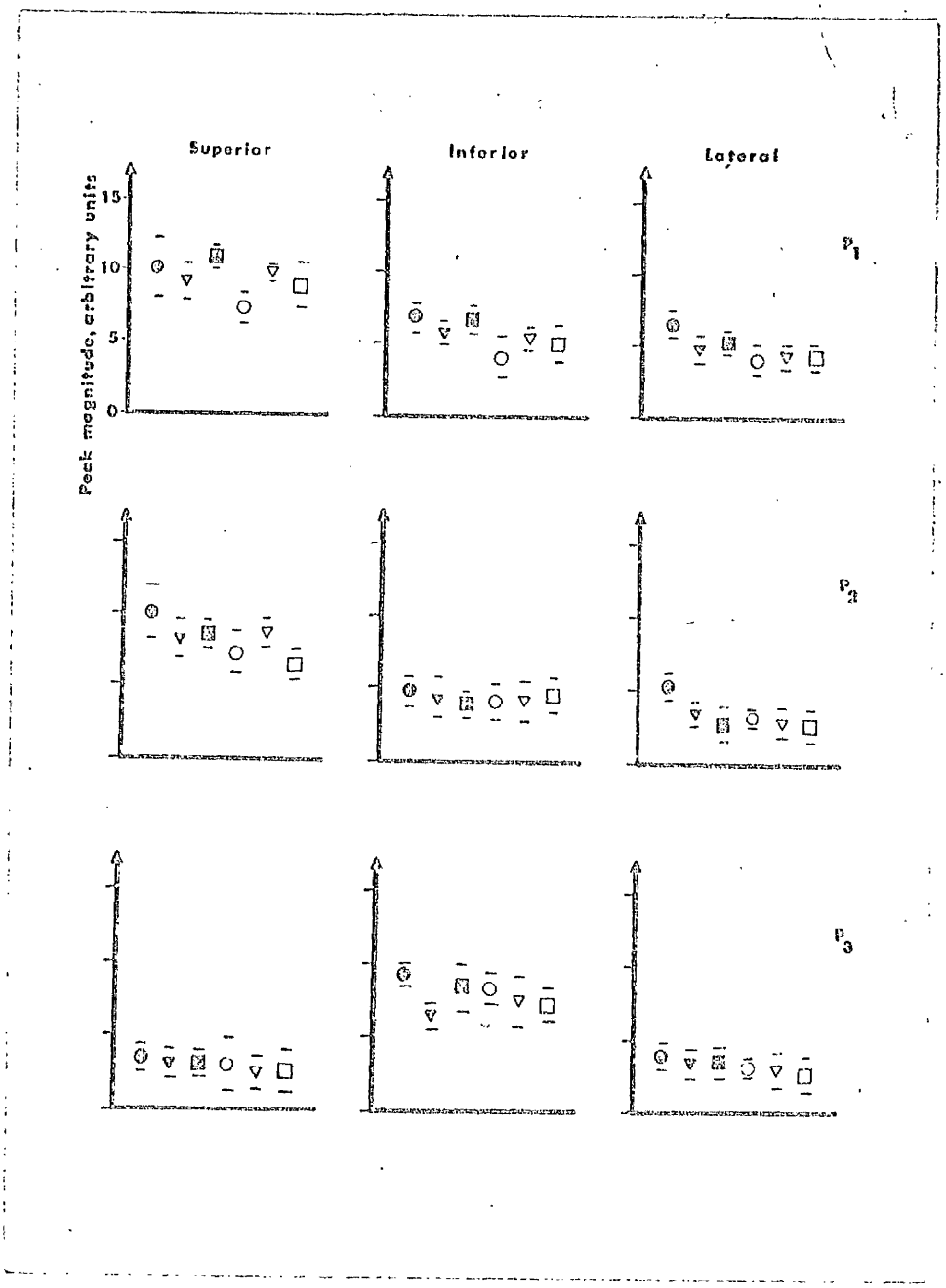


FIGURE 7.16.

MEAN PEAK MAGNITUDES OBTAINED FROM SUBJECT A (MONOPOLAR CONFIGURATION)

◍ - /pIp/, ▽ - /pʌp/, ▣ -/pIb/,
○ - /bIp/, ▽ - /bIb/, □ -/bʌb/.

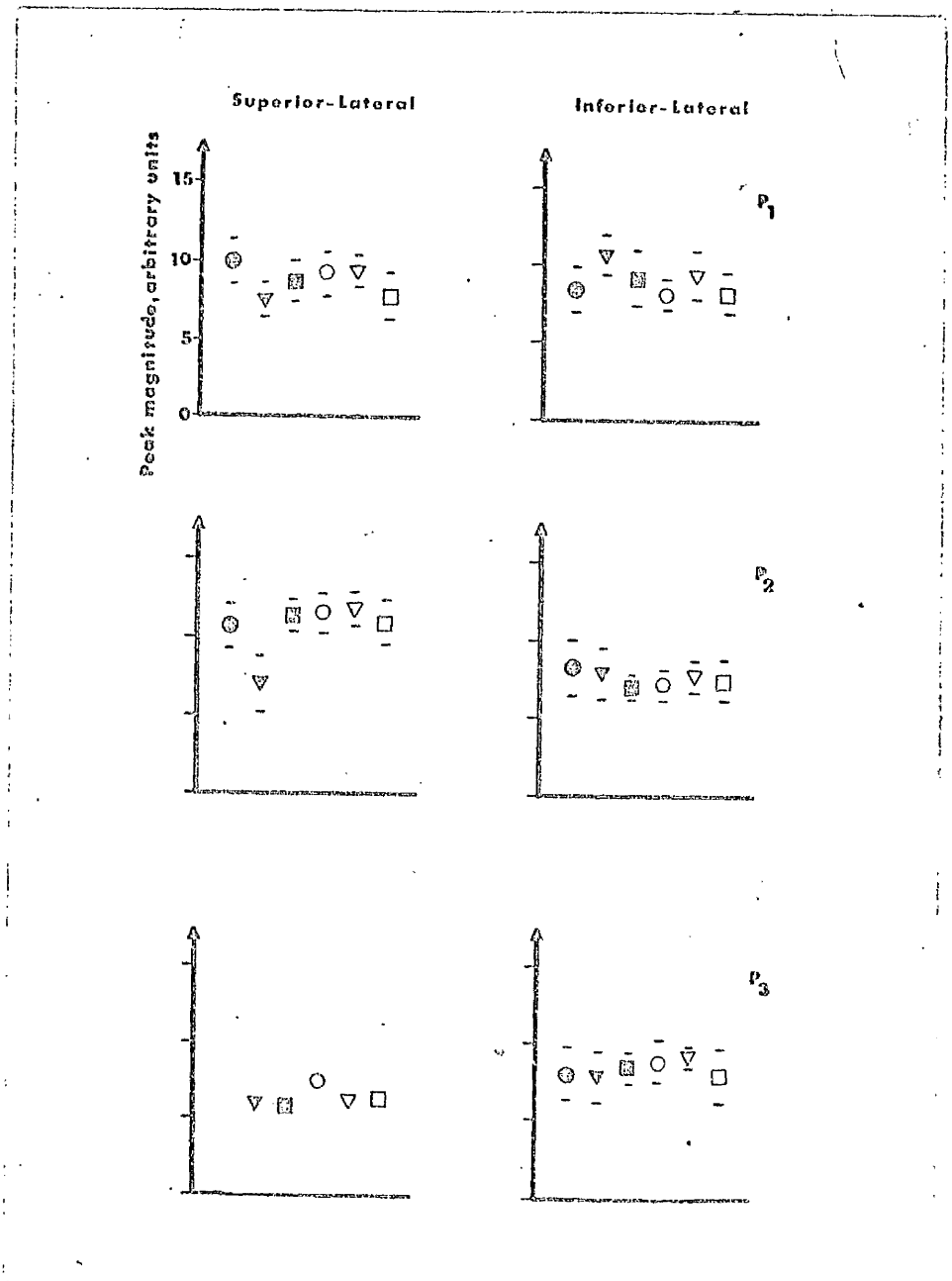


FIGURE 7.17.

MEAN PEAK MAGNITUDES OBTAINED FROM SUBJECT B (BIPOLAR CONFIGURATION)

◍ - /pIp/, ▽ - /p Λ p/, ▩ - /pIb/,
○ - /bIp/, ▽ - /bIb/, □ - /b Λ b/.

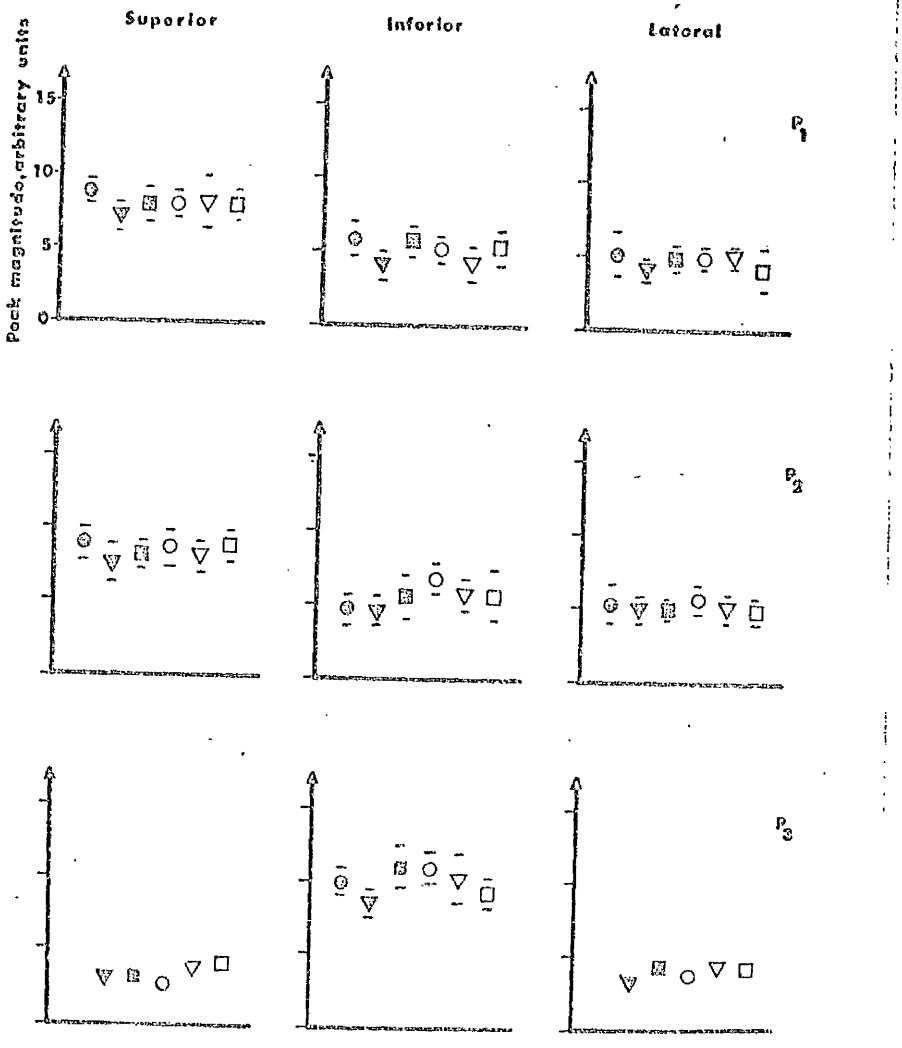


FIGURE 7.18

MEAN PEAK MAGNITUDES OBTAINED FROM SUBJECT B (MONOPOLAR CONFIGURATION)

◍ – /pIp/, ▽ – /pʌp/, ▩ – /pIb/,
○ – /bIp/, ▽ – /bIb/, □ – /bʌb/.

/p/ and /b/. This indicated either that such a discrimination was not possible on the basis of EMG activity, or that it was necessary in this part of the study to use needle electrodes to record from specific muscle groups, which individually might have actions dependent on the phonetic context.

7.9. THE MUSCLE ORIGINS OF THE MEASURED EMG ACTIVITY.

The EMG activity of the lip musculature has been studied during the production of CVC utterances using surface electrodes placed at fixed sites on the vermilion border, i.e. superior midline, inferior midline and lateral, and basic differences in the forms of the outputs recorded from these sites have been found. Since the surface electrodes recorded from finite volumes of muscle tissue surrounding each site, and not merely from a single or small group of motor units as would be the case with needle electrodes, the differences in the measured EMG activities reflected differences in the underlying muscle structure, i.e. the surface electrodes not only recorded activity from the muscle orbicularis oris (m.orb.oris), but also from some of the other facial muscles which were used to control lip movement. The study was extended to investigate whether it was possible, using different arrays, to map out spatially the regions of origin of the EMG activity giving rise to the various peaks, and from a knowledge of the facial muscle structure, to be able to attribute these peaks to particular muscle groups.

The superior and inferior portions of the lip musculature were investigated using the electrode arrays shown in Figure 7.19 A and B respectively. The electrodes were always connected in monopolar mode, and the utterance used was /pIp/. The subject was photographed at each recording session, and the position of each electrode was expressed as a percentage of the

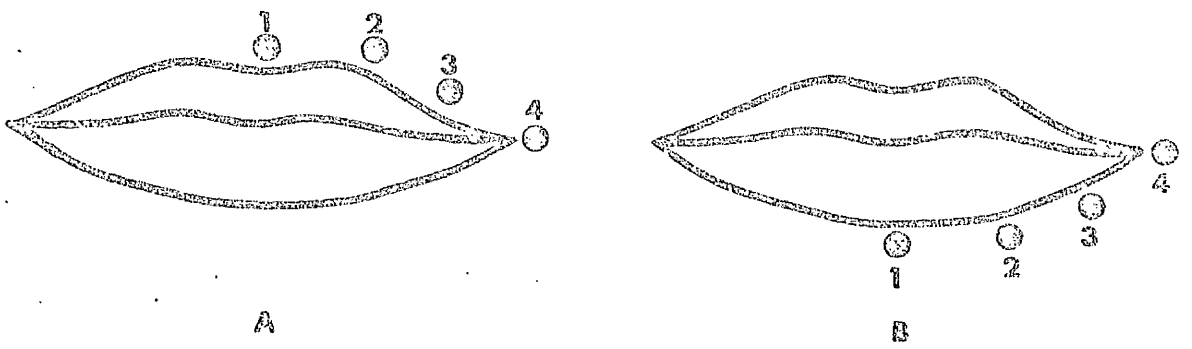


FIGURE 7.19.

ELECTRODE ARRAYS USED TO STUDY SPATIAL DISTRIBUTION OF EMG ACTIVITY.

midline - lateral distance of the vermilion border, with 0% representing midline and 100% the furthest lateral extremity.

The spatial distributions obtained for peaks P1 - P4 from subject A are shown in Figure 7.20. The P1 activity measured from the superior sites had a maximum magnitude at the midline, decreased steadily to a minimum of about half of its midline value, and then increased again as the lateral extremity was approached. The P1 distribution obtained from the inferior portion was similar but had a consistently smaller magnitude. The P2 activity followed very closely the spatial distribution of the peak, P1. The P3 activity attained a maximum magnitude about midway between the midline and lateral extremity on the superior vermilion border and decreased steadily both towards 0% and 100%. The inferior P3 activity showed a maximum magnitude at midline and decreased rapidly towards the lateral extremity, although the magnitude was always greater than that measured from the superior electrodes. The P4 distribution on the inferior border had a maximum at, or near, midline, and decreased rapidly to zero as the lateral extremity was approached.

The results obtained from subject B are given in Figure 7.21. The distributions obtained for P1, P2 and for P3, P4 on the inferior border were very similar to those described for subject A. There was no measurable activity corresponding to P3 or P4 from the superior border, but the P5 activity was at a maximum on midline and decreased steadily to zero as the lateral extremity was reached.

The muscle groups giving rise to the above activity could be identified with reference to the anatomy of the relevant musculature, shown in Figure 7.22 (Gray,1973; Palmer and La Russo,1965). The superior midline maximum in the P1 activity was due to the action of muscle levator labii superioris (m. lev. lab. sup.).

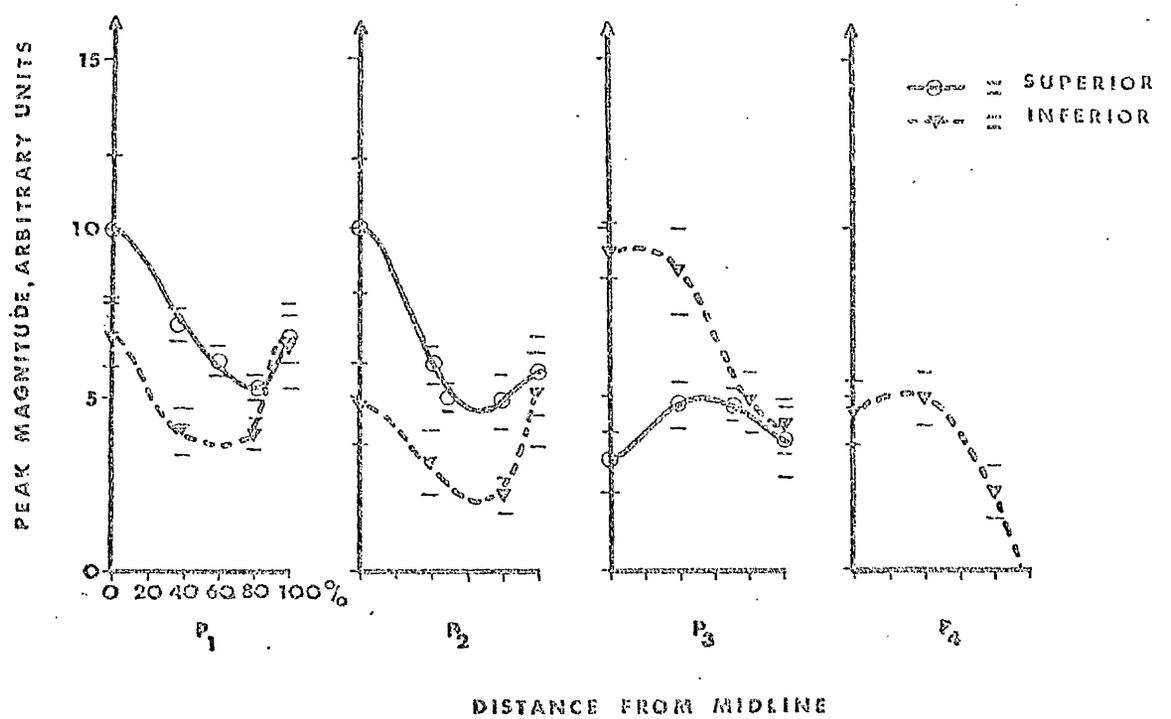


FIGURE 7.20.

SPATIAL DISTRIBUTION OF EMG ACTIVITY OBTAINED FOR SUBJECT A.

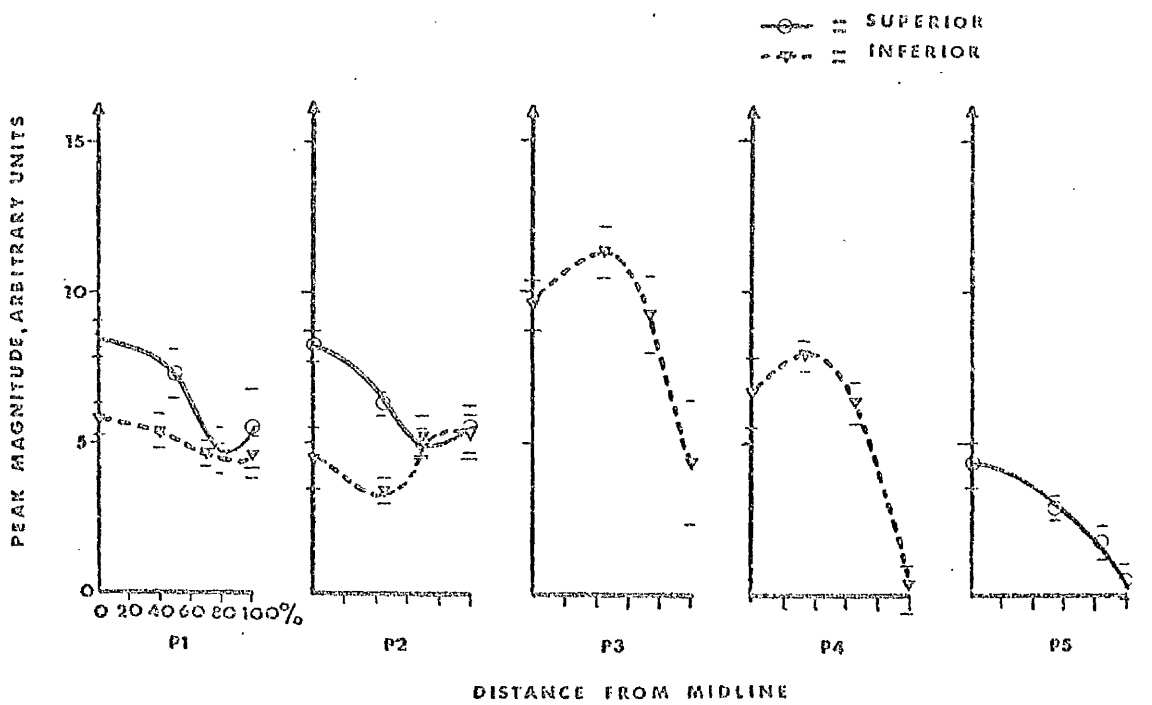


FIGURE 7.21.

SPATIAL DISTRIBUTION OF EMG ACTIVITY OBTAINED FOR SUBJECT B.

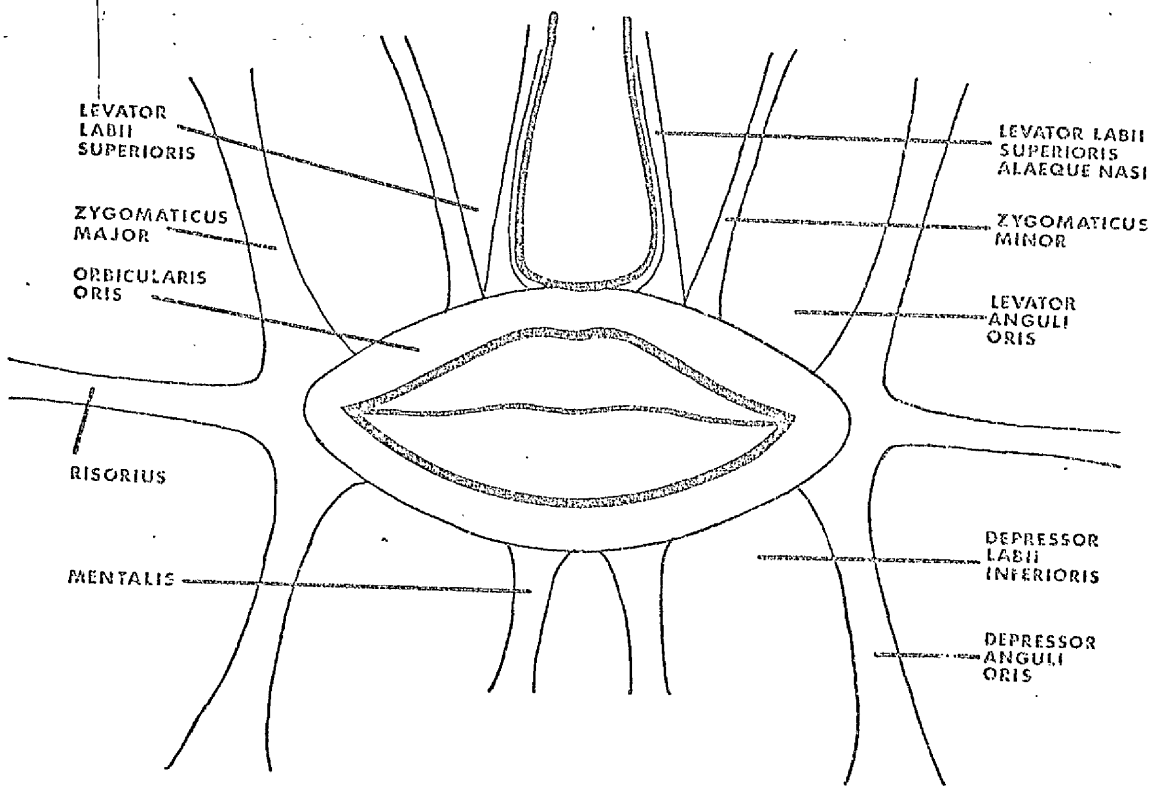


FIGURE 7.22.

DIAGRAM SHOWING THE MAIN MUSCLE GROUPS ASSOCIATED WITH LIP MOVEMENT.

This activity decreased rapidly to a minimum, but then increased again as the lateral extremity was approached, due to the action of the muscles zygomaticus major (m. zyg. maj.) and depressor anguli oris (m. dep. ang. oris). The inferior vermilion border gave rise to a similar distribution of activity, with the midline maximum being attributed to the muscle mentalis (m. ment.) and a probable contribution from the muscle depressor labii inferioris (m. dep. lab. inf.). The action of the m. orb. oris gave rise to a general background level of activity. The muscles giving rise to the peak, P2, were the same as described for P1. The P3 and P4 activities measured from the inferior electrode sites showed maxima due to the action of the m. dep. lab. inf. and probably the m. ment. The P3 activity obtained from the superior recordings of subject A reached a maximum between midline and the lateral extremity due to the m. lev. lab. sup. The m. zyg. maj. and m. dep. ang. oris possibly contributed to some of the activity measured at the lateral extremity. The P5 activity recorded from subject B was attributed to the m. lev. lab. sup. The differences between the outputs obtained from the two subjects were due to slight variations in the detailed structure of the lip musculature, and also to differences in the phonation of the test utterance.

7.10. DISCUSSION.

The experiments described in this chapter involved the measurement of the EMG activity of the lip musculature at a limited number of sites during the utterance of a CVC syllable. The results showed, as expected, that, in general, there was a lateral symmetry in the outputs measured, but that the activity measured from the superior portion of the vermilion border differed greatly to that obtained from the inferior border. Significant differences were also obtained depending upon the configuration of electrode

connection, i.e. bipolar or monopolar. These factors combined to explain some of the contradictory statements in the literature regarding the form and relative magnitudes of the various peaks, and stressed the necessity of taking great care in the comparison of results obtained from different studies. A more detailed study, using electrode arrays placed on the superior and inferior margins allowed the spatial distributions of the EMG activity to be measured, and the muscle groups giving rise to specific activity to be identified.

However, once the muscle groups giving rise to the activity measured in the present study had been identified, it was possible to compare these results with those of Leanderson and his co-workers, who had used arrays of needle electrodes implanted directly into specific muscle groups. The subjects in these studies were required to phonate longer and more complex phrases than those used in the present investigation, but comparisons with the activities associated with the voiceless plosive, /p/, could be achieved. In one study (Leanderson et al, 1967), the EMG signals were integrated, but not averaged, and recordings were made only from inferior electrode sites. However, peaks corresponding to P1 and P3 were associated with the consonant, /p/. P1 was present from the m. ment. and decreased in magnitude towards the more lateral sites (corresponding to the m. dep. lab. inf.) until the m. dep. ang. oris was reached, where it showed an increase in magnitude. P3 was just observable from the m. ment., increased to a maximum at the m. dep. lab. inf., and then decreased to zero at the m. dep. ang. oris.

In further studies (Leanderson et al,1971; Leanderson and Lindblom,1972) which did not involve any signal processing, the activities associated with the m. ment., m. dep. ang. oris and m. dep. lab. inf. were confirmed. Recordings were also obtained from

the m. orb. oris sup., m. orb. oris inf. and the m. lev. lab. sup. There were large amounts of activity measured from the m. orb. oris sites corresponding to P1, and the m. lev. lab. sup. electrode demonstrated activity consistent with P3.

Thus, the results of the present study agreed well with those obtained by other investigators who had used more invasive techniques involving needle electrodes.

7.11. SUMMARY.

The technique of recording EMG signals was described, and the analysis procedure involving signal averaging was discussed. The recordings obtained during the repeated utterance of the syllable /pIp/ were used to confirm lateral symmetry in the EMG activity, and to demonstrate the differences due to electrode position and also the configuration of electrode connection, i.e. bipolar or monopolar. The peaks in the activity measured were associated with the corresponding consonant of the CVC syllable or to muscle re-adjustments arising from CV or VC transitions. No consistent changes could be related to contextual effects when either the consonants /p/ and /b/ were interchanged, or the vowels /I/ and /ʌ/ were alternated. Arrays of electrodes situated on the superior and inferior vermilion borders were used to provide the spatial distributions for the EMG activity, and these were used to identify the muscle groups giving rise to the measured activity. The results agreed very well with those of other experimenters, who had used the invasive technique involving the use of needle electrodes.

CHAPTER 8.

REVIEW

CHAPTER 8. REVIEW

It was stated in CHAPTER 1 that speech formed the basis of man's ability to communicate, and that impairment of this function might lead to serious consequences for the subject, either clinical, social or economic. The department, within which the work described in this thesis was carried out, had between 200 and 300 new referals each year, with the number showing a steady increase. The majority of these cases arose from vocal misuse, sometimes accompanied by psychogenic factors, while others were caused by neurological or respiratory disease or carcinogenic involvement etc. Although some of these conditions required surgical intervention, the majority could be successfully treated by non-invasive techniques once the diagnosis and initial assessment had been carried out. However, the standard assessment consisted mainly of the visual inspection of the patient, with the result that it was entirely subjective and thus, unsatisfactory, as it gave little indication of the appropriate course of treatment. One of the aims of the work of this thesis was to provide an objective method of assessing vocal function, and the development of such a technique was described in CHAPTER 3.

The objective assessment was based on aerodynamic measurements and was divided into two series of tests, (Kelman et al, 1975). The first series measured the variability in a subject's respiratory pattern during a period of quiet respiration, and the results obtained from a control group of subjectively normal subjects were compared to those from a group of dysphonic patients. It was found that the dysphonic group showed greater variability than the control group, and, on the basis of a comparison between the groups, statistically significant "limits of normality" were calculated.

The second series of tests consisted of measuring air flow parameters while the subject sustained the vowel /a/. A normal range for the value of the mean

flow rate was obtained, and this agreed with results reported by Isshiki (Isshiki and Von Leden,1964). The variation of flow rate was estimated by calculating the parameter (mean flow rate/peak flow rate), and by the comparison of the data obtained from the control and dysphonic groups, a "limit of normality" was calculated. The volume of air expired during the phonation (phonation volume) showed a close correlation with the subject's vital capacity, and, using the data obtained from the control group, a "region of normality" was defined on the graph of phonation volume versus vital capacity. Since the air used during phonation was highly dependent on the vital capacity, it was necessary to investigate whether the measured vital capacity was as large as would be expected on the basis of the subject's physical characteristics. This was achieved by calculating the expected vital capacity for each subject using Baldwin's Formulae and obtaining the "limit of normality" for the ratio - (measured vital capacity/Baldwin's vital capacity).

Thus each series of tests consisted of the measurement of four parameters, the values of which were compared to the limits obtained by the statistical comparison between the data of the control group and those of the dysphonic group. A score of 1 or 0 was assigned to each test result which lay outwith or within the prescribed "limit of normality", and the total scores showed a close correlation with the results of the standard subjective assessment. However, the results of the objective assessment gave a quantitative measure of the degree of abnormality for each subject, and also indicated which aspect of treatment would be most appropriate, e.g. correction of the quiet respiratory pattern, control over air flow rate during phonation or increase in the vital capacity and air used for phonation. The recording session could be rapidly carried out (about 30 mins.), and the technique was totally non-invasive to the subject. The degree of patient co-operation required was

minimal as the phonation was sustained at the subjectively "normal" fundamental frequency, and at a "comfortable" sound level.

The "limits of normality" were calculated by firstly showing that the results obtained from both groups were highly significantly different, by using the CHI^2 technique (APPENDIX 1), and then optimising the limit value to allow for a level of about 10% for the number of false positive diagnoses (i.e. "normal" subjects who would be classified as "abnormal"). Thus the limits were, to some extent, dependent upon the distribution of results of the dysphonic subjects. If the number of subjects in the control group was considerably increased, it might be possible to fit the distributions of the results obtained with known statistical distributions and, thus, to obtain statistical limits, e.g. 95% limits, which would be totally independent of the results from the dysphonic group. This was not feasible in the present investigation. However, the assessment, as described in CHAPTER 3, is carried out routinely for every patient referred to the speech therapy department, and the results are proving of great benefit to patient management.

The frequency analysis of vowel sounds was described in CHAPTER 4, and the role of formants in vowel recognition discussed. The classification of hoarseness suggested by Yanagihara (1967) was explained. This was based on the presence of random noise components in the spectra of vowels phonated by subjects who were subjectively described as hoarse, and the intensity and frequency of these components indicated the degree of severity of hoarseness. A group of 85 dysphonic subjects was classified according to this system, and the results obtained were compared to those obtained by using the aerodynamic assessment. A highly significant correlation ($p < 0.0005$) was found between the results of the assessments. However, the analysis of the aerodynamic measurements was more easily carried out, and the results were of immediate clinical value,

as indicated above, whereas the frequency analysis was much more time consuming, and the clinical interpretation of the results was very difficult. The use of this technique to provide a routine clinical assessment was not practicable.

Although this was the case, the frequency spectra which were obtained in the course of the investigation illustrated the spectral correlates of several types of abnormal vocal function, e.g. pitch breaks, voice breaks, modulations in fundamental frequency and intensity, and, in particular, the presence in many spectra of sub-harmonic frequency components. These were found to be not uncommon in the spectra obtained from dysphonic subjects and indicated the non-linearity of the voice production mechanism. The origin of such components was investigated and was reported in CHAPTER 5.

In CHAPTER 5, the non-invasive laryngograph technique was used to study the vibratory pattern of the vocal folds. Since the development of this technique was comparatively recent, it was essential initially to obtain normative data before the technique would be of clinical use. The lengths of the different phases of the laryngograph (Lx) waveform were measured, and it was found that the closing time was always much shorter than either the opening time or period of no contact, typically being less than 20% of the period. The closing time corresponded to the period of maximum excitation of the vocal tract.

The build up of vocal fold vibration during the initiation of phonation was studied, and it required up to about 16 cycles until a steady amplitude of vibration was obtained. It was demonstrated that male subjects required a longer time to attain a steady vibratory pattern than females, this being attributed to the vocal folds of males being more massive, with lower natural frequencies of vibration and thus having longer periods. In the majority of cases studied, the phonations were initiated and maintained at the same

fundamental frequency, indicating that the vocal fold adjustment did not depend solely upon feedback via the auditory pathway, but that other control systems were also important.

A study of the frequency spectra of both simulated and real Lx waveforms revealed the importance of the length of the closing time (CL.T.) of the vocal fold vibratory pattern to the resultant spectra. The general shape of the spectra could be described by a logarithmic relationship of the form :

$$I(f) = m \cdot \ln f + C \quad \begin{pmatrix} I(f) = \text{intensity of} \\ \text{harmonic component} \\ \text{of frequency } f \end{pmatrix}$$

so that the parameter $a = m \ln 2$ (db/octave) could be used to define the spectral gradient and also be independent of frequency. The value of a was found to be relatively independent of the opening time (O.T.) over the relevant range of values, but to be dependent on the closing time, and the value of the closing time could be estimated from a knowledge of the spectral gradient. However, the envelope curve of the spectral gradient was not monotonic, but consisted of a series of maxima and minima, giving the curve fine structure. It was predicted from the analysis of asymmetrical triangular waveforms and confirmed by experiment, that the closing time could also be predicted from a knowledge of the fine structure from the relationship :

$$\text{CL.T.} = \frac{1}{F'}$$

where F' = frequency of the first subsidiary minimum.

The value of the closing time is thought to be of clinical significance, especially in some cases of vocal misuse in which the vocal folds are considered to close too rapidly, leading to irritation at the points of contact and the development of nodules etc. However, such measurements should be combined with aerodynamic measurements and this aspect was beyond the scope of this thesis.

The frequency spectra obtained from dysphonic

subject were studied, and abnormal components in voice output were found to have correlates in the Lx spectra. This was especially true for subharmonic components, proving that such components originated from the vibration of the folds. This was investigated further and was found to be caused by a regular amplitude modulation in the Lx signal and therefore in the area of contact of the folds. This was interpreted as resulting from an asymmetrical adjustment between the folds and perhaps even in the folds having slightly different vibrational frequencies. The extreme case of a subject with a unilateral vocal fold paralysis was used to illustrate a simple model of the folds, and predict the variations in Lx output which were experimentally observed.

The relationship between sound level, air flow rate and the vibratory pattern of the vocal folds during a sustained phonation were discussed in CHAPTER 6. Phonations sustained at constant fundamental frequency and sound level could be maintained over a very large range of flow rates, and produced no significant results. However, when the sound level was deliberately varied, while the fundamental frequency was still maintained constant, a relation of the form :

$$SL = A\,(1 - k\,e^{-\alpha . FR})$$

was found to exist between the sound level (SL) and air flow rate (FR). The interpretation of this type of curve was that it represented an equilibrium locus for the type of phonation under study. Phonations could be initiated at a point some distance from this curve in the sound level versus flow rate space, but the locus was always approached as the phonations continued. This curve also served as a lower limit of air flow rate which could maintain a particular sound level, and in that sense also represented a locus of maximum efficiency in terms of air flow rate. However, in practice, this locus may not correspond to

maximum efficiency when the vocal fold vibratory pattern and its consequences in terms of fold pathology are considered.

Data points in the sound level versus flow rate space corresponding to sound level values greater than those of the locus were only found during periods of rapidly varying fundamental frequency, and may have arisen entirely from such effects. A detailed study involving the effect of fundamental frequency was not possible because of the difficulty in controlling this parameter.

Variations in the vocal fold vibratory pattern, especially in the closing time, and hence the spectral gradient, correlated well with changes in sound level and air flow rate, and indicated the fine control which existed in the relationship between these parameters.

The important factor which it was not possible to measure during these investigations was the subglottic pressure. The techniques for measuring this parameter involve either the use of a tracheal puncture to sense the pressure directly, or the use of an oesophageal balloon which, after correction, permits the tracheal pressure to be estimated. Both techniques were considered to be too invasive and were not used in the work described in this thesis.

The electromyographic (EMG) activity associated with the lip musculature during the phonation of a consonant-vowel-consonant utterance was studied, and the results presented in CHAPTER 7. Surface cup electrodes were used in the investigation. In general, it was found that there was lateral symmetry in the activity measured, whereas the activity measured from the superior and inferior electrode sites were significantly different. (Kelman and Gatehouse,1975). The measured activities were highly dependent upon the electrode configuration, i.e. bipolar or monopolar, (Gatehouse and Kelman,1976). The use of monopolar electrode arrays allowed the activity associated with various

muscle groups to be identified (Kelman and Gatehouse, 1976), and the results showed close agreement with those obtained by other investigators who had used needle electrodes. The discrepancies in the data reported by some experimenters in the forms of the outputs obtained during the phonation of the same utterance were also explained. Different electrode sites and configurations had been used, and thus recordings had been obtained from different muscle groups.

EMG recording is an important tool in phonetic and physiological research, and it is important to use non-invasive techniques wherever possible. The use of such recording techniques may have clinical value in determining laterality or cerebral dominance in neurological cases, but such studies remain to be carried out. The lip muscle EMG activity is already being used to provide an auditory feedback in the treatment of stammerers, and it is therefore important to be aware of the activity expected at particular muscle sites and also its reproducibility.

The study of voice pathology is of relatively recent origin, although some of the techniques used have been available for many years. The aims of the work described in this thesis were to add to the understanding of the voice production mechanism, which would, in turn, provide an insight into pathological conditions, and also to use modern technology to provide objective means of assessing vocal function. This should aid in the management of such diseases.

APPENDIX 1

APPENDIX 1

A1.1. ESTIMATION OF OPTIMAL LIMITS OF NORMALITY

Consider two groups of subjects (1 and 2) having N_1 and N_2 members respectively, with $N = N_1 + N_2$. Assume that each group can be classified into two subgroups according to the criterion A (and also NOT A, and define B = NOT A), and that the numbers in each subgroup are as shown in TABLE A1.1. It is desired to test the Null Hypothesis that both groups belong to the same population. If this were true, then the proportion of each group being classified according to A would be the same (within statistical limits). This is a Binomial problem, but the CHI^2 distribution can be used to test the significance of the differences obtained in the proportions, once criterion A has been applied. It is necessary to apply Yates' correction to allow for the fact that the Binomial distribution is discrete, whereas the CHI^2 distribution is continuous (Moroney, 1970). If the numbers given in TABLE A1.1 already incorporate Yates' correction, then the CHI^2 value is given by :

$$\chi^2 = \frac{\left(n_{1a} \cdot n_{2b} - n_{2a} \cdot n_{1b}\right)^2 N}{N_1 \cdot N_2 \cdot N_a \cdot N_b}$$

with 1 degree of freedom. The Null Hypothesis must be rejected at the appropriate level of significance if this value exceeds that obtained from tables of the CHI^2 distribution.

In the part of the present investigation described in CHAPTER 3, the two groups were the control group and the dysphonic group, and the criterion A was that "a given parameter had a value less than a limit, L." Then, for a range of values of L, the CHI^2 value was calculated and plotted against L, as shown in Figure A1.1. The Null Hypothesis, that both groups came from the same population, had to be rejected most strongly at the value of L corresponding to the maximum value of χ^2. However, if

	CRITERION A	CRITERION B (=NOT A)	TOTALS
GROUP 1	n_{1a}	n_{1b}	$N_1 = n_{1a} + n_{1b}$
GROUP 2	n_{2a}	n_{2b}	$N_2 = n_{2a} + n_{2b}$
TOTALS	$N_a = n_{1a} + n_{2a}$	$N_b = n_{1b} + n_{2b}$	$N = N_1 + N_2$

TABLE A1.1

RESULTS OF DIVISION OF TWO GROUPS
ACCORDING TO CRITERION A

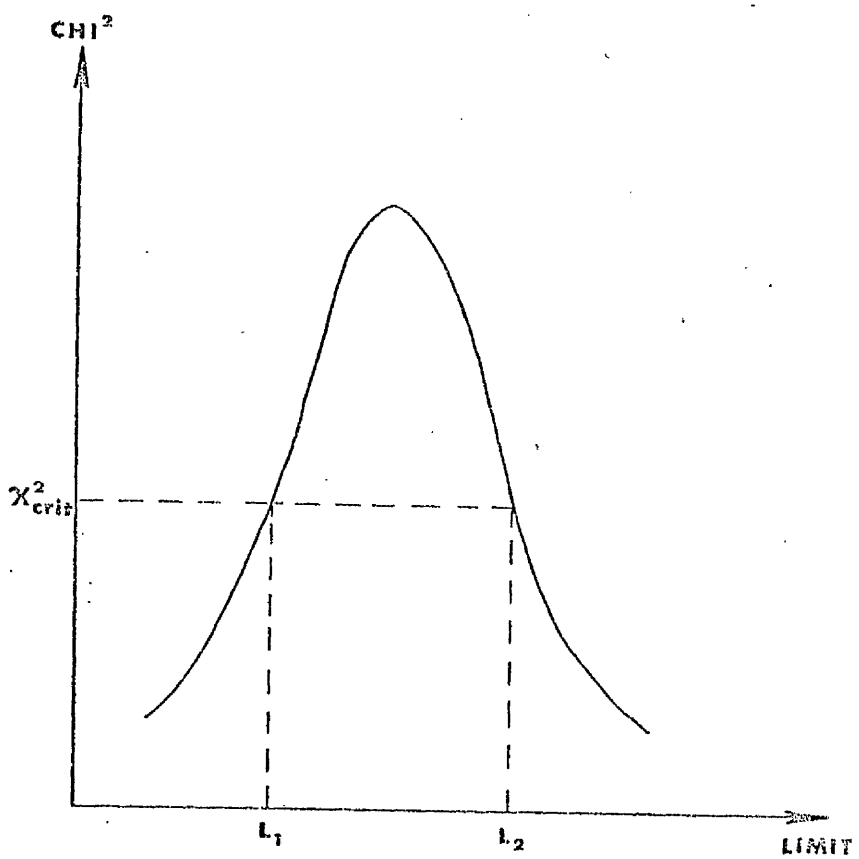


FIG.A1.1

Graph of CHI^2 value v limit value used for testing criterion A.

statistical significance was set at a particular level, corresponding to a χ^2 value, $\chi^2_{crit.}$, then the Null Hypothesis was rejected for all values of L in the range $L_1 < L < L_2$, and the most appropriate value of L selected with regard to other features of the investigation, e.g. the number of false negatives or false positives with which the clinical situation could cope. Thus, this technique used the CHI^2 coefficient as a variable with which to scan for the most appropriate value of the limit, L, at a given level of statistical significance. (Kirk, 1975).

A1.2 TEST OF INDEPENDENCE USING A CONTINGENCY TABLE.

Suppose that a group of N individuals can be classified according to two criteria, X and Y, and that there are r classes in X and s classes in Y. It is desired to test the Null Hypothesis that the classifications X and Y are independent. Let the data be plotted on a r x s contingency table, with the number of individuals belonging to class x_i and also y_j being denoted by n_{ij} (TABLE A1.2).

Thus $$N = \sum_{i,j} n_{ij}$$

Let the total number of subjects belonging to class x_i be T_{xi}, and the number belonging to y_j be T_{yj}, where:

$$T_{xi} = \sum_{j} n_{ij}$$

and

$$T_{yj} = \sum_{i} n_{ij}$$

An "independence value" can be calculated for each cell (i,j). This represents the number of individuals which would be expected in each cell if the criteria are independent. Let the independence

X \ Y	y_1	y_2	y_3	....	y_s	TOTALS
x_1	n_{11}	n_{12}	n_{13}	—	n_{1s}	T_{x1}
x_2	n_{21}	n_{22}	n_{23}	—	n_{2s}	T_{x2}
x_3	n_{31}	n_{32}	n_{33}	—	n_{3s}	T_{x3}
⋮	⋮	⋮	⋮		⋮	⋮
x_r	n_{r1}	n_{T2}	n_{r3}	....	n_{rs}	T_{xr}
TOTALS	T_{y1}	T_{y2}	T_{y3}	—	T_{ys}	N

TABLE A1.2

r x s CONTINGENCY TABLE FOR THE TWO CLASSIFICATIONS x AND y.

values be denoted by N_{ij}. The fraction of the total group belonging to class x_i is T_{xi}/N, and this applies for any size of population. In particular, it applies to a population of size T_{yj}, where the expected number also belonging to class x_i is $(T_{xi}/N).T_{yj}$. (Garret,1967). Thus, for all i, j :

$$N_{ij} = \frac{T_{xi} \cdot T_{yj}}{N} .$$

The CHI^2 coefficient is given by :

$$\chi^2 = \sum_{i,j} \frac{(n_{ij} - N_{ij})^2}{N_{ij}}$$

with $(r - 1).(S - 1)$ degrees of freedom. If the calculated CHI^2 value exceeds that obtained from tables for the required number of degrees of freedom and significance level, then the Null Hypothesis, that the classifications are independent, must be rejected.

The contingency coefficient is calculated from

$$C = \sqrt{\frac{\chi^2}{\chi^2 + N}}$$

However, the standard error in C is complex, and the most satisfactory method for testing the significance of a correlation is to test the χ^2 value, and this measure was used throughout the work described in this thesis.

APPENDIX 2

APPENDIX 2 - COMPUTER PROGRAMS.

A2.1 RESPIROMETRY PROGRAM.

Step	Prog.code	Key	
000	04 08	MARK	
001	10 00	1000	
002	04 12	WR ∝	
003	12 00	PRINTER ON	
004	01 08	CR/LF	
005	01 03	U/C	
006	01 04	I	
007	02 06	N	
008	02 07	T	INSPIRATION
009	01 14	V	TIDAL VOL.
010	01 02	L/C	CALCS.
011	04 13	END ∝	
012	07 00	0	
013	04 00	ST.DIR.	
014	00 05	REG.5	
015	00 01	SR 0001	
016	00 02	SR 0002	
017	07 01	1	
018	07 06	6	SD/M% LIMIT
019	07 12	.	
020	07 00	0	
021	00 03	SR 0003	
022	04 12	WR ∝	
023	01 03	U/C	
024	02 05	E	
025	02 15	X	
026	02 07	T	EXPIRATION
027	01 14	V	TIDAL VOL.
028	01 02	L/C	CALCS.
029	04 13	END ∝	
030	00 01	SR 0001	
031	00 02	SR 0002	
032	07 01	1	
033	07 06	6	SD/M% LIMIT
034	07 12	.	
035	07 00	0	
036	00 03	SR 0003	
037	04 12	WR ∝	
038	01 03	U/C	
039	02 07	T	
040	01 04	I	
041	01 15	M	
042	02 05	E	
043	01 01	S	
044	01 02	L/C	
045	01 08	CR/LF	
046	04 13	END ∝	
047	07 00	0	
048	04 04	ST.DIR.	PERIOD AND %
049	00 01	REG.1	INSPIRATION TIME
050	04 04	ST.DIR.	CALCS.

Step	Prog. code	Key	
051	00 02	REG.2	
052	04 04	ST.DIR.	
053	00 03	REG.3	
054	04 04	ST.DIR.	
055	00 06	REG.6	
056	04 04	ST.DIR.	
057	00 07	REG.7	
058	04 08	MARK	
059	15 03	1503	
060	05 15	STOP	INPUT
061	04 11	WR	INSPIRATION TIME
062	01 02	FORMAT	
063	06 04	↑	
064	04 11	WR	
065	15 02	2 SP.	
066	05 15	STOP	INPUT PERIOD
067	04 11	WR	
068	01 02	FORMAT	
069	04 00	+ DIR.	
070	00 01	REG.1	
071	06 03	÷	
072	07 13	x^2	
073	04 00	+ DIR.	
074	00 02	REG.2	
075	07 01	1	
076	04 00	+ DIR.	
077	00 03	REG.3	
078	06 05	↓	
079	04 12	WR	
080	07 02	2	
081	04 00	+ DIR.	
082	00 06	REG.6	
083	07 13	x^2	
084	04 00	+ DIR.	
085	00 07	REG.7	
086	04 07	SEARCH	
087	15 03	1503	
088	04 08	MARK	
089	07 01	1	
090	04 12	WR ∝	
091	01 08	CR/LF	
092	01 03	U/C	
093	00 05	P	
094	02 05	E	
095	01 13	R	
096	01 06	.	
097	01 02	L/C	
098	04 13	END ∝	
099	00 02	SR 0002	
100	07 01	1	
101	07 04	4	SD/M% LIMIT FOR
102	07 12	.	PERIOD
103	07 00	0	
104	00.03	SR 0003	

Step	Prog.code	Key	
105	04 12	WR ∝	
106	01 03	U/C	
107	03 05	%	
108	00 02	SP	
109	01 04	I	
110	02 06	N	
111	00 02	SP	
112	02 07	T	
113	01 02	L/C	
114	01 08	CR/LF	
115	04 13	END ∝	
116	07 01	1	
117	04 00	+ DIR.	
118	00 03	REG.3	
119	04 05	REC.DIR.	
120	00 06	REG.6	
121	04 04	ST.DIR.	
122	00 01	REG.1	
123	04 05	REC.DIR.	
124	00 07	REG.7	
125	04 04	ST.DIR.	
126	00 02	REG.2	
127	00 02	SR 0002	
128	07 01	1	
129	07 04	4	SD/M% LIMIT FOR
130	07 12	.	% INSP. TIME
131	07 00	0	
132	00 03	SR 0003	
133	04 12	WR ∝	
134	01 03	U/C	
135	02 07	T	
136	01 09	O	
137	02 07	T	
138	01 02	L/C	CALC. TOTAL
139	00 06	=	SCORE FOR QUIET
140	04 13	END∝	RESPIRATION
141	04 05	REC.DIR.	
142	00 05	REG.5	
143	04 11	WR	
144	01 01	FORMAT	
145	07 00	0	
146	04 04	ST.DIR.	
147	00 05	REG.5	
148	04 12	WR ∝	
149	01 08	CR/LF	
150	01 08	CR/LF	
151	01 03	U/C	
152	01 01	S	
153	00 14	F	
154	01 14	V	
155	02 12	C	CALC.MEASURED
156	01 02	L/C	VIT. CAP.
157	00 06	=	
158	04 13	END ∝	

Step	Prog.code	Key	
159	07 12	.	
160	07 08	8	
161	07 02	2	
162	07 06	6	
163	06 04	↑	
164	04 04	ST.DIR.	
165	00 00	REG.0	
166	05 15	STOP	INPUT SCALE FACTOR
167	06 02	x	FOR VIT.CAP. DATA
168	04 11	WR	
169	02 02	FORMAT	
170	04 11	WR	
171	15 04	4 SP	
172	04 12	WR ∝	
173	01 03	U/C	
174	01 14	V	
175	02 12	C	
176	01 02	L/C	
177	00 06	=	
178	04 13	END ∝	
179	05 15	STOP	INPUT VIT.CAP.
180	06 02	x	
181	06 05	↓	
182	04 04	ST.DIR.	
183	00 01	REG.1	
184	04 11	WR	
185	02 02	FORMAT	
186	04 12	WR ∝	
187	01 08	CR/LF	
188	01 03	U/C	
189	01 12	A	
190	00 15	G	INPUT DATA FOR CALC.
191	02 05	E	OF "BALDWIN'S"
192	01 02	L/C	VIT.CAP.
193	00 06	=	
194	04 13	END ∝	
195	05 15	STOP	INPUT AGE
196	04 11	WR	
197	02 00	FORMAT	
198	04 04	ST.DIR.	
199	00 02	REG.2	
200	04 11	WR	
201	15 04	4 SP	
202	04 12	WR ∝	
203	01 03	U/C	
204	02 01	H	
205	02 07	T	
206	01 02	L/C	
207	00 06	=	
208	04 13	END ∝	
209	05 15	STOP	INPUT HEIGHT
210	04 11	WR	
211	02 00	FORMAT	
212	04 04	ST.DIR.	
213	00 03	REG.3	
214	04 11	WR	

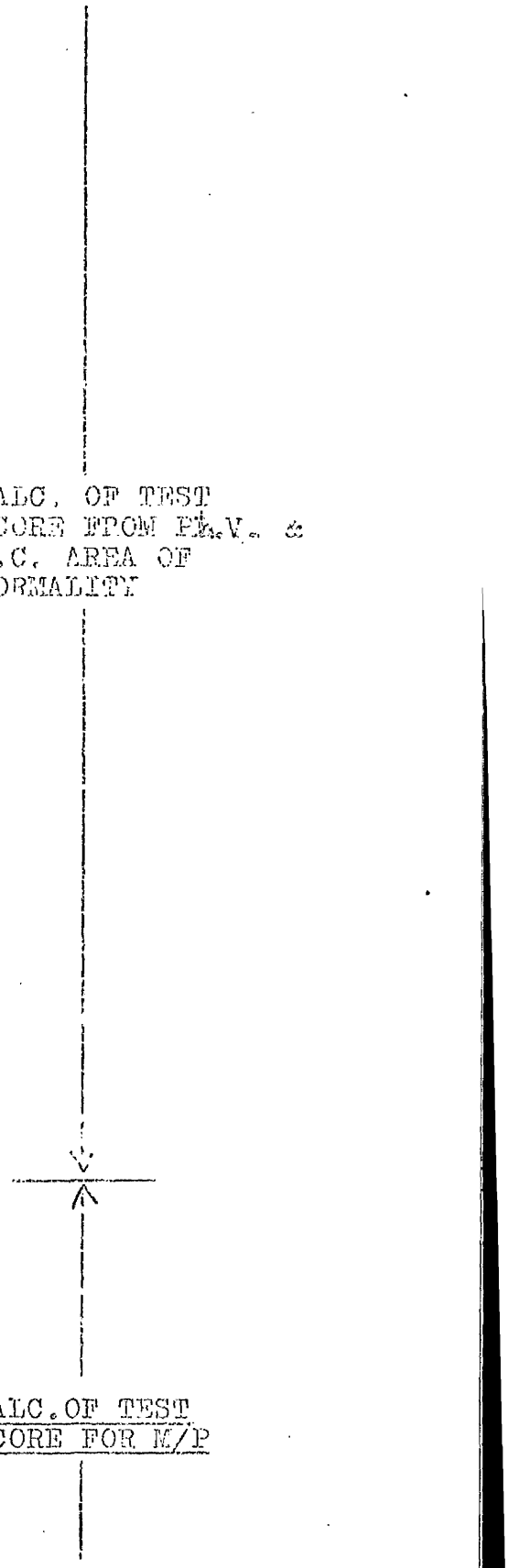
Step	Prog.code	Key	
215	15 06	6 SP	
216	04 12	WR ∝	
217	01 03	U/C	
218	02 00	B	
219	01 14	V	
220	02 12	C	
221	01 02	L/C	
222	00 06	=	
223	04 13	END ∝	
224	05 15	STOP	
225	04 08	MARK	
226	07 02	2	
227	07 02	2	
228	07 08	8	
229	04 12	)	$\div 10^5$
230	04 05	)	MALE SUBJECT
231	04 02	x DIR.	
232	00 02	REG.2	
233	07 12	.	
234	07 00	0	
235	07 07	7	
236	07 00	0	
237	07 01	1	
238	07 08	8	
239	06 04	↑	
240	04 05	REC.DIR.	
241	00 02	REG.2	
242	06 01	-	
243	04 07	SEARCH	
244	15 15	1515	
245	04 08	MARK	
246	07 03	3	
247	07 02	2	
248	07 06	6	FEMALE SUBJECT
249	04 12	)	$\div 10^5$
250	04 05	)	
251	04 02	x DIR.	
252	00 02	REG 2	
253	07 12	.	
254	07 00	0	
255	07 05	5	
256	07 05	5	
257	07 03	3	CALC. OF B.V.C.
258	07 02	2	
259	06 04	↑	
260	04 05	REC.DIR.	
261	00 02	REG.2	
262	06 01	-	
263	04 08	MARK	
264	15 15	1515	
265	04 05	RED.DIR.	
266	00 03	REG.3	

Step	Prog.code	Key	
267	06 02	x	
268	06 05	↓	
269	04 11	WR	
270	01 02	FORMAT	
271	04 11	WR	
272	15 06	6 SP	
273	04 12	WR ∝	
274	01 03	U/C	
275	01 14	V	
276	02 12	C	
277	01 02	L/C	
278	00 09	/	
279	01 03	U/C	CALC. VC/EVC
280	02 00	B	
281	01 14	V	
282	02 12	C	
283	01 02	L/C	
284	00 06	=	
285	04 13	END ∝	
286	04 15	REC.Y	
287	00 01	REG.1	
288	06 03	÷	
289	06 05	↓	
290	04 11	WR	
291	01 02	FORMAT	
292	04 11	WR	
293	15 04	4 SP	
294	04 12	WR ∝	
295	01 03	U/C	
296	01 01	S	
297	02 12	C	
298	01 02	L/C	
299	00 06	=	
300	04 13	END ∝	
301	07 12	.	
302	07 07	7	LIMIT FOR VC/EVC
303	06 06	↑↓	
304	00 03	SR 0003	
305	04 08	MARK	
306	07 04	4	
307	04 12	WR ∝	
308	01 08	CR/LF	
309	01 08	CR/LF	
310	04 13	END ∝	
311	04 11	WR	
312	15 15	15 SP	
313	04 12	WR ∝	
314	01 03	U/C	
315	00 05	P	
316	02 01	H	
317	02 07	T	
318	04 13	END ∝	
319	04 11	WR	

Step	Prog. Code	Key	
320	15 08	8 SP	
321	04 12	WR ∝	
322	00 05	P	
323	02 01	H	
324	01 14	V	
325	04 13	END ∝	
326	04 11	WR	
327	15 07	7 SP	
328	04 12	WR ∝	
329	01 15	M	
330	00 14	F	
331	01 13	R	
332	04 13	END ∝	
333	04 11	WR	
334	15 08	8 SP	
335	04 12	WR ∝	
336	00 05	P	
337	00 14	F	
338	01 13	R	PHON. HEADINGS
339	04 13	END ∝	
340	04 11	WR	
341	15 10	10 SP	
342	04 12	WR ∝	
343	01 15	M	
344	00 14	F	
345	01 13	R	
346	01 02	L/C	
347	00 09	/	
348	01 03	U/C	
349	00 05	P	
350	00 14	F	
351	01 13	R	
352	04 13	END ∝	
353	04 11	WR	
354	15 07	7 SP	
355	04 12	WR ∝	
356	01 14	V	
357	01 14	V	
358	01 04	I	
359	01 02	L/C	
360	01 08	CR/LF	
361	04 13	END ∝	
362	04 08	MARK	
363	07 05	5	
364	04 12	WR ∝	
365	01 08	CR/LF	
366	01 03	U/C	HEADING NP
367	02 06	N	
368	00 05	P	
369	01 02	L/C	
370	04 13	END ∝	
371	00 04	SR 0004	
372	04 07	SEARCH	

Step	Prog. Code	Key	
373	15 13	1513	
374	04 08	MARK	
375	07 06	6	
376	04 12	WR ∝	
377	01 08	CR/LF	
378	01 03	U/C	HEADING HP
379	02 01	H	
380	00 05	P	
381	01 02	L/C	
382	04 13	END ∝	
383	00 04	SR 0004	
384	04 07	SEARCH	
385	15 13	1513	
386	04 08	MARK	
387	07 07	7	
388	04 12	WR ∝	
389	01 08	CR/LF	
390	01 03	U/C	HEADING LP
391	02 09	L	
392	00 05	P	
393	01 02	L/C	
394	04 13	END ∝	
395	00 04	SR 0004	
396	04 08	MARK	
397	15 13	1513	
398	04 12	WR ∝	
399	01 04	i	
400	04 13	END ∝	
401	04 11	WR	
402	15 13	13 SP	
403	00 05	SR 0005	
404	04 12	WR ∝	PHON. DATA FOR
405	02 05	e	VOWELS
406	02 01	h	/i/, /ɛ/, /a/.
407	04 13	END ∝	
408	04 11	WR	
409	15 12	12 SP	
410	00 05	SR 0005	
411	04 12	WR ∝	
412	01 12	a	
413	04 13	END ∝	
414	04 11	WR	
415	15 13	13 SP	
416	00 05	SR 0005	
417	05 15	STOP	
418	04 08	MARK	
419	07 08	8	
420	04 12	WR ∝	
421	01 03	U/C	
422	00 05	P	
423	02 01	H	
424	01 14	V	
425	01 02	L/C	

Step	Prog. Code	Key	
426	04 13	END ∝	
427	04 05	REC. DIR.	
428	00 01	REG. 1	
429	04 15	REC. Y	
430	00 06	REG. 6	
431	06 00	+	
432	07 03	3	
433	07 12	.	
434	07 05	5	
435	05 07	SKIP Y ≠ X	
436	04 07	SEARCH	
437	15 11	1511	
438	04 15	REC. Y	
439	00 01	REG. 1	
440	07 01	1	
441	06 01	-	
442	07 12	.	CALC. OF TEST
443	07 06	6	SCORE FROM F.V. &
444	07 05	5	V.C. AREA OF
445	06 02	x	NORMALITY
446	04 05	REC. DIR.	
447	00 06	REG. 6	
448	05 07	SKIP Y ≠ X	
449	04 07	SEARCH	
450	15 10	1510	
451	04 08	MARK	
452	15 11	1511	
453	07 01	1	
454	04 00	+ DIR.	
455	00 05	REG. 5	
456	04 07	SEARCH	
457	15 09	1509	
458	04 08	MARK	
459	15 10	1510	
460	07 00	0	
461	04 00	+ DIR.	
462	00 05	REG. 5	
463	04 08	MARK	
464	15 09	1509	
465	00 06	SR 0006	
466	04 11	WR	
467	15 04	4 SP	
468	04 12	WR ∝	
469	01 03	U/C	
470	01 15	M	
471	01 02	L/C	
472	00 09	/	
473	01 03	U/C	CALC. OF TEST
474	00 05	P	SCORE FOR M/P
475	01 02	L/C	
476	04 13	END ∝	
477	04 15	REC. Y	



Step	Prog.Code	Key	
478	00 07	REG.7	
479	07 12	.	
480	07 06	6	LIMIT FOR M/P
481	07 05	5	
482	05 07	SKIP Y≥X	
483	04 07	SEARCH	
484	15 08	1508	
485	07 00	0	
486	04 00	+ DIR.	
487	00 05	REG.5	
488	04 07	SEARCH	
489	15 07	1507	
490	04 08	MARK	
491	15 08	1508	
492	07 01	1	
493	04 00	+ DIR.	
494	00 05	REG.5	
495	04 08	MARK	
496	15 07	1507	
497	00 06	SR 0006	
498	04 12	WR ∝	
499	01 03	U/C	
500	01 15	M	
501	00 14	F	
502	01 13	R	
503	01 02	L/C	
504	04 13	END ∝	
505	04 15	REC.	
506	00 04	REG.4	
507	07 05	5	LOWER LIMIT
508	05 07	SKIP Y≥X	FOR MFR
509	04 07	SEARCH	
510	15 06	1506	
511	07 01	1	UPPER LIMIT
512	07 01	1	FOR MFR
513	05 08	SKIP Y≥X	
514	04 07	SEARCH	
515	15 06	1506	
516	07 00	0	
517	04 00	+ DIR.	
518	00 05	REG.5	
519	04 07	SEARCH	CALC. OF TEST
520	15 05	1505	SCORE FOR MFR
521	04 08	MARK	
522	15 06	1506	
523	07 01	1	
524	04 00	+ DIR.	
525	00 05	REG.5	
526	04 08	MARK	
527	15 05	1505	
528	00 06	SR 0006	
529	04 12	WR ∝	
530	01 03	U/C	

Step	Prog. Code	Key	
531	02 07	T	
532	01 09	O	
533	02 07	T	
534	01 02	L/C	CALC. OF TOTAL
535	00 06	=	SCORE FOR
536	04 13	END ∝	PHON. TESTS
537	04 05	REC. DIR.	
538	00 05	REG. 5	
539	04 11	WR	
540	01 01	FORMAT	
541	04 12	WR ∝	
542	01 08	CR/LF	
543	04 13	END ∝	
544	05 15	STOP	
545	04 08	MARK	
546	00 01	0001	
547	04 11	WR	
548	15 04	4 SP	
549	04 12	WR ∝	
550	01 03	U/C	
551	01 01	S	
552	00 14	F	
553	01 02	L/C	
554	00 06	=	
555	04 13	END ∝	
556	05 15	STOP	INPUT VOL. SCALE
557	04 11	WR	FACTOR
558	02 02	FORMAT	
559	06 04	↑	SUBROUTINE
560	04 11	WR	0001
561	15 04	4 SP	
562	04 12	WR ∝	
563	01 03	U/C	
564	02 07	T	
565	02 12	C	
566	01 02	L/C	
567	00 06	=	
568	04 13	END ∝	
569	05 15	STOP	INPUT TEMP.
570	04 11	WR	CORRECTION
571	01 03	FORMAT	
572	06 02	X	
573	04 14	ST Y	
574	00 00	REG. 0	
575	04 12	WR ∝	
576	01 08	CR/LF	
577	04 13	END ∝	
578	07 00	0	
579	04 04	ST. DIR.	
580	00 01	REG. 1	
581	04 04	ST. DIR.	
582	00 02	REG. 2	
583	04 04	ST. DIR.	

Step	Prog. Code	Key	
584	00 03	REG.3	
585	04 08	MARK	
586	15 00	1500	
587	05 15	STOP	INPUT TIDAL VOL.
588	04 11	WR	DATA
589	02 03	FORMAT	
590	04 15	REC.Y	
591	00 00	REG.0	
592	06 02	X	
593	06 05	↓	
594	04 00	+ DIR.	
595	00 01	REG.1	
596	07 13	X^2	
597	04 00	+ DIR.	
598	00 02	REG.2	
599	07 01	1	
600	04 00	+ DIR.	
601	00 03	REG.3	
602	04 11	WR	
603	15 02	2 SP	
604	04 07	SEARCH	
605	15 00	1500	
606	04 08	MARK	
607	07 00	0	
608	05 11	RETURN	
609	04 08	MARK	
610	00 02	0002	
611	04 12	WR ∝	
612	01 08	CR/LF	
613	02 06	M	
614	00 06	=	
615	04 13	END ∝	
616	04 05	REC.DIR.	
617	00 03	REG.3	
618	04 11	WR	
619	02 00	FORMAT	
620	04 11	WR	
621	15 04	4 SP	
622	04 12	WR ∝	
623	01 03	U/C	
624	01 15	M	
625	01 02	L/C	SUBROUTINE
626	00 06	=	0002
627	04 13	END ∝	
628	04 15	REC.Y	
629	00 01	REG.1	
630	06 03	÷	CALC.MEAN
631	06 05	↓	
632	04 11	WR	
633	02 03	FORMAT	
634	04 04	ST.DIR.	
635	00 04	REG.4	

Step	Prog. Code	Key	
636	04 15	REC.Y	
637	00 01	REG.1	
638	06 02	X	
639	04 05	REC.DIR.	
640	00 02	REG.2	
641	06 06	↑↓	
642	06 01	−	
643	07 01	1	
644	04 01	− DIR.	
645	00 03	REG.3	
646	04 05	REC.DIR.	
647	00 03	REG.3	
648	06 03	÷	
649	06 05	↓	
650	06 12	$\sqrt{x}$	
651	04 11	WR	
652	15 04	4 SP	
653	04 12	WR ∝	
654	01 03	U/C	CALC. S.D.
655	01 01	S	
656	02 13	D	
657	01 02	L/C	
658	00 00	=	
659	04 13	END ∝	
660	04 11	WR	
661	02 03	FORMAT	
662	04 11	WR	SUBROUTINE
663	15 04	4 SP	0002
664	04 12	WR ∝	
665	01 03	U/C	
666	01 01	S	
667	02 13	D	
668	01 02	L/C	
669	00 09	/	
670	01 03	U/C	
671	01 15	M	
672	03 05	%	
673	01 02	L/C	
674	00 06	=	
675	04 13	END ∝	
676	06 04	↑	CALC.SD/M%
677	04 05	REC.DIR.	
678	00 04	REG.4	
679	06 03	÷	
680	06 05	↓	
681	04 12	)	
682	07 02	) x 100	
683	04 11	WR	
684	02 01	FORMAT	
685	06 04	↑	
686	04 11	WR	
687	15 04	4 SP	

Step	Prog.Code	Key	
688	04 12	WR ∝	
689	01 03	U/C	
690	01 01	S	
691	02 12	C	
692	01 02	L/C	
693	00 06	=	
694	04 13	END ∝	
695	05 11	RETURN	
696	04 08	MARK	
697	00 03	0003	
698	05 07	SKIP Y ≥ X	
699	04 07	SEARCH	
700	15 01	1501	
701	07 01	1	
702	04 00	+ DIR.	CALC.SCORE(0 or 1)
703	00 05	REG.5	
704	04 07	SEARCH	
705	15 02	1502	
706	04 08	MARK	
707	15 01	1501	
708	07 00	0	SUBROUTINE
709	04 00	+ DIR.	0003
710	00 05	REG.5	
711	04 08	MARK	
712	15 02	1502	
713	04 11	WR	
714	01 01	FORMAT	
715	04 12	WR ∝	
716	01 08	CR/LF	
717	01 08	CR/LF	
718	04 13	END ∝	
719	05 11	RETURN	
720	04 08	MARK	
721	00 04	0004	
722	04 12	WR ∝	
723	01 08	CR/LF	
724	01 03	U/C	
725	01 01	S	
726	00 14	F	
727	01 14	V	
728	01 02	L/C	
729	00 06	=	
730	04 13	END ∝	
731	05 15	STOP	INPUT SCALE
732	04 11	WR	FACTOR PHON.VOL.
733	02 02	FORMAT	
734	04 15	REC. Y	
735	00 00	REG. 0	
736	06 02	X	
737	04 14	ST.Y	
738	00 02	REG.2	
739	04 12	WR ∝	
740	01 08	CR/LF	

Step	Prog.Code	Key	
741	01 03	U/C	
742	01 01	S	
743	00 14	F	
744	00 14	F	SUBROUTINE
745	01 13	R	0004
746	01 02	L/C	
747	00 06	=	
748	04 13	END ∝	
749	05 15	STOP	INPUT SCALE
750	04 11	WR	FACTOR FLOW RATE
751	03 01	FORMAT	
752	04 15	REC. Y	
753	00 00	REG. 0	
754	06 02	X	
755	04 14	ST. Y	
756	00 03	REG.3	
757	04 12	WR ∝	
758	01 08	CR/LF	
759	04 13	END ∝	
760	05 11	RETURN	
761	04 08	MARK	
762	00 05	0005	
763	05 15	STOP	INPUT PHON.TIME
764	04 11	WR	
765	02 02	FORMAT	
766	04 04	ST.DIR.	
767	00 04	REG.4	
768	04 11	WR	
769	15 05	5 SP.	
770	05 15	STOP	INPUT PHON.VOL.
771	04 15	REC. Y	
772	00 02	REG. 2	
773	06 02	X	
774	06 05	↓	
775	04 11	WR	
776	01 03	FORMAT	
777	04 11	WR	
778	15 04	4 SP.	
779	04 04	ST.DIR.	
780	00 06	REG.6	
781	06 04	↑	SUBROUTINE
782	04 05	REC.DIR.	0005
783	00 04	REG.4	
784	06 03	÷	
785	07 06	6	
786	07 00	0	
787	06 02	X	CALC.MFR
788	06 05	↓	
789	04 11	WR	
790	02 02	FORMAT	
791	04 04	ST.DIR.	
792	00 04	REG.4	

Step	Prog. Code	Key	
793	04 11	WR	
794	15 05	5 SP.	
795	05 15	STOP	INPUT PFR
796	04 15	REC. Y	
797	00 03	REG. 3	
798	06 02	X	
799	06 05	↓	
800	04 11	WR	
801	02 02	FORMAT	
802	04 11	WR	
803	15 07	7 SP.	
804	04 15	REC. Y	
805	00 04	REG. 4	
806	06 03	÷	
807	06 05	↓	CALC. MFR/PFR
808	04 11	WR	
809	01 03	FORMAT	
810	04 04	ST. DIR.	
811	00 07	REG. 7	
812	04 11	WR	
813	15 06	6 SP.	
814	04 05	REC. DIR.	
815	00 01	REG. 1	
816	04 15	REC. Y	
817	00 04	REG. 4	
818	06 03	÷	
819	06 05	↓	
820	04 11	WR	CALC. VOCAL VEL.
821	03 03	FORMAT	INDEX (VVI)
822	04 12	WR ∝	
823	01 08	CR/LF	
824	04 13	END ∝	
825	05 11	RETURN	
826	04 08	MARK	
827	00 06	0006	
828	04 12	WR ∝	
829	01 03	U/C	
830	00 0?	SP.	
831	01 02	S	SUBROUTINE
832	02 12	C	0006
833	01 02	L/C	
834	00 06	=	
835	04 13	END ∝	SCORE FOR EACH
836	04 11	WR	PHON. TEST
837	01 01	FORMAT	
838	04 11	WR	
839	15 04	4 SP.	
840	05 11	RETURN	
841	05 15	STOP	
842	05 12	END PROG.	

A2.2

FOURIER SERIES PROGRAM

Let the periodic function Y(t) satisfy the continuity conditions for Fourier Analysis to be valid, and let one cycle of Y(t) be defined by 2n+1 consecutive, equally spaced data points $(Y_0, Y_1, Y_2 \ldots Y_{2n})$. Then Y(t) can be described by a summation of trigonometric terms, such that

$$Y(t) = A_o + \sum_{r=o}^{2n} (A_r \cos r\Theta + B_r \sin r\Theta)$$

where $$A_o = \frac{1}{2n+1} \cdot \sum_{s=o}^{2n} Y_s = \overline{Y} \quad ,$$

$$A_r = \frac{2}{2n+1} \cdot \sum_{s=o}^{2n} Y_s \cos rS\Theta \quad ,$$

$$B_r = \frac{2}{2n+1} \cdot \sum_{s=1}^{2n} Y_s \sin rS\Theta \quad ,$$

and $$\Theta = \frac{2\pi}{2n+1}$$

Let $m = 2n+1$, then

$$A_r = \frac{2}{m} \cdot \sum_{s=o}^{m-1} Y_s \cos rS\Theta$$

and $$B_r = \frac{2}{m} \cdot \sum_{s=1}^{m-1} Y_s \sin rS\Theta = \frac{2}{m} \cdot \sum_{s=o}^{m-1} Y_s \sin rS\Theta$$

Let $t = s + 1$ i.e. $s = t-1$

Then
$$A_r = \frac{2}{m} \sum_{t=1}^{m} Y_{t-1} \cos r(t-1)\,\theta \quad \ldots\ldots(A2.1)$$

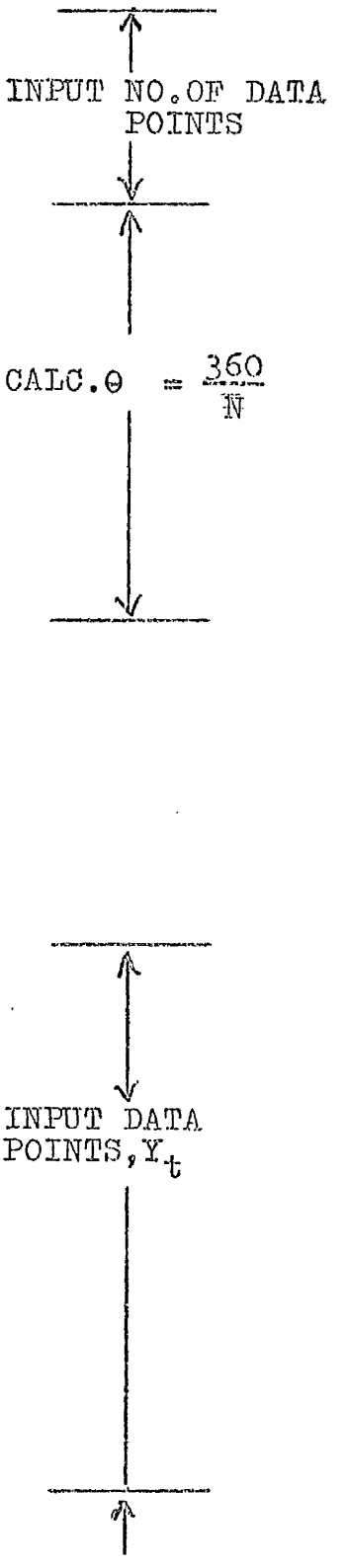
And
$$B_r = \frac{2}{m} \sum_{t=1}^{m} Y_{t-1} \sin r(t-1)\,\theta \quad \ldots\ldots(A2.2)$$

The intensity , I_r, of the r^{th} harmonic component is given by

$$I_r = A_r^2 + B_r^2 \ .$$

In the computer program the A_r, B_r coefficients were calculated from the forms given in eqns (A2.1) and (A2.2)

Step	Prog. Code	Key
000	04 08	MARK
001	01 00	0100
002	04 12	WR ∝
003	12 00	PRINTER ON
004	01 08	CR/LF
005	01 02	L/C
006	02 06	N
007	00 02	SPACE
008	04 13	END ∝
009	05 15	STOP
010	04 11	WR
011	02 00	FORMAT
012	04 12	WR ∝
013	01 08	CR/LF
014	04 13	END ∝
015	04 04	ST.DIR.
016	06 09	REG. 69
017	06 04	↑
018	07 03	3
019	07 06	6
020	07 00	0
021	06 06	↑↓
022	06 03	÷
023	04 14	ST. Y
024	00 00	REG. 0
025	07 00	0
026	06 04	↑
027	04 04	ST.DIR.
028	06 01	REG. 61
029	04 04	ST.DIR.
030	06 03	REG. 63
031	04 04	ST.DIR.
032	07 00	REG.70
033	04 04	ST.DIR.
034	07 01	REG. 71
035	04 08	MARK
036	15 00	1500
037	07 01	1
038	06 00	+
039	05 15	STOP
040	05 04	ST.INDIR.
041	04 11	WR
042	02 04	FORMAT
043	04 11	WR
044	15 02	2 SP.
045	04 00	+ DIR.
046	06 01	REG.61
047	04 05	REC.DIR.
048	06 09	REG.6
049	05 09	SKIP IF Y=X
050	04 07	SEARCH
051	15 00	1500
052	04 15	REC. Y



Step	Prog. Code	Key	
053	06 01	REG. 61	
054	06 03	÷	
055	06 05	↓	
056	04 12	WR ∝	
057	01 08	CR/LF	
058	01 08	CR/LF	
059	00 02	SP	
060	00 02	SP	CALC. A_o
061	01 03	U/C	
062	01 12	A	
063	01 02	L/C	
064	03 01	0	
065	04 13	END ∝	
066	04 11	WR	
067	15 04	4 SP	
068	04 11	WR	
069	02 04	FORMAT	
070	04 11	WR	
071	15 15	15 SP	
072	07 13	x^2	CALC. A_o^2
073	04 12	WR ∝	
074	01 03	U/C	
075	01 12	A	
076	01 12	A	
077	01 02	L/C	
078	04 13	END ∝	
079	04 11	WR	
080	15 04	4 SP	
081	04 11	WR	
082	02 04	FORMAT	
083	04 11	WR	
084	15 05	5 SP.	
085	04 12	WR ∝	
086	01 03	U/C	
087	01 04	I	
088	01 02	L/C	
089	00 02	SP.	
090	00 02	SP.	
091	04 13	END ∝	
092	04 12	WR ∝	
093	07 03	3	
094	06 10	LOG X	CALC. I_o
095	04 12	WR ∝	
096	07 01	1	
097	04 11	WR	
098	04 04	FORMAT	
099	04 12	WR	
100	01 08	CR/LF	
101	04 13	END ∝	
102	04 08	MARK	
103	15 01	1501	
104	07 01	1	

Step	Prog. Code	Key
105	04 00	+ DIR.
106	06 03	REG.63 ——— r
107	07 00	0
108	04 04	ST.DIR.
109	06 02	REG.62
110	04 04	ST.DIR.
111	06 05	REG.65
112	04 04	ST.DIR.
113	06 08	REG.68
114	04 08	MARK
115	15 02	1502
116	07 01	1
117	04 00	+ DIR.
118	06 02	REG.62 ——— t
119	04 15	REC. Y
120	06 02	REG. 62
121	06 01	-
122	04 05	REC.DIR.
123	00 00	REG. 0
124	06 02	X
125	04 05	REC.DIR.
126	06 03	REG.63
127	06 02	X
128	06 05	↓
129	04 04	ST.DIR.
130	06 06	REG.66 ——— r(t-1)θ
131	00 03	SR.0003
132	04 04	ST.DIR.
133	06 04	REG.64 ——— cos r(t-1)θ
134	04 15	REC.Y
135	06 02	REG.62
136	05 05	REC.INDIR.
137	04 15	REC.Y
138	06 04	REG.64
139	06 02	X
140	06 05	↓
141	04 00	+ DIR.
142	06 05	REG.65 ——— $\sum y_{t-1} \cos r(t-1)\theta$
143	04 05	REC.DIR.
144	06 06	REG.66
145	00 02	SR.0002
146	04 04	ST.DIR.
147	06 07	REG.67 ——— sin r(t-1)θ
148	04 15	REC.Y
149	06 02	REG.62
150	05 05	REC.INDIR.
151	04 15	REC.Y
152	06 07	REG.67
153	06 02	X
154	06 05	↓
155	04 00	+ DIR.
156	06 08	REG.68 ——— $\sum y_{t-1} \sin r(t-1)\theta$

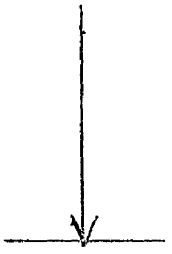
Step	Prog. Code	Key	
157	04 05	REC.DIR	
158	06 09	REG.69	
159	04 15	REC.Y	CALC. A_r, B_r
160	06 02	REG.62	
161	05 09	SKIP IF Y=X	
162	04 07	SEARCH	
163	15 02	1502	
164	07 02	2	
165	06 03	÷	
166	06 05	↓	
167	04 03	÷ DIR.	
168	06 05	REG.65 ——— A_r	
169	04 03	DIR.	
170	06 08	REG.68 ——— B_r	
171	04 12	WR.∝	
172	00 02	SP.	
173	00 02	SP.	
174	01 03	U/C	
175	01 12	A	
176	01 02	L/C	
177	04 13	END∝	
178	04 05	REC.DIR.	
179	06 03	REG.63	
180	04 11	WR.	
181	02 00	FORMAT	
182	04 11	WR.	
183	15 01	1 SP.	
184	04 05	REC.DIR.	
185	06 05	REG.65	
186	04 11	WR.	
187	02 04	FORMAT	
188	07 13	X^2 ——— A_r^2	
189	06 04	↑	
190	04 11	WR.	
191	15 03	3 SP.	
192	04 12	WR.∝	
193	01 03	U/C	
194	02 00	B	
195	01 02	L/C	
196	00 02	SP.	
197	00 02	SP.	
198	04 13	END∝	
199	04 05	REC.DIR.	CALC. $A_r^2 + B_r^2$
200	06 08	REG.68	
201	04 11	WR.	
202	02 04	FORMAT	
203	04 11	WR.	
204	15 02	2 SP.	
205	07 13	X^2 ——— B_r^2	
206	06 00	+	
207	06 05	↓	
208	04 12	WR.∝	

Step	Prog. Code	Key	
209	01 03	U/C	
210	01 12	A	
211	01 12	A	
212	00 06	+	
213	02 00	B	
214	02 00	B	
215	01 02	L/C	
216	00 02	SP	
217	04 13	END ∝	
218	04 11	WR.	
219	02 04	FORMAT	
220	04 11	WR.	
221	15 05	5 SP.	
222	04 12	WR. ∝	
223	01 03	U/C	
224	01 04	I	
225	01 02	L/C	
226	04 13	END ∝	
227	04 12	WR. ∝	SCALING FACTOR
228	07 03	3	
229	06 10	LOG_{10} X	
230	04 12	WR. ∝	
231	07 01	1	
232	04 11	WR.	CALC. I_r
233	04 04	FORMAT	
234	04 12	WR. ∝	
235	01 08	CR/LF	
236	04 13	END ∝	
237	04 15	REC.Y	
238	06 03	REG.63	
239	07 01	1	
240	07 05	5	
241	05 09	SKIP IF Y=X	
242	04 07	SEARCH	
243	15 01	1501	
244	04 07	SEARCH	
245	15 03	1503	
246	04 08	MARK	
247	00 02	SR.0002	
248	06 04	↑	SUBROUTINE 0002
249	07 09	9	(SINE X)
250	07 00	0	
251	06 01	-	
252	06 05	↓	
253	04 08	MARK	
254	00 03	0003	
255	06 04	↑	
256	07 03	3	
257	07 06	6	
258	07 00	0	
259	06 03	÷	
260	06 05	↓	
261	06 08	INT. X	
262	06 01	-	
263	07 04	4	SUBROUTINE 0003

Step	Prog. Code	Key
264	06 02	X
265	06 05	↓
266	06 08	INT.X
267	06 01	−
268	04 12	WR.∝
269	06 12	$\sqrt{x}$
270	06 09	$\overline{\pi}$
271	06 02	X
272	07 02	2
273	06 03	÷
274	06 05	↓
275	07 13	x^2
276	04 04	ST.DIR.
277	07 00	REG.70
278	07 01	1
279	07 06	6
280	06 04	↑
281	07 01	1
282	04 04	ST.DIR.
283	07 01	REG.71
284	04 08	MARK
285	15 14	1514
286	04 05	REC.DIR.
287	07 00	REG.70
288	04 02	X DIR.
289	07 01	REG.71
290	06 05	↓
291	04 03	÷ DIR.
292	07 01	REG.71
293	07 01	1
294	06 01	−
295	06 05	↓
296	07 11	CH.SIGN
297	04 03	÷ DIR.
298	07 01	REG.71
299	07 01	1
300	06 01	−
301	04 00	+ DIR.
302	07 01	REG.71
303	04 12	WR.∝
304	04 11	WR.
305	04 07	SEARCH
306	15 14	1514
307	04 15	REC.Y
308	07 01	REG.71
309	07 12	.
310	07 05	5
311	07 10	SET EXP.
312	07 11	CH.SIGN
313	07 01	1

SUBROUTINE
0003
(COSINE X)

Step	Prog. Code	Key
314	07 01	1
315	06 01	-
316	06 01	-
317	06 05	↓
318	04 12	WR.∝
319	05 12	END PROG.
320	05 11	RETURN
321	04 08	MARK
322	15 03	1503
323	05 15	STOP
324	05 12	END PROG.



A2.3 RESPIROMETRY PROGRAM (MK II) (FORTRAN IV)

```
      1 FORMAT('RESPIROMETRY PROGRAM')
      2 FORMAT('QUIET RESPIRATION')
      3 FORMAT('INSPIR. TIDAL VOL.')
      4 FORMAT('EXPIR. TIDAL VOL.')
      5 FORMAT('TEMPORAL DATA')
      6 FORMAT('PERIOD')
      7 FORMAT('INSPIR. FRACTION')
      8 FORMAT('MEAN=',F6.3,4X,'SD=',F6.3,4X,'SD/MN=',F5.2,
     1 14X,'SCORE=',F3.1)
      9 FORMAT(I2)
     10 FORMAT('NO.OF DATA POINTS=',I2)
     11 FORMAT('QUIET RESPIRATION  TOTAL SCORE=',F3.1)
     12 FORMAT(10X,'PHON.T',3X,'PHON.V',3X,'MFR',3X,'PFR',
     1 13X,'MFR/PFR',3X,'VVI')
     13 FORMAT(F4.2,1X,F4.2)
     14 FORMAT('SEVC=',F4.2,4X,'MVC=',F4.2,4X,'VITAL CAP.=',
     1 F4.2)
     15 FORMAT('AGE=',I2,4X,'HT=',I2,4X,'BVC=',F4.2,4X,'VC/BVC
     1=',F4.2,4X,'SCORE=',F3.1)
     16 FORMAT('NORMAL PITCH')
     17 FORMAT('HIGH PITCH')
     18 FORMAT('LOW PITCH')
     19 FORMAT(F4.2,1X,F5.2)
     20 FORMAT('SFV=',F4.2,4X,'SFFR=',F5.1)
     21 FORMAT('I')
     22 FORMAT('E')
     23 FORMAT('A')
     24 FORMAT(I2,1X,I2,1X,I1)
     27 FORMAT('PH.V.SCORE=',F3.1,4X,'MFR/PFR SCORE=',F3.1,
     1 14X,'MFR SCORE=',F3.1)
     28 FORMAT('PHONATION-(A)  TOTAL SCORE=',F3.1)
        INTEGER AGE,HT
        REAL MP,MFR,MPHONV,MPFR,MVC
        DIMENSION X(24),T(24)
        COMMON X,T,SP,TC,XM,SD,PC,SFV,SFFR,SCORE,PHONV,MP,
     1 MFR,VC,J
        WRITE(6,1)
        WRITE(6,2)
        READ(1,9) J
        WRITE(6,10) J
        WRITE(6,3)
        CALL SCALE1
        CALL DATA1
        CALL MEAN
```

```
      J
>     45          CALL TEST1
>     46          RESULT=SCORE
>     47          XX=X1*SI*TC
>     48          S=SD*SI*TC
>     49          WRITE(6,8) XX,S,PC,SCORE
>     50          WRITE(6,4)
>     51          CALL SCALE1
>     52          CALL DATA1
>     53          CALL MEAN
>     54          CALL TEST1
>     55          RESULT=RESULT+SCORE
>     56          XX=X1*SI*TC
>     57          S=SD*SI*TC
>     58          WRITE(6,8) XX,S,PC,SCORE
>     59          WRITE(6,5)
>     60          WRITE(6,6)
>     61          CALL DATA1
>     62          CALL MEAN
>     63          CALL TEST2
>     64          WRITE(6,8) X1,SD,PC,SCORE
>     65          RESULT=RESULT+SCORE
>     66          DO 30 I=1,J
>     67       30 T(I)=X(I)
>     68          CALL DATA1
>     69          DO 31 I=1,J
>     70       31 X(I)=X(I)/T(I)
>     71          CALL MEAN
>     72          CALL TEST2
>     73          X1=100*X1
>     74          SD=100*SD
>     75          WRITE(6,8) X1,SD,PC,SCORE
>     76          RESULT=RESULT+SCORE
>     77          WRITE(6,11) RESULT
>     78          READ(1,13) SFVC,AVC
>     79          VC=AVC*SFVC*TC
>     80          WRITE(6,14) SFVC,AVC,VC
>     81          READ(1,24) AGE,HT,L
>     82          IF(L.EQ.1)GO TO 35
>     83          BVC=(0.05532-0.00026*AGE)*HT
>     84          GO TO 40
>     85       35 BVC=(0.07018-0.00028*AGE)*HT
>     86       40 RVC=VC/BVC
>     87          IF(RVC.GE.0.7) GO TO 41
>     88          SCORE=1.0
>     89          GO TO 42
>     90       41 SCORE=0.0
>     91       42 WRITE(6,15) AGE,HT,BVC,RVC,SCORE
>     92          WRITE(6,12)
>     93          WRITE(6,16)
>     94          READ(1,19) SFV,SFFR
>     95          WRITE(6,20) SFV,SFFR
>     96          WRITE(6,21)
>     97          CALL PAU1
>     98          WRITE(6,22)
>     99          CALL PAU1
>    100          WRITE(6,23)
```

```
      CALL PHON
      U=PHOTV+VC
      IF(U.LT.3.5) GO TO 45
      V=0.64*(VC-1.0)
      IF(PHOTV.LT.V) GO TO 45
      SC1=0.0
      GO TO 50
   45 SC1=1.0
   50 IF(AP.LT.0.65) GO TO 55
      SC2=0.0
      GO TO 60
   55 SC2=1.0
   60 IF(SFR.GT.11.0) GO TO 65
      IF(SFR.LT.5.0) GO TO 65
      SC3=0.0
      GO TO 70
   65 SC3=1.0
   70 RESULT=SCORE+SC1+SC2+SC3
      WRITE(6,27) SC1,SC2,SC3
      WRITE(6,28) RESULT
      WRITE(6,17)
      READ(1,19) SFV,SFFR
      WRITE(6,20) SFV,SFFR
      WRITE(6,21)
      CALL PHON
      WRITE(6,22)
      CALL PHON
      WRITE(6,23)
      CALL PHON
      WRITE(6,18)
      READ(1,19) SFV,SFFR
      WRITE(6,20) SFV,SFFR
      WRITE(6,21)
      CALL PHON
      WRITE(6,22)
      CALL PHON
      WRITE(6,23)
      CALL PHON
      END
      SUBROUTINE SCALE1
      DIMENSION X(24),T(24)
      COMMON X,T,SF,TC,XM,SD,PC,SFV,SFFR,SCORE,PHOTV,AP,
     1AFR,VC,J
      READ(1,75) SF,TC
      WRITE(6,80) SF,TC
   75 FORMAT(F5.2,1X,F5.3)
   80 FORMAT('SC.F=',F5.2,4X,'TEMP.CORR.=',F5.3)
      RETURN
      END
      SUBROUTINE DATA1
```

```
      DIMENSION X(24),T(24)
      COMMON X,T,SF,TC,XM,SD,PC,SFV,SHFR,SCORE,PHONV,MP,
     1MFR,VC,N
   90 FORMAT(12(F4.2,1X))
   91 FORMAT(1X,12(F4.2,1X))
      IF(N.GT.12) GO TO 85
      READ(1,90) (X(I),I=1,N)
      WRITE(6,91) (X(I),I=1,N)
      GO TO 95
   85 READ(1,90) (X(I),I=1,12)
      WRITE(6,91) (X(I),I=1,12)
      READ(1,90) (X(I),I=13,N)
      WRITE(6,91) (X(I),I=13,N)
   95 RETURN
      END
      SUBROUTINE MEAN
      DIMENSION X(24),T(24)
      COMMON X,T,SF,TC,XM,SD,PC,SFV,SHFR,SCORE,PHONV,MP,
     1MFR,VC,N
      A=0.0
      B=0.0
      DO 100 I=1,N
      A=A+X(I)
  100 B=B+X(I)**2
      XM=A/N
      SDD=(B-A**2/N)/(N-1)
      SD=SQRT(SDD)
      PC=100*SD/XM
      RETURN
      END
      SUBROUTINE TEST1
      DIMENSION X(24),T(24)
      COMMON X,T,SF,TC,XM,SD,PC,SFV,SHFR,SCORE,PHONV,MP,
     1MFR,VC,N
      IF(PC.GT.16.0) GO TO 105
      SCORE=0.0
      GO TO 110
  105 SCORE=1.0
  110 RETURN
      END
      SUBROUTINE TEST2
      DIMENSION X(24),T(24)
      COMMON X,T,SF,TC,XM,SD,PC,SFV,SHFR,SCORE,PHONV,MP,
     1MFR,VC,N
      IF(PC.GT.14.0) GO TO 115
      SCORE=0.0
      GO TO 120
  115 SCORE=1.0
  120 RETURN
      END
```

```
  201            SUBROUTINE PHOT
>   202             DIMENSION X(24),T(24)
>   203             COMMON X,T,SP,TC,XM,SD,PC,SFV,SFFR,SCORE,PHOTV,MP,
>   204            1MFR,VC,N
>   205             REAL MP,MFR,MPHOTV,MPFR,MVC
>   206       125  FORMAT(F5.2,1X,F5.3,1X,F4.2)
>   207       130  FORMAT(10X,F5.2,4X,F5.3,3X,F5.2,1X,F5.2,3X,F5.3,3X
>   208            1,F6.3)
>   209             READ(1,125) PHOTT,MPHOTV,MPFR
>   210             PHOTV=SFV*TC*MPHOTV
>   211             PFR=SFFR*TC*MPFR
>   212             MFR=PHOTV*60/PHOTT
>   213             MP=MTR/PFR
>   214             VVI=MFR/VC
>   215             WRITE(6,130) PHOTT,PHOTV,MFR,PFR,MP,VVI
>   216             RETURN
>   217             END
#END OF FILE
#
```

APPENDIX 3

APPENDIX 3

A3.1. FOURIER SERIES EXPANSION FOR ASYMMETRICAL TRIANGULAR WAVEFORM

Let f(x) be a periodic waveform defined in the interval $o \leqslant x \leqslant 2\pi$ and satisfying the condition that f(x) and its first derivative are piecewise continuous, then f(x) can be described by the Fourier series

$$f(x) = \sum_{n=o}^{\infty} (a_n \cos nx + b_n \sin nx)$$

where $$a_n = \frac{1}{\pi} \int_o^{2\pi} f(x) \cdot \cos nx \cdot dx$$

and $$b_n = \frac{1}{\pi} \int_o^{2\pi} f(x) \cdot \sin nx \cdot dx.$$

If such a waveform is analysed, then a series of components corresponding to n = 1, 2, 3 ..... are obtained, whose intensities are given by

$$I_n = a_n^2 + b_n^2 .$$

In the special case of the asymmetrical triangular waveform shown in figure A3.1, f(x) is defined by the equations :

$$f(x) = \frac{A}{t_1} \cdot x \quad , \; o \leqslant x < t_1$$

$$= \frac{A}{t_2} \cdot (t_1 + t_2) - \frac{Ax}{t2} \quad , \; t_1 \leqslant x < t_1 + t_2$$

$$= 0 \quad , \; t_1 + t_2 \leqslant x \quad 2 \leqslant 2\pi$$

and it can be shown (SECTION A3.2) that :

$$a_n = \frac{A}{\pi n^2} \left(\frac{1}{t_1} + \frac{1}{t_2} \right) \cos nt_1 - \frac{A}{\pi n^2} \cdot \frac{1}{t_2} \cos n(t_1 + t_2) - \frac{A}{\pi n^2} \cdot \frac{1}{t_1}$$

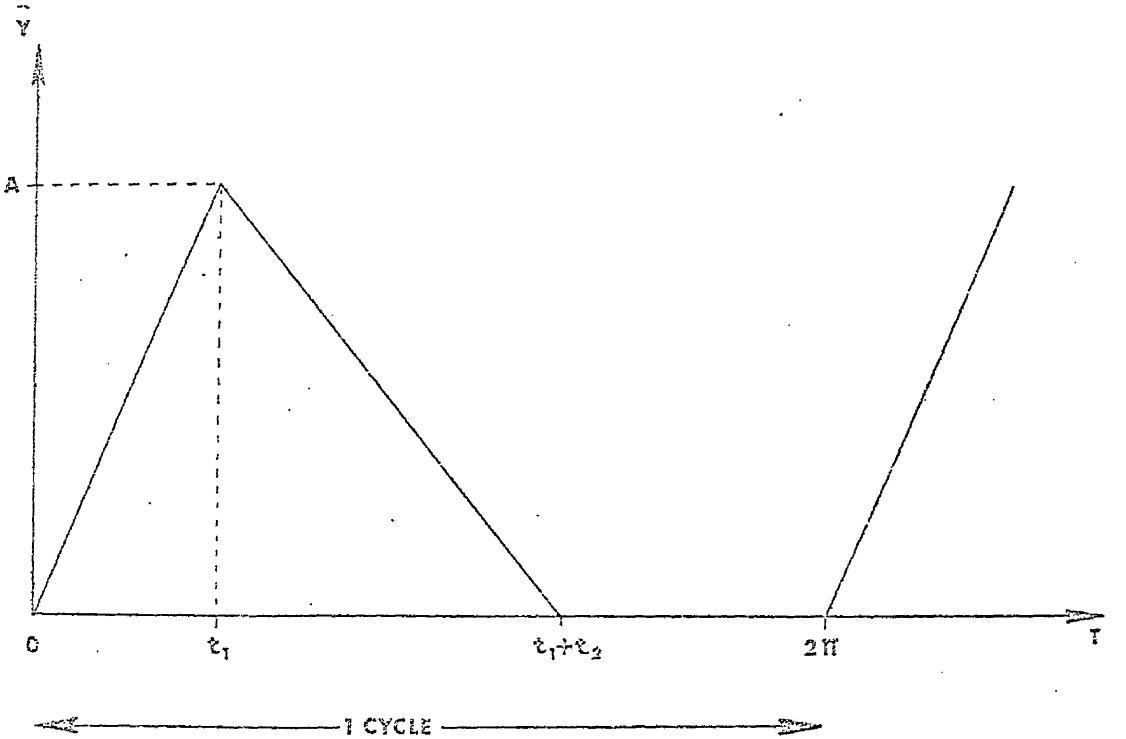


FIG.A3.1. TYPE OF TRIANGULAR WAVEFORM USED TO SIMULATE THE Lx WAVEFORM.

and
$$b_n = \frac{A}{\pi n^2}\cdot\left(\frac{1}{t_1} + \frac{1}{t_2}\right) \sin nt_1 - \frac{A}{\pi n^2}\cdot\frac{1}{t_2}\sin n(t_1 + t_2)$$

Thus I_n can be calculated for each value of n, if t_1 and t_2 are known. The values of t_1 and t_2 were expressed as percentages of the period, and the spectral components calculated for a range of values of t_1 and t_2. The harmonic component which corresponded to the first minimum of the envelope curve of each spectrum was noted. The value of $(100/t_1)$ was also calculated, and the results are tabulated in TABLE A3.1. The values of t_1 were in the range 3-20% of the period, and for t_2 20-67% of the period, comparable to the range of values found in Lx waveforms. The table shows that there was very good agreement between the value of $(100/t_1)$ and the harmonic number of the first minimum of the envelope curve. It was possible that if a smooth curve was fitted through the harmonic components, a more precise estimation of the position of the minimum could be obtained. Thus for a waveform of the required shape, if the frequency spectrum was known, an estimation of the value of t_1 could be obtained. If the fundamental frequency was also known, the absolute time for the t_1 phase of the waveform could be calculated.

A3.2 CALCULATION OF a_n AND b_n

(a): a_n

From $a_n = \frac{1}{\pi}\int_0^{2\pi} y \cos nx \cdot dx$,

and substituting the value of y

$$\frac{\pi}{A}\cdot a_n = \frac{1}{t_1}\int_0^{t_1} x\cos nx\,dx - \frac{1}{t_2}\int_{t_1}^{(t_1+t_2)} x\cos nx\,dx.$$

$$+ \frac{(t_1+t_2)}{t_2}\int_{t_1}^{(t_1+t_2)} \cos nx\,dx$$

t_1 (% OF PERIOD)	t_2 (% OF PERIOD)	$\frac{100}{t_1}$	HARMONIC NO.AT 1ST MINIMUM
3.33	20.00	30	30
	26.67	30	30
	33.33	30	30
	40.00	30	30
	46.67	30	30
	50.00	30	30
	60.00	30	30
	66.67	30	30
6.67	20.00	15	15
	26.67	15	15
	33.33	15	15
	40.00	15	15
	46.67	15	15
	50.00	15	15
	60.00	15	15
	66.67	15	15
10.00	20.00	10	10
	26.67	10	10,11
	33.33	10	10
	40.00	10	10
	46.67	10	10
	50.00	10	10
	60.00	10	10
	66.67	10	10
13.33	20.00	7.5	8
	26.67	7.5	8
	33.33	7.5	9
	40.00	7.5	8
	46.67	7.5	8
	50.00	7.5	8
	60.00	7.5	8
	66.67	7.5	8
16.67	20.00	6	6
	26.67	6	8
	33.33	6	6
	40.00	6	6
	46.67	6	6
	50.00	6	6
	60.00	6	6
	66.67	6	6
20.00	20.00	5	5
	26.67	5	5
	33.33	5	6
	40.00	5	5
	46.67	5	5
	50.00	5	5
	60.00	5	5
	66.67	5	5

TABLE A3.1

DATA ON WAVESHAPE PARAMETERS AND HARMONIC NUMBER OF FIRST MINIMUM IN THE ENVELOPE CURVE OF THE FREQUENCY SPECTRUM.

giving

$$a_n = \frac{A}{\pi n^2}\left(\frac{1}{t_1} + \frac{1}{t_2}\right)\cos nt_1 - \frac{A}{\pi n^2}\frac{1}{t_2}\cos n(t_1+t_2) - \frac{A}{\pi n^2}\cdot\frac{1}{t_1}$$

(b) : b_n

$$b_n = \frac{1}{\pi}\int_o^{2\pi} y \sin nx.dx \quad ,$$

so that :

$$\frac{\pi}{A} \cdot b_n = \frac{1}{t_1}\int_o^{t_1} x.\sin nx.dx - \frac{1}{t_2}\int_{t_1}^{t_1+t_2} x.\sin nx.dx$$

$$+ \frac{(t_1+t_2)}{t_2}\int_{t_1}^{t_1+t_2} \sin nx.dx$$

giving:

$$b_n = \frac{A}{\pi n^2}\left(\frac{1}{t_1} + \frac{1}{t_2}\right)\sin nt_1 - \frac{A}{\pi n^2}\cdot\frac{1}{t_2}\sin n(t_1+t_2)$$

Then, intensity of each harmonic component is given by

$$I_n = a_n^2 + b_n^2 \quad .$$

In the special case of the symmetrical triangular wave,

$t_1=t_2=t$, and a_n and b_n can be simplified to give

$$a_n = \frac{2A}{\pi n^2 t}(1 - \cos nt).\cos nt$$

$$b_n = \frac{2A}{\pi n^2 t}(1 - \cos nt).\sin nt$$

Thus, in this case,

$$I_n = \frac{4A^2}{\pi^2 n^4 t^2}(1 - \cos nt)^2$$

giving minima in the envelope curve when cos nt = 1

i.e. $$n = \frac{2\pi}{t}J \quad (J = 1,2,3,.....)$$

and these minima are equally spaced.

REFERENCES

REFERENCES

Baldwin, E.D.F., Cournand, A., Richards, D.W.: Pulmonary Insufficiency. Medicine 1938,27,243-278.

Campbell,E.J.M.: Muscular Activity in Normal and Abnormal Ventilation. Ventilatory and Phonatory Control Systems (ed.Wyke), (Oxford Univ. Press, 1974), 1-11.

Cohen, M.I.: Discussion. Ventilatory and Phonatory Control Systems (ed.Wyke), (Oxford Univ. Press, 1974), 108.

Dedo, H.H., Dunker, E.: Husson's Theory. Arch.Otolaryng. 1967, 85, 303-313.

Draper, M.H., Ladefoged, P., Whitteridge, D.: Respiratory Muscles in Speech. J. Speech Hearing Res. 1959, 2, 16-27.

Fant, G.: Acoustic Theory of Speech Production. (Mouton, The Hague, 1970).

Flanagan, J.L.: Some Properties of the Glottal Sound Source. J. Speech Hearing Res. 1958, 1, 99 - 116.

Fourcin, A.J.: Laryngographic Examination of Vocal Fold Vibration. Ventilatory and Phonatory Systems (ed.Wyke), (Oxford Univ.Press, 1974), 315 - 326.

Fourcin, A.J., Abberton, E.: First Applications of a New Laryngograph. Med.Biol.Illust.1971, 21, 172 - 182.

Fromkin, V.: Neuro-muscular Specification of Linguistic Units. Lang. Speech 1966, 7, 170 - 199.

Fromkin, V., Ladefoged, P.: Electromyography in Speech Research. Phonetica 1966, 15, 219-242.

Ganong, W.F.: Review of Medical Physiology (3rd Edition), (Blackwell Scientific Publications, Oxford, 1967.)

Garret, H.E.: Statistics in Psychology and Education. (Longmans, Edinburgh, 1967).

Gatehouse, S., Kelman, A.W.: Comparison of the Electromyographic Activity of the Muscle Orbicularis Oris Using Different Electrode Configurations. Folia Phoniat.1976, 28, 52-62.

Gray, H.: in Warwick and William's "Gray's Anatomy", 35th Edit. (Longmans, Edinburgh, 1973).

Greene, M.C.L.: The Voice and its Disorders. (Pitmans, London, 1957).

Guitar, B.: Reduction of Stuttering Frequency Using Analog Electromyographic Feedback. J.Speech Hearing Res. 1975, 18, 672 - 685.

Hirano, M.: Morphological Structure of the Vocal Cord as a Vibrator and its Variations. Folia Phoniat.1974, 26, 89 - 94.

Husson, R.: Sur la Physiologie Vocale. Ann.Otolaryng. 1952, 69, 124 - 137.

Isshiki, N.: Regulatory Mechanism of Voice Intensity Variation. J.Speech Hearing Res. 1964, 7, 17 - 29.

Isshiki, N.: Vocal Intensity and Air Flow Rate. Folia Phoniat. 1965, 17, 92 - 104.

Isshiki,N., von Leden,H.: Hoarseness: Aerodynamic Studies. Arch. Otolaryng. 1964, 80, 206-213.

Iwata,S., von Leden,H.: Voice Prints in Laryngeal Disease. Arch.Otolaryng.1970, 91, 346-351.

Kelman,A.W., Gatehouse,S.: A Study of the Electromyographic Activity of the Muscle Orbicularis Oris. Folia Phoniat. 1975,27,177 - 189.

Kelman,A.W., Gatehouse,S.: Origins in the Muscle Structure of the Electromyographic Activity Associated with the Plosive /p/ in a CVC Utterence. Folia Phoniat. 1976,28,85-93.

Kelman,A.W., Gordon, M.T., Simpson, I.C., Morton,F.M.: Assessment of Vocal Function by Air Flow Measurements. Folia Phoniat. 1975, 27, 250 - 262.

Kirchner,J.A., Suzuki, M.: Laryngeal Reflexes and Voice Production. Ann.N.Y.Acad.Sci. 1968, 155, 98 - 109.

Kirchner,J.A., Wyke,B.D.: The Innervation of the Laryngeal Joints in the Cat. J. Anat. 1964, 98, 684 - 694.

Kirchner,J.A., Wyke,B.D.: Articular Reflex Mechanisms in the Larynx. Ann.Otol.Rhinol.Laryngol. 1965, 74, 749 - 769.

Kirk,J.: Personal Communications, 1975.

Kitzing,P., Sonesson,B.: A Photoglottographical Study of the Female Vocal Folds during Phonation. Folia Phoniat. 1974, 26, 138-149.

Koike,Y., Hirano,M., von Leden,H.: Vocal Initiation: Acoustic and Aerodynamic Investigations of Normal Subjects. Folia Phoniat. 1967, 19, 173 - 182.

Kotby,M.N., Haugen,L.K.: The Mechanics of Laryngeal Function. Acta Otolaryng. 1970,70, 203-211.

Leanderson,R., Lindblom,B.E.F.: Muscle Activation for Labial Speech Gestures. Acta Otolaryng. 1972, 73, 362 - 373.

Leanderson,R., Persson,A., Ohman,S.: Electromyographic Studies of Facial Muscle Activity in Speech. Acta Otolaryng. 1971, 72, 361 - 369.

Leanderson,R., Persson,A., Ohman,S.: Electromyographic Studies of the Function of the Facial Muscles in Dysarthria. Acta Otolaryng. Suppl. 263, 1970, 89 - 94.

Leanderson,R., Ohman,S., Persson,A.: Electromyographic Studies of Facial Muscle Co-ordination during Speech. Acta Otolaryng. Suppl. 224, 1967, 307 - 310.

Lecluse,F.L.E., Brocaar,M.P., Verschure,J.: The Electroglottography and its relation to Glottal Activity. Folio Phoniat. 1975, 27, 215-224.

Lieberman,P.: Vocal Cord Motion in Man. Ann.N.Y. Acad. Sci. 1968, 155, 28 - 38.

Lubker,J.F., Parris,P.J.: Simultaneous Measurements of Intraoral Pressure, Force of Labial Contact, and Labial Electromyographic Activity during Production of the Stop Consonant Cognates /p/ and /b/. J. Acoust. Soc. Am. 1970, 47, 625 - 633.

MacNeilage,P.F., De Clerk,J.L.: On the Motor Control of Coarticulation in CVC Monosyllables. J. Acoust. Soc. Am. 1969, 45, 1217 - 1233.

Moroney,M.J.: Facts from Figures. (Penguin, 1970).

Palmer,J.M., La Russo,D.A.: Anatomy for Speech and Hearing. (Harper and Row, New York, 1965).

Perkins,W.H., Yanagihara,N.: Parameters of Voice Production: I Regulation of Pitch. J. Speech Hearing Res. 1968, 11, 246-267

Peterson,G.E.: Parameters of Vowel Quality. J. Speech Hearing Res. 1961, 4, 10 - 20

Peterson,G.E., Barney,H.L.: Control Methods Used in a Study of the Vowels. J. Acoust. Soc. Ann. 1952, 24, 175 - 184

Potter,R.K., Kopp,A.G., Green,H.C.: Visible Speech. (Van Nostrand, New York,1947).

Proctor,D.F.: Breathing Mechanics During Phonation and Singing. Ventilatory and Phonatory Control Systems (Ed.Wyke), (Oxford Univ.Press, 1974), 39 - 57.

Seymour,J.: Acoustic Analyses of Singing Voices. III Spectral Components, Formants and the Glottal Source. Acoustica 1972, 27 218-227

Stevens,K.N., Klatt,D.H.: Current Models of Sound Sources for Speech Ventilatory and Phonatory Control Systems (Ed.Wyke), (Oxford Univ.Press,1974), 279 - 292.

Tatham,M.A.A., Morton,K.: Some Electromyography Data towards a Model of Speech Production. Lang. Speech 1969, 12, 39 - 53.

Tatham,M.A.A., Morton,K.: Electromyographic and Intraoral Air Pressure Studies of Bi-Labial Stops. Lang.Speech 1973, 16, 336 - 350.

Timcke,R., von Leden,H., Moore,P.: Laryngeal Vibrations: Measurements of the Glottic Wave,Part I. Arch.Otolaryng.1958, 68, 1-19.

Timcke,R., von Leden,H., Moore,P.: Laryngeal Vibrations: Measurements of the Glottic Wave, Part II. Arch.Otolaryng. 1959, 69, 438 - 444.

Titze,I.R.: The Human Vocal Cords: A Mathematical Model Part I. Phonetica 1973, 28, 129 - 170.

Titze,I.R.: The Human Vocal Cords: A Mathematical Model Part II. Phonetica 1974, 29, 1-21.

Van den Berg, J.: Myoelastic - Aerodynamic Theory of Voice Production. J.Speech Hearing Res. 1958,1, 227 - 244.

Van den Berg,J., Zantema,J.T., Doornenbal,P.: On the Air Resistance and the Bernoulli Effect of the Human Larynx. J.Acoust. Soc. Ann. 1957, 29, 626 - 631

von Leden,H.: The Mechanism of Phonation. Arch. Otolaryng. 1961, 74, 660 - 676.

von Leden,H.: Objective Measures of Laryngeal Function and Phonation. Ann.N.Y. Acad.Sci. 1968, 155, 56 - 66.

von Leden,H., Moore,P.: Vibratory Pattern of the Vocal Cords in Unilateral Laryngeal Paralysis. Acta Oto-larying. 1960, 53, 493 - 506.

von Leden,H., Moore,P., Timcke, R.: Measurements of the Glottic Wave, part III. Arch.Otolaryng. 1960, 71, 16-35.

Wegel,R.L.: Theory of Vibration of the Larynx. Bell Syst.Tech. J. 1930, 9, 207 - 227.

Wyke,B.D.: Laryngeal Myotatic Reflexes and Phonation. Folia Phoniat, 1974, 26, 249 - 264.

Yanagihara,N.: Significance of Harmonic Changes and Noise Components in Hoarseness. J.Speech Hearing Res. 1967, 10, 531 - 541.

Yanagihara,N., Koike,Y.: The Regulation of Sustained Phonation. Folia Phoniat. 1967,19, 1-18.

Yanagihara,N., Koike,Y., von Leden,H.: Phonation and Respiration. Folia Phoniat. 1966, 18, 323 - 340.

Yanagihara,N., von Leden, H.: Respiration and Phonation. Folia Phoniat 1967, 19, 153 - 166.