A STUDY OF VOICE QUALITY IN A GROUP OF IRRADIATED LARYNGEAL CANCER PATIENTS TUMOUR STAGES T1 AND T2.

by

Eva Ingrid Carlson

A thesis submitted for the degree
of Ph.D in the University of London.

1995

ABSTRACT

A STUDY OF VOICE QUALITY IN IRRADIATED LARYNGEAL CANCER PATIENTS. TUMOUR STAGES T1 AND T2.

This is a longitudinal study of voice quality in a group of 35 patients irradiated for early vocal fold tumours, stages T1 and T2. Electrolaryngograph (ELG) based analyses were used to obtain objective measurements of speaking fundamental frequency parameters over a wide range of time intervals following radiotherapy. Lx waveforms were also analysed.

Perceptual evaluation of voice quality and patients' self assessments of their experience of vocal symptoms and limitations in vocal function after radiotherapy, were carried out. The relationship between perceptual and self assessment parameters and objective voice quality measurements was determined.

A few patients underwent periods of voice therapy. A comparison is made of their voice measurements before and after therapy intervention with a group of patients, who did not receive voice therapy.

The findings in this study show that, contrary to some early reports that the voice returns to normal in the majority of patients after radiotherapy, most patients' show evidence of residual abnormal voice quality and symptoms as measured and as rated by clinicians and by patients themselves. The majority of patients do not consider these a major problem, however.

Evidence is presented of the beneficial effect of voice therapy to help patients compensate for the inevitable tissue damage caused by radiotherapy to the larynx.

Electrolaryngograph generated objective measures and Lx waveforms proved sensitive, reliable and clinically applicable for objective voice analysis.

ACKNOWLEDGEMENTS.

This being a longitudinal study spanning over ten years, all my friends and colleagues in and out of St Thomas' Hospital, where I have been employed for the duration, have been helpful and supportive in various ways, all of whom I thank.

I thank St Thomas' Hospital Research Endowments Committee and the Ear, Nose and Throat Department through Mr Alec Fitzgerald O'Connor for supporting me financially through their generous grants.

A few individuals stand out for their active encouragement and assistance:

Dr Basil Stoll for his encouragement that I undertake this study.
Mr Valentine Hammond and Dr David Collins for giving me access to their patients and for their advice and support in the early stages of the Pilot study.
Mr John Evans and Mr Alec Fitzgerald O'Connor who also allowed me to include their patients in the study.
Dr Ron Beaney for spending time reading drafts of text and advising me on matters concerning radiotherapy.
Nick Taub and Fiona Warburton in the Department of Public Health Medicine for their time and patience in teasing out problems and leading me through the jungle of statistical analyses and computing.
David Miller of 'Laryngograph Ltd.' for being at the end of the telephone when instant advice was needed.

Last but not least I thank my tutor Dr Evelyn Abberton of University College London who, with her endless patience and helpful advice, has supported me throughout the long process of completing this study.

A STUDY OF VOICE QUALITY IN IRRADIATED LARYNGEAL CANCER PATIENTS TUMOUR STAGES T1 AND T2.

VOLUME ONE

CONTENTS.

III VOICE QUALITY

IV DIRECT AND INDIRECT ASSESSMENT OF LARYNGEAL FUNCTION

A ACOUSTIC MEASUREMENT TECHNIQUES

B GLOTTOGRAPHIC TECHNIQUES

V ELECTROGLOTTOGRAPHY (EGG) and ELECTROLARYNGOGRAPHY (ELG)

VI THE EFFECT OF RADIOTHERAPY ON LARYNGEAL FUNCTION

VII ELECTROLARYNGOGRAPHY (ELG) IN ASSESSMENT AND MEASUREMENT OF VOCAL FOLD FUNCTION

VIII THE RADIOTHERAPY STUDY

A THE PILOT STUDY

B THE MAIN STUDY

IX REPEATABILITY OF AND AGREEMENT BETWEEN SOME Lx BASED MEASUREMENTS OF FUNDAMENTAL FREQUENCY, REGULARITY AND IRREGULARITY.

A

B

X SIGNIFICANCE OF THE DIFFERENCES IN Fx and %TS BETWEEN THE NORMAL SPEAKERS AND THE T1 SUBJECTS

XI PERCEPTUAL EVALUATION OF VOICE QUALITY

XV DISCUSSION

LIST OF FIGURES.

CHAPTER I   Page

CHAPTER III

CHAPTER IV

CHAPTER V

CHAPTER VI

CHAPTER VII

CHAPTER VIII

CHAPTER IX

CHAPTER XI

CHAPTER XII

CHAPTER XIII

CHAPTER XIV

LIST OF TABLES.

CHAPTER II

CHAPTER V

CHAPTER VI

CHAPTER VIII

CHAPTER IX

CHAPTER X

CHAPTER XI

CHAPTER XII

CHAPTER XIII

CHAPTER XIV

VOLUME TWO.

LIST OF APPENDICES.

INTRODUCTION.

A Pilot study was devised to develop standard techniques and baselines for using Electrolaryngography (ELG) as a routine means of objective voice analysis in voice therapy and to interpret Dx and Cx plots and Lx waveforms (Fourcin and Abberton, 1971, Abberton and Fourcin, 1972, 1976, 1984, Abberton 1976, Wechsler, 1977, Abberton, Howard and Fourcin, 1989, Fourcin, 1974, 1981, 1982, 1989 a,b). The subjects chosen for the study, were patients who had received radiotherapy for laryngeal carcinoma, who were not routinely referred to the Speech and Language Therapy department for voice therapy or advice.

Electrolaryngography (ELG) is one of few clinically applicable methods for reliable, objective voice analysis. It is non-invasive and measures voice fundamental frequency through registering individual vocal fold contacts via two superficial electrodes placed either side of the thyroid cartilage. This facilitates calculation of fundamental frequency generated by the voice source, the vocal folds. It avoids the need to derive the fundamental from the complex acoustic wave spectrum radiating at the lips that results from the filtering of the glottal source waveform in the supraglottic vocal tract.

The advantages of Electrolaryngography for clinical voice assessment are the following:

a) ELG is completely non-invasive, simple and quick to apply in a clinic situation, where there is often a lack of ideal acoustic recording facilities.

b) ELG allows the recording and fundamental frequency analysis of speech samples several minutes in duration thereby increasing the possibility of capturing a wide range of habitual linguistic and paralinguistic phonatory behaviours characteristic for an individual.

c) ELG allows the recording of a subject in conversation, interacting with another speaker, a more 'natural', habitual use of the voice for that individual than samples of reading or sustained phonation (Hirson and Roe, 1993). Not all patients attending a Speech and Language Therapy department will be able to read confidently or at all.

d) Lx waveforms may be used to reflect and monitor changes in vocal fold contact behaviour as a result of a disease process, medical or therapeutic intervention. Lx can also be used to offer a patient visual feedback of changing laryngeal behaviour e.g. as a result of the learning process that is our aim in voice therapy (Carlson, 1986, 1988 a, c, 1993 a, b)

Phonation is the result of complex interaction between laryngeal muscular-elastic, mucosal and aerodynamic forces effecting vocal fold vibration in the pulmonic egressive airstream. The complexity of normal vocal fold vibration has been studied in detail by high speed filming and stroboscopic observation of vocal fold vibration in slow motion. Laryngeal anatomy and aerodynamics, the structure of the vocal fold and theories of vocal fold vibration will be described in Chapter I.

Control and stability of the voice fundamental frequency and the effect on this of sex, age, smoking and laryngeal pathology will be described in Chapter II.

'Voice quality' is, however, not solely dependent on vocal 'pitch', the perceptual equivalent of fundamental frequency. Vocal fold pathology or malfunctioning may lead to disturbances in the regularity of vibration, the completeness and mode of vocal fold contact. This is likely to have an effect on vocal fold speed of vibration, registered as fundamental frequency, but also give rise to a variety of other perceptual phenomena, often described as degrees of 'hoarseness', 'roughness, or 'breathiness', not neccessarily fundamental frequency dependent. Many schemes have been devised to assist voice clinicians in perceptually describing and communicatiing degrees of such voice quality features to gain a more complete picture of 'voice quality' than what is purely described by fundamental frequency parameters. Research into the advantages and problems with such systems and their relationship to acoustic measures will be described in Chapter III.

Techniques for objective measurement of voice parameters will be reviewed in chapter IV. Particular reference to the interpretation of Electroglottographic (EGG) and Electrolaryngographic (ELG) data is offered in Chapter V.

Radiotherapy is the treatment of choice for early tumours of the larynx, stages T1 and T2, as they give symptoms early and are thereby highly curable (Dickens, Cassisi, Million and Bova, 1983, Kaplan, Johns, Slaughter-Fitzhugh, 1983, Botnik, Rose, Goldberg and Recht, 1984, Mendenhall, Parsons and Stringer, 1988). Chapter VI reviews the use of radiotherapy for laryngeal carcinoma and studies examining the effect of radiotherapy on voice quality and function as rated by professionals and by irradiated speakers themselves.

Studies have been reported which claim that the voice quality returns to 'normal' within approximately two months after radiotherapy (Mendonca, 1975, Stoicheff, 1975, Karim, Snow, Diek and Hanjo, 1983). There is, however, also evidence that acute and late laryngeal tissue reactions may produce significant longterm effects on voice quality and vocal function (Fu, Woodhouse, Quivey et al, 1982, Lo, Salzman and Schwartz, 1985, Mendenhall et al, 1988, Benninger, Gillen, Thieme et al, 1994).

One early study is reported where voice therapy was offered to help alleviate the effects of radiotherapy on patients' voices before during and after treatment. It was found to be beneficial and patients gave a positive response to such therapy (Fex and Henriksson, 1966). Other researchers have expressed opinions that voice therapy is likely to help patients adjust to unfavourable laryngeal conditions after radiotherapy. Some do so more easily than others and may not experience much difficulty after treatment. Others adjust less well and may experience severe limitations in voice use and quality (Stoicheff, 1975, Lehmann, Bless and Brandenburg, 1988).

Chapter VII elaborates on particular Electrolaryngographic voice measurements used in this study.

As the Pilot study, described in Chapter VIII A, got under way, it became evident that some of the irradiated speakers needed help for residual voice problems. All were grateful for explanations of laryngeal structure and function and how the radiotherapy had affected their laryngeal tissues. A number of subjects were offered periods of voice therapy and there was indication of both perceptually and objectively measureable improvement.

After the Pilot study had been completed, irradiated laryngeal cancer patients continued to be referred for voice therapy because of residual voice problems. A modified protocol was devised for the 'Main study', introduced in Chapter VIII B, to follow up speakers irradiated for early glottic tumours stages T1 and T2, in the long term, some of whom received courses of voice therapy. Some of the original Pilot study subjects were also reassessed at gradually longer intervals after radiotherapy.

During the course of this study a new fractionation regime for radiotherapy for glottic carcinoma was introduced. A comparison is made of voice quality in the short and long term under the two different fractionation schedules, 3 fractions per week versus 5 fractions per week.

The decision to continue data collection from old and new subjects at increasing intervals after radiotherapy meant that subjects were recorded many times over periods spanning several years. The ELG instrumentation and software for voice analysis was further developed during this period. The original Voiscope was replaced by the Laryngograph Processor. The BBC microcomputer based TPS and ROM system for voice analysis were replaced by PCLx. The original recording equipment was also replaced.

Apart from variability in individual speakers' objective voice measurements over time, there was a risk that changes in recording instrumentation and software for voice analysis might have introduced artefacts into measurements, which would affect the comparability. It seemed prudent to find out the size of any systematic measurement error resulting from such changes.

Chapter IX tests the repeatability of and agreement between objective measurements carried out on the same recorded voice samples using the same and different software programs on two different occasions. Most were found satisfactory for the analysis systems used on speech and reading samples from each of a random sample of ten irradiated T1 subjects.

A random sample of ten 'Normal' speakers was drawn from the data collected by Kramer (1989) using the same recording and analysis instrumentation. The repeatability and agreement of their measurements was also determined. No major artefacts seemed to have been introduced by changes in instrumentation or software.

A measure of vocal fold regularity of vibration calculated as the proportion of period, Tx, samples carried into 2nd order Dx distributions (%TS), was developed in the early days of the experimenter using ELG for voice analysis. It was found useful as an objective measure of 'voice quality' other than fundamental frequency and was sensitive to improvement as a result of intervention, surgical and therapeutic (Carlson, 1993 a and b).

No ELG based norms for speaking fundamental frequency parameters exist for the age group concerned in this study. The randomised 'Normal' speakers' and T1 subjects' speech and reading samples used in the repeatability study, were therefore also used to determine if any significant differences could be detected between the irradiated and 'Normal speakers' objective voice measurements. Chapter X shows that, although the irradiated speakers showed consistently lower central and range fundamental frequency measurements than the 'normal' speakers, most of these differences did not reach significances due to large intra-group variation within both groups. However, the regularity measure, %TS and the 2nd order fundamental frequency range maximum were significantly lower in the irradiated speakers. So was the % Irregularity measure which was introduced with the PCLx software program, employed in more recent voice analyses.

The difficulties with perceptual evaluation of such a great number of voice recordings as have been recorded and analysed here, are described in Chapter XI. Inter- and intra-rater agreements are determined and findings are reported showing significant correlations between some of the objective measurements and perceptual and self assessment parameters used in this study.

Chapter XII describes trends in the total voice quality data collected as a function of a) time elapsed since radiotherapy b) tumour stage and fractionation c) continued employment.

Chapter XIII describes the voice therapy method used and the effect of voice therapy on objective and perceptual voice quality measures in a number of irradiated subjects. The use of visual feedback in therapy with particular reference to Lx waveforms is also described.

Chapter XIV, finally reports on the results of qualitative Lx waveform analyses. Initially, waveforms had to be photographed 'live' during phonation from an oscilloscope screen and visually inspected and qualitatively judged for speed of vocal fold closure, uniformity of waveform peaks and duration of closed and open phases of the vibratory cycle. In recent years computerbased analysis of waveforms have become available including some quantitative measurements.

The Discussion in Chapter XV concerns problems with the methodolgy used in this study and suggests ideas for future study using Electrolaryngography.

CHAPTER I -VOCAL FOLD VIBRATION.

i) The gross anatomy of the larynx and intrinsic laryngeal musculature.

The primary function of the larynx is to act as a protective valve for the airway. As such it has a separate and very different function in swallowing, coughing and throat clearing from that observed in phonation. Only laryngeal function in relation to phonation will be described here.

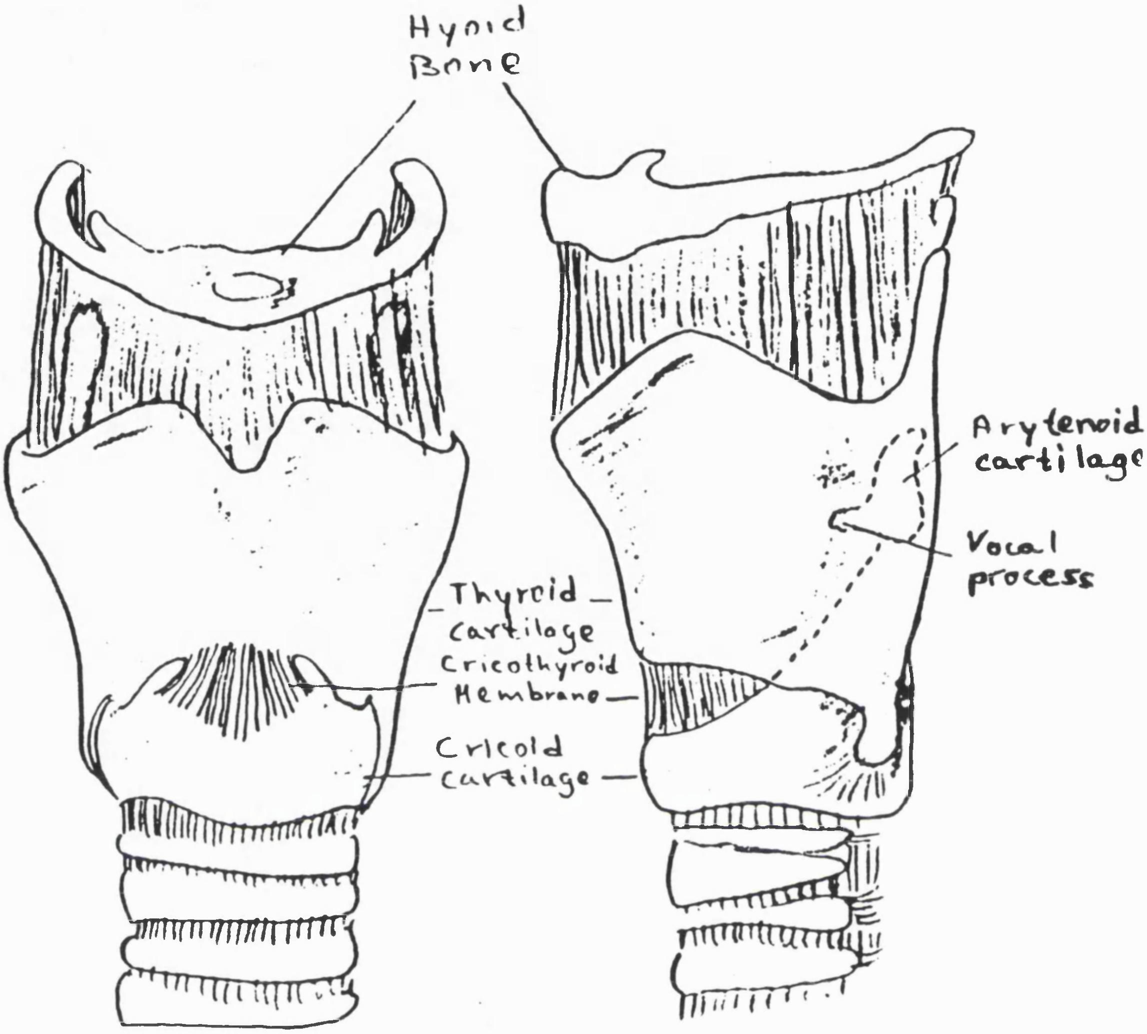
The main sources of this account are Hirano's (1981) 'Clinical Examination of Voice' and Laver's (1980) 'The Phonetic Description of Voice Quality'. Most references to fundamental frequency control are taken from Shipp and McGlone, (1971) and Berke and Gerratt (1993)

Figure 1 shows the relationships between the cartilages constituting the larynx and the trachea and the hyoid bone at the base of the tongue. It shows the thyroid cartilage, with its V-shaped notch in the midline and wide 'wings', or alae, either side. This is a useful 'landmark' in the application of the superficial laryngeal electrodes employed for objective voice measurements using Electrolaryngography in this study. The thyroid cartilage articulates on the cricoid, with its 'signet ring' shape posteriorly clearly demonstrated.

The arytenoid cartilages articulate with this and are more clearly seen in Fig.2, which illustrates the intrinsic laryngeal musculature, whose complex interaction with aerodynamic forces from an egressive pulmonic airstream, effects phonation during speech.

There are five sets of paired muscles and one unpaired muscle, which contribute to control the vibration of the vocal folds during phonation, through their effect on tension and compression of the folds and related structures.

Contraction of the paired **Cricothyroid Muscles** (Fig.2a) results in the lengthening, tensing and thinning of the vocal folds by pulling

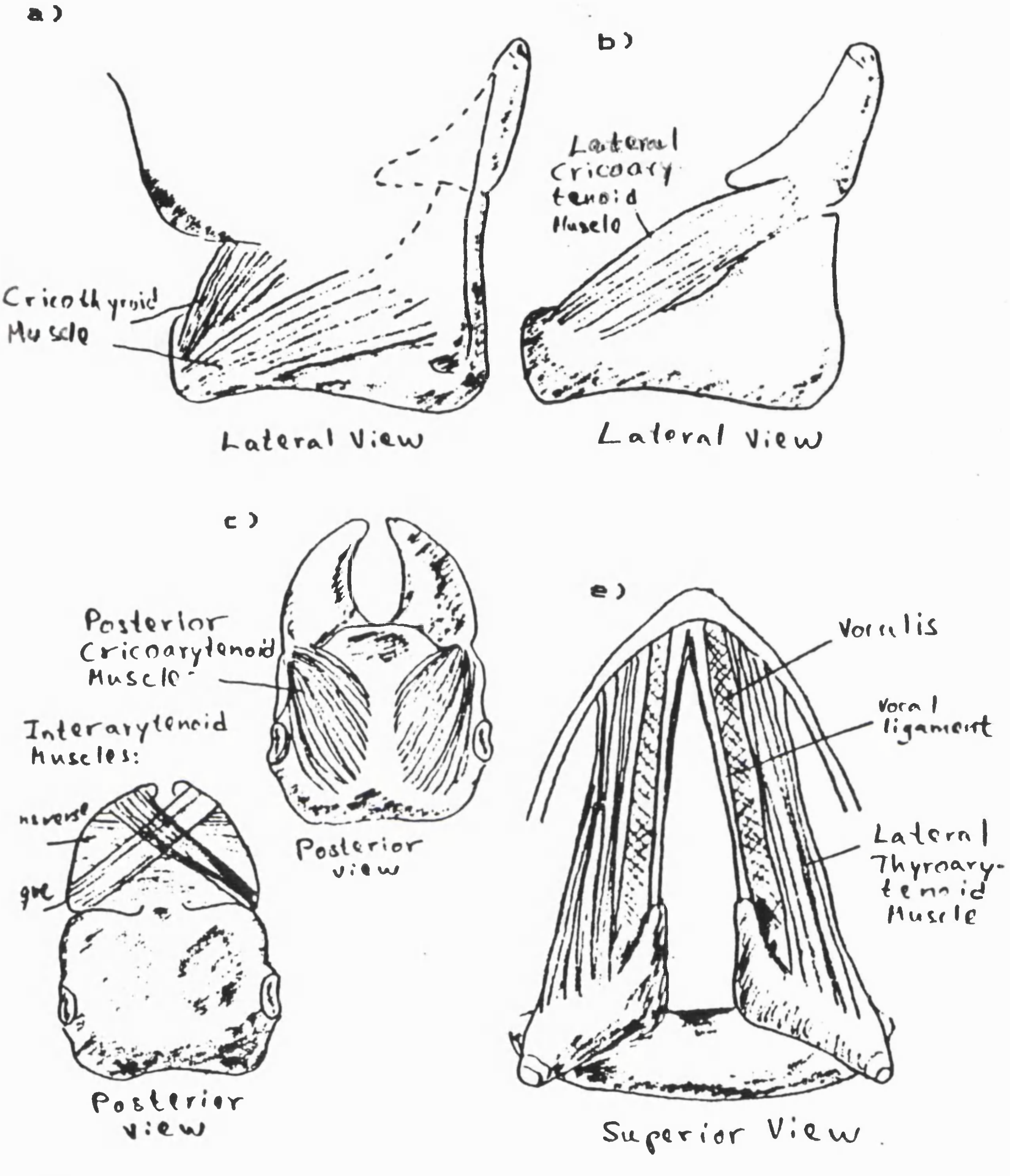


Anterior view　　Lateral view

THE LARYNX

(From: Palmer, 1972)

Figure 1



Intrinsic Laryngeal muscles

(From: Palmer, 1972)

Figure 2

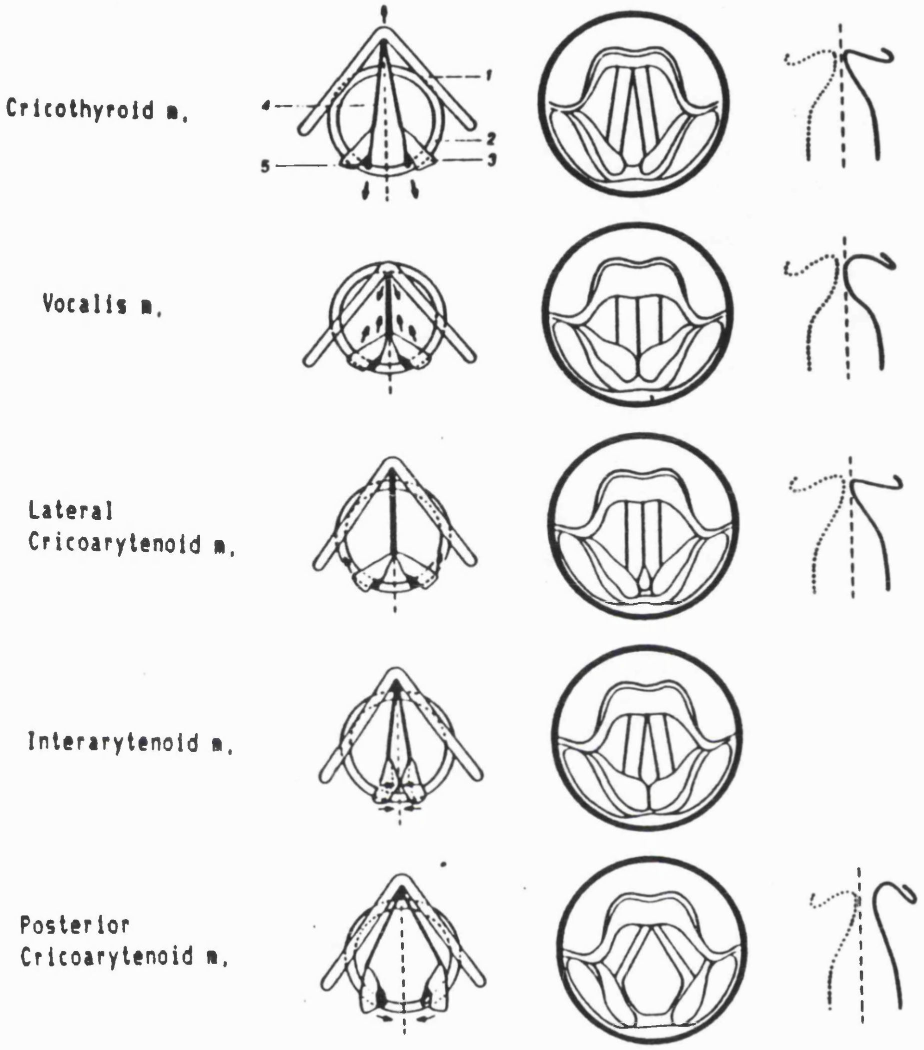
the front of the cricoid up towards the thyroid cartilage, thereby tilting the posterior part of the cricoid down and back. This results in a raising of the fundamental frequency (Hirose and Sawashima, 1981).

The **Thyroarytenoid Muscles** (Fig.2 e), which are also paired, consist of an upper portion constituting the ventricular folds, a lower medial portion constituting the **Vocalis Muscle** or **Vocal Fold** and a lateral portion. They are attached to the thyroid cartilage in front, at the anterior commissure, and to the vocal processes of the arytenoid cartilages posteriorly. Contraction of the thyroarytenoid muscles result in a shortening of the glottis by pulling the arytenoids forward towards the thyroid cartilage. Their activity is also closely involved in fundamental frequency control in co-ordination with Cricothyroid activity.

The **Glottis** is the space between the vocal folds. It is helpful to think of the glottis as divided into the 'ligamental glottis' bordered by the vocal ligaments or membranous portions of the vocal folds, and extending from the anterior commissure to the vocal processes, and the 'cartilaginous glottis', constituting the part of the glottis between the vocal processes of the arytenoids (Fig. 2e).

The action of the **Posterior Cricoarytenoid muscles** (Fig.2c and Fig.3) abducts and lengthens the vocal folds by pulling the muscular processes of the arytenoids, rotating them backwards, pivoting the vocal processes outwards and upwards. Their main function is in opening the glottis bringing about voicelessness.

The **Lateral Cricoarytenoid muscles** (Fig.2 b) act in opposition to the posterior cricoarytenoids in that contraction results in adduction of the vocal folds along their full length by swinging the arytenoids inwards and forwards. The vocal fold is lowered, elongated and thinned (Fig.3).



Schematic representation of the function of the laryngeal muscles. (From: Hirano, 1981)

Figure 3

The paired **Oblique Arytenoid Muscles** (Fig.2 d), tilt the tops of the cartilages towards each other and also contribute to vocal fold adduction.

Contraction of the **Transverse Arytenoid Muscle**,the only unpaired muscle (Fig.2 d), has the effect of drawing the arytenoid cartilages together medially, contributing to vocal fold adduction (Kaplan,1960). It mainly affects adduction at the cartilaginous end of the glottis and does not affect the dimensions of the vocal folds. The latter muscles are often referred to simply as the **Interarytenoid muscles**, and oppose the action of the lateral cricoarytenoids (Fig. 3).

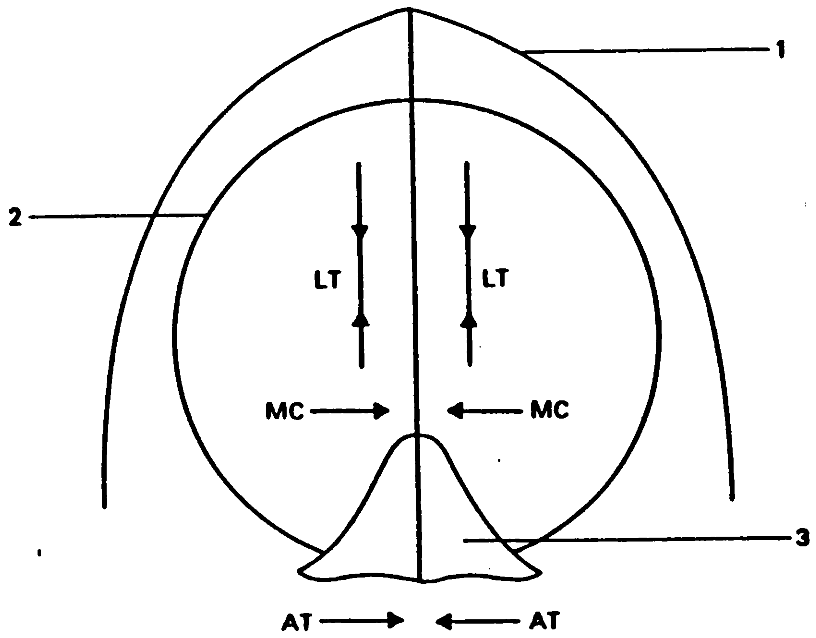
An early summary of how glottal closure is achieved was given by van den Berg (1968):

> *'A contraction of the (powerful) interarytenoid muscles*
> *primarily adducts the apexes of the arytenoids*
> *and closes the back part of them so that no wild air*
> *can escape...A contraction of the lateral*
> *cricoarytenoid muscles adducts the vocal processes*
> *of the arytenoids and therefore the body of*
> *the vocal folds. This adduction is augmented*
> *by a contraction of the lateral parts of the*
> *thyroarytenoid muscles (.....along with an adduction*
> *of the vocal folds). These adductional forces*
> *provide a medial compression of the vocal*
> *folds and reduce the length of the glottis,*
> *which is effectively free to vibrate.'*

ii) Tension characteristics of the vocal folds.

The source of the following account is Laver (1980). Figure 4 illustrates the general direction of forces, which achieve three main laryngeal tension characteristics Laver uses to define his 'Phonatory settings', which will be explained in detail in Chapter III. It shows a diagram of modes of control which contraction of the intrinsic laryngeal musculature may have over laryngeal tension, thereby effecting changes in voice quality and/or fundamental frequency.

Adductive tension (A.T.) is described as resulting from the contraction of the interarytenoid muscles, bringing the arytenoids



Geometric relationship between three laryngeal parameters.

LT-Longitudinal Tension	1. Thyroid cartilage
MC-Medial Compression	2. Cricoid cartilage
AT-Adductive Tension	3. Arytenoid cartilage

(From: Laver, 1980)

Figure 4

together and closing both the cartilaginous and the ligamental glottis (Fig. 2e and Fig. 3).

Medial Compression (M.C.) is defined as the *'compressional pressure on the vocal processes of the arytenoid cartilages achieved by contraction of the lateral cricoarytenoid muscles and reinforced by tension in the lateral Thyroarytenoid muscles'* (Fig. 3) (Van den Berg, 1968, Hardcastle, 1976). M. C. will result in the closure of the ligamental glottis *'but whether the cartilaginous glottis also closes will depend on the analytically separate adductive tension achieved by the interarytenoid muscles'*.

The latter comment confirms the possibility of varying degrees of closure of the glottis in the anterior-posterior dimension, giving rise to different perceived voice qualities. This will be described in Chapter III.

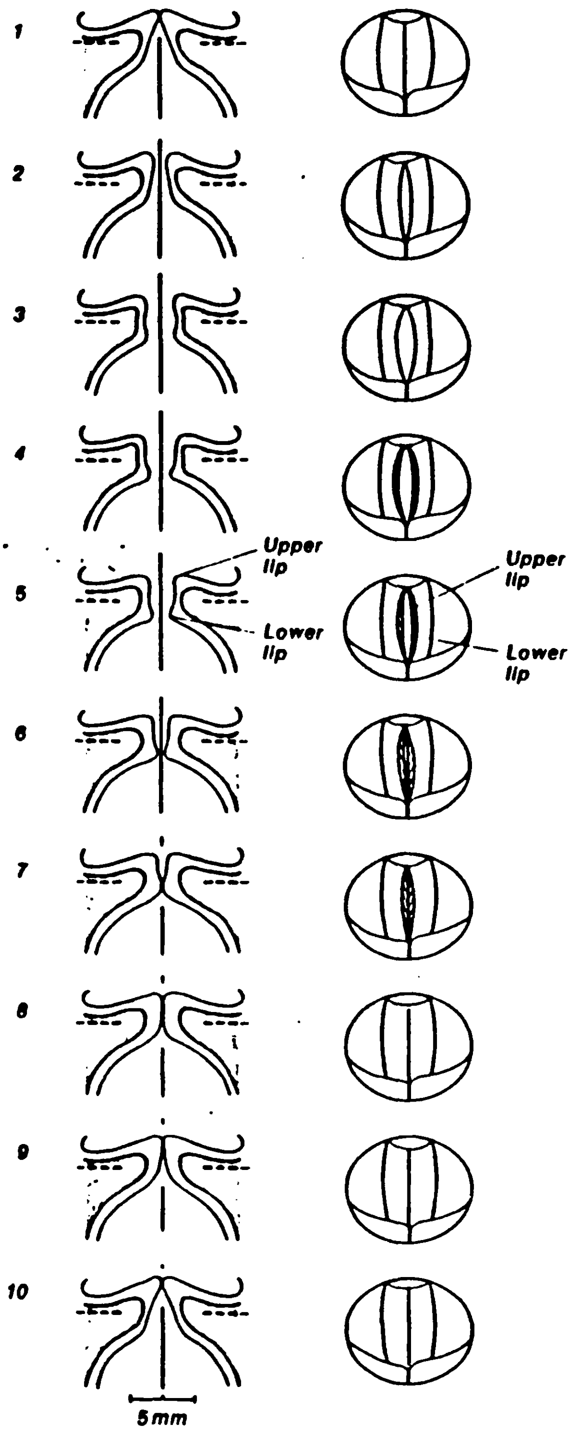
Longitudinal tension (L.T.), described by Van den Berg, (1968), results from the contraction of the vocalis muscle with or without the help of the cricothyroid muscles (Fig.3).

Vocal fold adjustments should be considered in three dimensions; anterior - posterior, left - right and also the vertical dimension. Laver (1980) writes:

> *'The changing vertical thickness of the vocal folds from the outer wall inwards to the vocal ligaments at the edge of the glottal space reflects the interplay of the different tensions that are exerted in and on the folds by the laryngeal musculature, and this third, vertical dimension is one factor among others which differentiates the major settings of the phonatory mechanism.'*

A schematic representation of the changing vertical dimensions of the vocal folds during vibration is shown in Fig.5.

The degree of vocal fold contact during phonation is of crucial importance in effective voice production. It determines the quality of vocal tract excitation at the moment of vocal fold adduction



Schematic representation of vocal fold vibration. (From: Hirano, 1981)
Left column-frontal section
Right column -view from above

Figure 5

(Fourcin, 1974). Degrees of incomplete closure will allow acoustic energy to be absorbed in the subglottis.

Relative degrees of vocal fold contact during phonation are reflected in Electroglottographic waveforms (Lecluse, Brocaar and Verschuure, 1975, Childers, Smith and Moore, 1984, Gilbert. Potter and Hoodin, 1984)). This will be gone into in greater detail in Chapters IV and V.

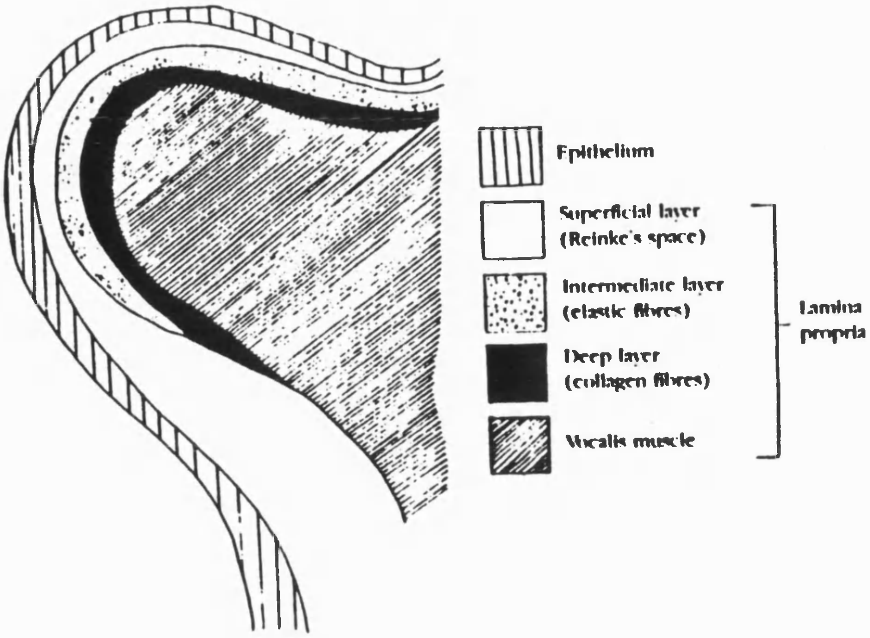
iii) The structure of the Vocal Fold.

Hirano (1977) has provided a crucial, detailed analysis and description of the structure of the vocal fold. A cross section is shown in Figure 6. It illustrates the layered structure of the vocal fold tissues covering the body of the vocalis muscle.

Moving inwards from the epithelium, each layer is looser than the one deeper to it. The four outermost layers constitute the mucous membrane, over which we have no active control. Contraction of the deepest 'layer', the Vocalis Muscle, effects a reduction of the length and an increase in thickness of the crossection of the membranous portion of the vocal folds as illustrated in fig. 3.

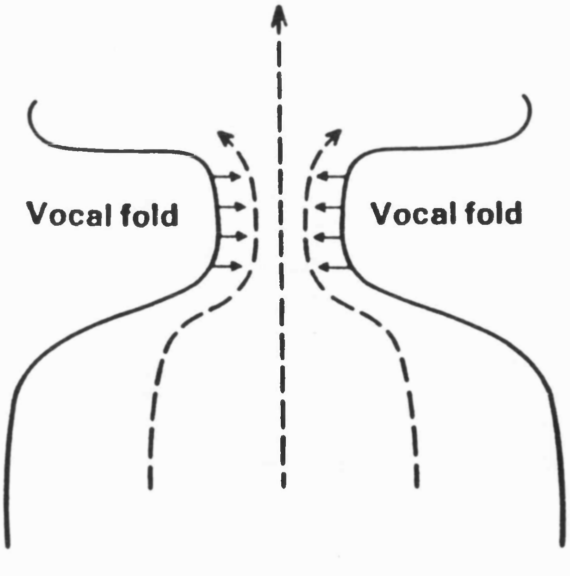
Hirano (1981) divided the five layers into three sections: 'The Cover', consisting of the epithelium and the superficial layer of the lamina propria or Reinke's space forming a semifluid layer (Laver, Hiller and MacKenzie-Beck, 1992); 'The Transition' made up of the intermediate and deep layers of the lamina propria forming a stiffer layer within the Cover and 'The Body' consisting of the Vocalis muscle.

The different layers behave as three relatively independent masses. According to Laver et al (1992), a change in the stiffness, mass or geometry of any of the layers will result in a change in the vibratory pattern of the vocal folds which will have acoustic consequences.



The structure of the vocal fold.
(From; Greene and Mathieson, 1989)

Figure 6



Schematic illustration of the Bernoulli effect. The layer of the airstream in the center between the vocal folds travels a shorter distance than adjacent layers, which are forced to bend around the vocal fold contours. The difference in traveling distance generates an underpressure, or a sucking force, as shown by the small arrows. This sucking force strives to close the glottis as soon as an airstream passes.

The Bernoulli effect.
(From;Sundberg, 1987)

Figure 7

The recognition of the vocal fold as a layered structure is crucial to all recent studies and explanations of the aerodynamics of vocal fold vibration as described below and is essential for sustaining vocal fold oscillation (Titze, 1980, Childers and Krishnamurthy, 1985, Titze, Jiang and Drucker, 1988, Berke and Gerratt, 1993).

Fourcin (1981) suggests the mucous covering the vocal folds should be considered a separate layer, as it may influence glottal measurements using Electrolaryngography. A crucial factor in normal vocal fold vibration, is the condition of the mucous membrane (Titze, 1980, Finkelhor, Titze and Durham,1988). Dryness, inflammation or oedema will have a significant effect on the vibration of the vocal folds and thereby on voice quality, as they affect the undulating movement of the cover over the body of the vocal fold (Hiroto, 1981).

Titze (1980), using a theoretical model of the vocal folds, predicted that:

> *'as viscous damping of the vocal fold tissue increases, the required lung pressure to maintain the same oscillation pattern must also increase if the same glottal configuration is to be maintained'.*

Finkelhor et al, (1988) confirmed this in an experiment using excised canine larynges, which they bathed in different osmotic solutions. They measured what they called the 'oscillation threshold pressure' defined as the subglottic pressure that barely produced sustained vocal fold oscillation. They found that with decreased hydration the range of oscillation decreased; with increased hydration the range increased.

iv) Theories of vocal fold vibration.

The Aerodynamic-Myoelastic theory of vocal fold vibration was a comprehensive theory formulated by Van den Berg (1968) It explained vocal fold vibration as initiated and sustained by the effect of the pulmonary airflow on the muscle systems of the larynx.

The main mechanism was the effect of the egressive pulmonary airflow, travelling up the wide tube of the trachea and into the narrow glottis. This causes a sudden pressure drop between the vocal folds and results in a 'suction effect' from below upwards, the so called **Bernoulli effect** illustrated in Fig.7.

After making contact in the midline, the folds separate in a rolling motion from below upwards (frames 2-9, Fig.5). This is achieved through the combined effect of the increased airpressure from below, and the vocal folds' natural elasticity and tendency to spring back to their original position, likened to a mass on a spring. According to the Aerodynamic-Myoelastic theory, as long as subglottal airpressure is maintained against a balanced glottal resistance by means of vocal fold adduction, the oscillation is sustained by myoelastic forces.

However, using stroboscopy, where a light source is set to flash at a rate slightly lagging the frequency of vibration of the vocal folds, vocal fold vibration has been observed in slow motion. This enabled researchers to observe that complex vocal fold vibration is further added to by a wavemotion in the wet mucosal 'Cover' (Fig.5) (Hirano, 1981).

Titze and Strong (1975) observed that:

> *'the 'undulation' in the mucosal layer involves the relative motions between the mucosa and the ligament vocalis and occurs whenever the vocal cord is unstretched. A surface wave is seen to propagate laterally from the glottis toward the vocal cord boundary...Due to the high surface tension of the mucosa the surface wave is readily dispersed, but occasionally gets reflected from the boundary and travels back towards the glottis."*

On the basis of this observation, they formulated a complementary theory to the Aerodynamic Myoelastic theory, which they call the **Muco - Undulatory** theory of vocal fold vibration.

Titze (1980) also emphasised the significance of the vertical motions in the mucosa for initiating and maintaining oscillation which, he believes, facilitates the interaction between aerodynamic energy and tissue dynamics. Vertical displacement can have a significant effect on vocal tract excitation.

Laver (1980) also states:

> *'During phonation, the crossectional shape of each vocal fold is subject to continuously changing deformation. Part of the dynamic deformation is attributable to the mucosal wavemotion travelling up the external surfaces of the vocal folds and into the ventricles, and part to the more gross displacements involved in the vertical phase differencs.'*
>
> (Fig. 5).

The effect of the airflow on the mucosa creates the lower and upper 'lip' in the Cover illustrated in Fig.5, and results in the vertical phase difference in the contact between the lower and upper part of the vocal fold. It is this which is held to be essential for the maintenance of vocal fold oscillation (Titze, 1980, Fujimura, 1981, Hiroto, 1981, Ishisaka,1981).

Using computer models of vocal fold displacement during phonation, Titze (1985) questions the need, stated in the original theory of Van den Berg (1968), for the pressure drop between the vocal folds to be negative to effect closure. He has demonstrated that:

> *'as long as the pressure varies in phase with the vibrations, this can be sufficient to cause glottal closure. The glottis may be shaped like a cone - convergent during opening and like a funnel - divergent - during closure'.*

By 1993 Berke and Gerratt categorically state:

> *'Mucosal wave movement is the primary means by which the larynx transforms the egressive pulmonary air flow into sound'.*

CHAPTER II - FUNDAMENTAL FREQUENCY.

1 Fundamental frequency control.

The physical correlate of perceived pitch is fundamental frequency. It is generated by the periodic oscillation of the vocal folds in the pulmonary egressive airstream. Electrolaryngography is used in this study for measurement of fundamental frequency as it allows reliable registration of successive vocal fold contacts and calculation of fundamental frequency (Fx).

Hollien (1981) summarising findings related to control of vocal fundamental frequency concluded it was achieved by interrelationships between the mass and stiffness of the vocal folds and the subglottal airpressure. Hirose and Sawashima (1981) suggested the main mechanism for pitch change was a variation in longitudinal tension (L.T.) of the vocal folds (Fig. 4, p.36).

Titze, Jiang and Drucker (1988) summarise the complex relationship in Fo control between stiffness and tension of the vocal folds and the 'Body' and 'Cover' structure (Fig. 6, p.40):

> *"The Cover is very pliable and has no contractile properties...being heavily irrigated with liquid, it can propagate a surface wave that facilitates energy transfer from the glottal airstream to the vocal fold tissues. Its tension is controlled by vocal fold length...The Body...is less deformable and has active contractile properties. Its tension is therefore not only determined by length but also by active stiffening of the muscle internally. The combined tension of the portions of the Body and Cover in vibration, would seem to regulate the fundamental frequency."*

The earlier researchers were thinking along the right lines. However, taking into account the complex influence of the layered structure of the vocal folds on vibratory mechanics, our understanding of fundamental frequency control has been taken a step further.

Primarily, Titze et al (1988) suggest, the fundamental frequency of vocal fold oscillation is controlled by the cricothyroid muscle (Fig. 2 and 3, p. 32 and 34). Contraction will always result in an increase in tension as a result of elongation of the vocal folds, and in combination with an increase in subglottal presssure (P_s), this will always result in an increase in pitch.

Next most influential in pitch control would be the Thyroarytenoid (T.A) (Vocalis) muscle (Fig. 2 and 3). Activity in this muscle, however, may either increase or decrease fundamental frequency, because of the complex way in which the effective length, tension and mass of the vocal folds change, when the (T.A) muscle contracts. Contraction i.e shortening of the vocal fold may reduce the effective tension *"...because the increased slackness of the vocal fold Cover, may dominate over the increased stiffness of the Body."*

The third most important factor in fundamental frequency control Titze et al suggest (1988) is subglottal pressure (P_s), which affects vocal fold tension as it regulates the amplitude of vibration.

Contraction of the cricothyroid muscles (Fig. 2 and 3) increases the distance between the thyroid cartilage and the vocal processes of the arytenoids.

In summary, changes in fundamental frequency perceived as changes in pitch, are achieved through modification of vocal fold length, mass and tension achieved by adjustments in the intrinsic laryngeal musculature, for extremes of pitch, also in the extrinsic musculature, and adjustment of the expiratory effort to control subglottal airpressure.

ii The role of respiration in Phonation and Fundamental frequency control.

Expiratory airflow provides the driving force which creates and sustains the subglottal pressure necessary to set the vocal folds into motion and maintain vibration. The mode of vocal fold contact and vibration is determined by the rate and volume of the airflow and the tension characteristics of the vibrating folds as described above.

The ability to sustain and control phonation for any vocal task, depends on the ability to maintain subglottal pressure by control of the expiratory airflow. The main source of the following account is Sundberg (1987).

In breathing, inspiration is an active process involving the contraction of the inspiratory intercostal muscles between the ribs resulting in an increase in ribcage volume. Relaxation of the same muscles results in a passive expiratory force. The lung volume at which the passive inspiratory and expiratory forces are equal, is called the **Functional Residual Capacity (FRC)**. If the lungs expand or contract beyond this point, passive forces try to restore them back to this volume.

The **Vital Capacity** is the difference between the amount of air that fills the lungs after a maximally deep inbreath, the **Total Lung Volume**, and the air left in the lungs after a maximal outbreath, the **Residual Lung Volume**. The Vital capacity varies between the sexes, with body size and with age. Women tend to have smaller vital capacity than men. Older people tend to have smaller vital capacity than younger people. (Ptacek, Sander, Maloney and Jackson, 1966).

Part of the vital capacity is air available for phonation. Only about 10 % of the vital capacity is used in quiet breathing, a volume just over the Functional Residual Capacity in the average adult male However, when we speak, we take a breath which allows us the use of approximately 50% of our vital capacity. The passive expiratory force is used to maintain subglottal pressure for phonation as we speak.

An interesting experiment reported in Sundberg (1987) shows the lung volumes used by a male subject speaking at normal volume and reading at normal, loud and at very loud volume. During spontaneous speech he used lung volumes between 55 % and 10% of his vital capacity. His Functional residual capacity was 35% of the Vital Capacity, and most of the time he used lung volumes below this value to speak on. In loud reading the subject used vital capacity values between 10% and 70%. In very loud reading these values varied between 15% and 95%.

The lung volumes used during loud, and very loud reading reflect the need for increased subglottal pressure against the increased adduction and tension in the vocal folds to produce large vibrational amplitude, which results in increased vocal intensity. It also tends to lead to increased air consumption.

Another important observation was that in reading aloud (not speaking) at normal and loud volume, the subject tended to replenish his airsupply when his lung volume was close to the Functional residual capacity: ***'...in reading, this subject tended to take advantage of the passive forces of exhalation in maintaining a suitable subglottal pressure for phonation.'*** Sundberg suggests the reason for replenishing the airsupply around the FRC in reading aloud may be because we feel physically more comfortable at lung volumes near to or above the FRC.

Grosjean and Collins (1979) reported that in reading aloud subjects tended to *'match inspiratory refills to constituent or sentence boundaries'*.

The comfortable maintenance of subglottal pressure during reading aloud, by taking a breath before the lung volume is below a value, where active expiratory force is needed, may explain the common finding of slightly higher mean fundamental frequency measurements during reading aloud than during conversational speech (Hollien and Jackson 1973, Schulz-Coulon, 1975, Ramig and Ringel, 1983). This has been found to be true in several languages. Hanley, Snidecor and Ringel (1966) compared Japanese, German and Spanish speakers, who all demonstrated this effect. The effect is noticeable in Tables 1 and 2 in the findings by Mysak (1959) and by Barry et al, (1990).

More consistent and better controlled subglottal pressure may result in more regular phonation during reading compared to conversation (Carlson, 1988, Kramer, 1989), which may be the explanation of the higher mean fundamental frequency measurements.

Schultz-Coulon, (1975) who first commented on the tendency for reading Fo being higher than speaking Fo, hypothesised that it might simply be due to subjects feeling more relaxed during spontaneous speech and counting numbers, than they did during reading aloud.

It is possible to change pitch by respiratory means, i.e. by varying subglottal pressure, rather than through laryngeal manoeuvres. It has been found however, to only enable a very small change in fundamental frequency (Hixon, Klatt and Mead, 1971). Netsell, (1989) found changes of only 2-6 Hz per cm H_2O. Sundberg (1987) mentions an increase in frequency by ½ a semitone per dB increase in speech loudness.

Titze (1980), in his comments on van den Berg's Myoelastic Aerodynamic theory of vocal fold vibration, concludes that *'muscular adjustments in the larynx acting in synchrony with abdominal effort can account for slow and deliberate fundamental frequency control in marked breath groups, but rapid variations, similar to those imposed artificially and involuntarily during phonation, can perhaps be explained by a myoelastic amplitude effect.'*

This indicates that Titze considers Fo changes mainly implemented by myoelastic, i.e. not aerodynamic, means *'...partially as a result of deliberate or reflex adjustments of laryngeal muscles, and partially as a result of nonlinear tissue strain over the vibrational cycle'.*

A comprehensive study by Shipp and McGlone (1971) combined electromyography, subglottal airpressure and airflow measurements in determining the contribution of muscle activity and subglottal pressure variation to changes in vocal pitch.

As subjects increased their pitch through their modal register into falsetto, there was first a gradual and parallell increase in Cricothyroid (CT) and Thyroarytenoid (TA) muscle activity, 28% and 22% respectively, up to about the 50% frequency point of the subjects' phonation range. This was also the point at which most subjects switched from modal to falsetto voice. At the 70% and 90%

frequency points the activity in the CT muscle increased by 49% and in the TA by 21 %. Subglottal airpressure levels increased consistently as the pitch increased. Sundberg (1987) seems to support this finding although, he does not prove it with electromyographic data, by stating that as soon as the vocalis muscles stop tensing the vocal fold, the voice changes into falsetto, i.e activity in the vocalis muscles ensures the voice is in modal register.

Shipp and McGlone's (1971) results, which also include electromyographic data from the activity in the Interarytenoid muscles, make them dispute van den Berg's (1968) suggestion that the Interarytenoid muscles are of any major importance in control of fundamental frequency. Nor did they find any consistent correlation between airflow measures and increase in pitch. There was no significant rise in airflow until the subjects phonated at their 70% frequency point.

iii Sex differences in voice fundamental frequency.

The different size, shape and tissue characteristics, of the male and female larynx give rise to the most obvious difference between male and female voices, the difference in pitch. This difference becomes dramatically emphasised at puberty with a more or less sudden drop in pitch of the male voice.

Measurements of laryngeal dimensions in men and women by Kahane (1978) on human cadavers, showed a 60% greater length of the membranous part of male vocal folds compared to the female. The anterior-posterior dimensions of the Thyroid cartilage showed only a 20% greater size in the male compared to the female. According to Titze (1989 a) this also applies to the lateral and vertical dimensions of the thyroid cartilage in Kahane's measurements:

> *'This suggests that the male vocal folds grow disproportionately in the antero-posterior direction. The primary growth is located in the anterior two thirds of the larynx, between the vocal processes and the anterior commissure.'* (Titze, 1989 a).

Titze concludes that the main factor determining the difference in fundamental frequency between male and female voices is the difference in length of the membranous portion of the vocal folds. This is on average 16 mm in the adult male and 10 mm in the female. He derives from this a 'scale factor' of 1.6 which explains the difference in fundamental frequency. The same scale factor also accounts for the differences in mean airflow and aerodynamic power, by which is understood the mean subglottal pressure multiplied by the mean airflow (Titze, 1989a).

Earlier data produced by Hollien (1962) comparing male and female vocal folds in terms of membranous length and thickness (depth) of the vocal folds shows a 20-30% greater thickness of the male folds. He shows how vocal fold thickness and length are inversely proportional in fundamental frequncy control. As frequency increases, the length of the vocal folds increase and the thickness decreases. Titze (1989a) suggests this is a result of 'conservation of tissue volume' and not of muscular adjustment. At low fundamental frequencies he found the mean thickness and the length were almost equal, indicating that the medial surface of the vocal folds in crossection, was nearly square. He suggests:

> *'If the effective mass per unit length is profoundly greater in the male than in the female, the difference is likely to be associated with depth of vibration rather than anatomical thickness. This could in turn be explained by a larger amplitude of vibration for the male.'*

Much research has been devoted to establishing normative data for vocal fundamental frequency in men, women, children and babies using a variety of speech tasks; speaking, counting, reciting, reading, sustaining vowels and singing. A number of different recording and analysis techniques have been employed to derive the fundamental frequency measurements from the complex acoustic data collected. Electroglottography (EGG) or Electrolaryngography (ELG), which is used in this study, was used for data collection in the studies by Abberton (1976), Pegoraro-Krook (1986) and Barry et al (1990) shown in Tables 1 and 2 below. Most of the data relates to reading, a few studies report fundamental frequency data relating to conversational speech or monologues.

Hirson and Roe (1993) point out, however, that a rigorous definition of 'normal voice' has still not been achieved because of a failure to control for all the variables that may influence the voice in all its aspects. These variables will be closer examined in the discussion on stability of voice fundamental frequency.

In this study subjects' voices are recorded using ELG and Fx parameters derived both for conversational speech and reading aloud are used to reflect a range of habitual phonatory behaviours.

iv The effect of advancing age on voice fundamental frequency.

The male voice deepens towards middle age and then tends to increase in fundamental frequency towards old age. This is noticeable in Table 1 below in the data reported by Mysak, (1959), Hollien and Shipp (1972) and Pegoraro-Krook (1986). Among women (Table 2) the fundamental frequency also tends towards a slight lowering towards middle age but then a further decrease with increasing age (Abberton 1976, Stoicheff, 1981, Pegoraro-Krook, 1986, Brown, Morris, Hollien and Howard, 1991).

The reasons, for the different development in fundamental frequency characteristcs with increasing age between men and women, are put down to different physiological changes in the sexes with advancing age. One contributing factor may be the calcification of the laryngeal cartilages, which proceeds at different rates between individuals and particularly, happens more slowly and proceeds less far in women than in men (Kahane, 1983). Another factor is the atrophy of laryngeal muscles with advancing age. This results in reduced tension and sometimes bowing of the vocal folds on phonation in elderly men. The mucous membrane covering the folds may also change colour as a result of fat degeneration or keratosis.

TABLE 1

Speaking Fundamental Frequency for normal adult males.

AGE Mean	 Range	N	MEAN Fo Hz	Language	Reported by
	19-24	8	123,0	Br. English	Barry et al, 1990
20,3	17,9-25,8	157	123,3	Am. English	Hollien and Jackson, 1973
	20-29		112,0	Swedish	Pegoraro-Krook (1986)
47,9	32-62	15	100,0	Am. English	Mysak, 1959
73,3	65-79	12	119,3	- " -	-"-
85,0	80-92	12	136,2	- " -	-"-
	80-89		124,0	Swedish	Pegoraro-Krook (1986)

Reading Fundamental Frequency for normal adult males.

AGE Mean	 Range	N	MEAN Fo Hz	Language	Reported by
	19-24	8	132,0	Br. English	Barry et al, 1990
24,4	20-29	25	119,5	Am. English	Hollien and Shipp, 1972
26,0	20-35	15	118,0	Am English	Brown et al, 1991
34,9	30-39	25	112,2	- " -	Hollien and Shipp, 1972
45,4	40-49	25	107,1	- " -	- " -
44,0	40-55	15	100,0	- " -	Brown et al, 1991
47,9	32-62	15	113,2	- " -	Mysak, 1959
54,1	26-79	65	112,5	Japanese	Horii, 1975
54,3	50-59	25	118,4	Am. English	Hollien and Shipp, 1972
64,6	60-69	25	112,2	- " -	- " -
75,0	65-85	15	127,0	- " -	Brown et al, 1991
74,7	70-79	25	132,1	- " -	Hollien and Shipp, 1972
83,6	80-89	25	146,3	- " -	- " -
85,0	80-92	12	141,0	- " -	Mysak, 1959

TABLE 2

Speaking Fundamental Frequency for normal adult females.

AGE *Mean*	*Range*	*N*	*MEAN Fo* *Hz*	*Language*	*Reported by*
	19-24	*10*	*207.0*	*Br English*	*Barry et al, 1990*
	20-29		*196.0*	*Swedish*	*Pegoraro-Krook (1986)*
	80-89		*188.0*	*- " -*	*Pegoraro-Krook (1986)*

Reading Fundamental Frequency for normal adult females.

AGE *Mean*	*Range*	*N*	*MEAN Fo* *Hz*	*Language*	*Reported by*
	18-19	*10*	*241.6*	*Br.English*	*Abberton, 1976*
	19-24	*10*	*207.0*	*Br English*	*Barry et al, 1990*
24.6	*20-29*	*21*	*224.3*	*Am. English*	*Stoicheff, 1981*
28	*20-35*	*20*	*192*	*- " -*	*Brown et al, 1991*
33.5	*30-40*	*9*	*196.3*	*- " -*	*Saxman and Burk, 1967*
35.4	*30-39*	*18*	*213.3*	*- " -*	*Stoicheff, 1981*
44	*40-55*	*10*	*195*	*- " -*	*Brown et al, 1991*
	45-50	*6*	*180.3*	*Br.English*	*Abberton, 1976*
54.5		*17*	*199.3*	*Am. English*	*Stoicheff, 1981*
65.8	*60-69*	*15*	*199.7*	*- " -*	*- " -*
75.4	*over 70*	*19*	*202.2*	*- " -*	*- " -*
79	*65-85*	*19*	*175*	*- "-*	*Brown et al, 1991*

In old women, the vocal folds tend to be oedematous, possibly related to hormonal changes and hormone imbalance during the menopause (Honjo and Isshiki, 1980).

Ptacek et al (1966) examined vocal differences between younger adults, aged under 40, and 'geriatric' adults, aged 65 and above, using measurements other than mean fundamental frequency. They chose to examine changes with advancing age in respiratory, articulatory and phonatory ability. Subjects were asked to perform tasks at the limits of their functioning as opposed to habitual functioning. The number of subjects in each of four groups of men and women was 31.

Findings showed a significant reduction in aged subjects' total pitch range, vital capacity, maximum vowel duration and intensity and in maximum intraoral pressure. This was the case for both men and women. Ptacek et al conclude that these differences may be explained by a decrease in the power of the respiratory muscles, a loss of elasticity of lung tissue and degenerative changes in laryngeal musculature as a result of ageing.

The trends in changing fundamental frequency with age shown in Tables 1 and 2 were confirmed in a recent comparison of fundamental frequency during a reading task by 60 professional singers and 94 non-singers in the age groups 20-35, 40-55 and 65-85 (Brown et al,1991). The fundamental frequency patterns for the untrained subjects followed predicted trends according to human ageing (Tables 1 and 2). However, trained singers were found to maintain essentially unchanged speaking fundamental frequency through most of their adult life.

v The effect of smoking on fundamental frequency.

Long term smoking may produce generalised thickening and oedema of the larynx. It may lead to localized oedema of the vocal folds and/or polypoidal changes (Myerson, 1964). It may further lead to epithelial changes, leukoplakia and hyperkeratosis, all recognised as precursors of laryngeal carcinoma (Wynder, Covey, Mabuch and Mushinski, 1976, Wynder and Stellman. 1977, Shaw, 1979, Burch, 1981, De Stefani, Correa, Oreggia et al, 1987)

In the long term vocal fold mass increases in smokers and chronic inflammation, particularly oedema on the free vibrating edge of the vocal fold, may lead to changes in the vibratory pattern (Abberton, 1976, Rainbow, 1985, Murphy and Doyle, 1987, Comins, 1988). The effect of the nicotine on the cardiovascular system and on vasoconstriction results in poorly oxygenated blood supply to the vocal folds and may contribute to mucosal atrophy (Myerson, 1964).

Gilbert and Weismer (1974) found significantly lower fundamental frequency in women who smoked. This has since been confirmed by Abberton (1976) and Stoicheff (1981). Sorensen and Horii (1982) analysed fundamental frequency parameters in men and women, smokers and non-smokers, during reading, conversation and sustained phonation. They found significant differences between smoking and non-smoking men for reading and conversation, but not between the women. Comins (1988) found significant differences in fundamental frequency measures in reading aloud between her 'old' male smokers and non-smokers, aged 45-60 (N=8), but not between her 'young' smokers and non-smokers, aged 21-30 (N=8). This may indicate, that the effect of lowering fundamental frequency increases with years spent smoking.

Instead of studying groups of men and women, smokers and non-smokers, Murphy and Doyle (1987) studied the effect of temporary cessation of smoking in two adults, one woman aged 42 and one man aged 30, who were longstanding smokers, 24 and 17 years respectively, and smoked at least 25 cigarettes per day. They were compared to two age and sex matched non-smoking subjects.

Results indicated that cessation of smoking for 40 hours resulted in an increase in fundamental frequency in both smokers. This effect remained for the first day after resumption of smoking, after which the fundamental frequency returned to its previous level. The effect was observed again on a second trial with the same subjects. The speaking tasks consisted of a monologue, reading of the first paragraph of the Rainbow passage (Fairbanks, 1960) and sustained phonation on [ɑ], at comfortable pitch and loudness.

Comins (1988) used a questionnaire to compare the reported number of cigarettes smoked by her subjects to an objective measure of the amount of Carbon Monoxide in exhaled air. She found that the number of reported cigarrettes smoked per day was not a reliable measure. There was a highly positive correlation, however, between measured levels of Carbon Monoxide and fundamental frequency variance, indicating greater Fo variability in the heaviest smokers.

vi Stability of voice fundamental frequency in normal speakers.

> *"Fluctuations of laryngeal measures are dependent on the prosodic structure of the speech being analysed. Given the limited number of intonational patterns that exist in a language, there is a statistical element in the amount of speech that is neccessary to achieve a representative sample of these patterns. However there is no way of calculating the amount because the units of the patterns are not defined in physical terms, and are dependent on the intonational model adopted"* (Barry et al, 1990 a).

This paragraph summarises some of the important reasons for the difficulties encountered by anyone attempting to develop norms for speaking fundamental frequency. A number of studies report on the optimum length of the sample needed to get a representative sample of an individual's habitual voice. It has been found to be 2 minutes of speech or reading (Green, 1972, Mead, 1974, Markel and Davis, 1979, Hiller, Laver and MacKenzie, 1984, Barry et al, 1990 b).

In an in-depth study of four normal speakers, two women and two men, Barry et al (1990 b, 1991) using the Fourcin Electrolaryngograph recorded speakers reading their 'Environmental Passage', followed by a 13 minute reading of a book of short stories. These recordings were interspersed with the reading, eight times in succession of the 'Environmental Passage' to avoid "fluctuation in interest and excitement" for comparison with the unprepared reading passage. Another task consisted of a free monologue of 15 minutes on a topic of the speaker's choosing and finally a dialogue was recorded with two of the speakers in conversation.

Analysis of one minute samples of the longer texts showed fluctuations in mean Fx values in the short term of up to 23 Hz in the women and up to 18 Hz in the men. However, it was found that the effect on the overall mean of extremes in short term fluctuation, decreased with the duration of the sample. They confirmed that a two minute passage was a safe basis from which to calculate personal voice frequency characteristics (Barry et al, 1990 b, 1991).

However, further analysis of mean fundamental frequency differences across and within tasks, lead them to conclude that:

> *"The differences between the speakers in the repeated passage task, and the random distribution of large and small fundamental frequency differences for two minute stretches across the different tasks, indicate that there is strictly speaking no such thing as a generally valid personal voice frequency value. Each situation has its specific and individual effect on a speaker's voice frequency."*

In an earlier study comparing Fx values for two groups of normal speakers aged 19-24, significant differences were found between the Mean and Modal values of the 'Environmental Passage' a 'Numbers Passage' containing listings of numbers and a 'Free Monologue' (Table 3) (Barry, Goldsmith, Fourcin and Fuller, 1990 a). The differences between the 'Numbers Passage' and the 'Free Monologue' were not significant.

Table 3

Task	Environmental Passage		Numbers Passage		Free Monologue	
	Mean (Hz)	Mode (Hz)	Mean (Hz)	Mode (Hz)	Mean (Hz)	Mode (Hz)
Women (N=10)	216	210	210	202	207	202
Men (N=8)	132	126	126	120	123	117

Another measure investigateed by Barry et al, 1990 b) was frequency range across different speech tasks. Similar to the mean, there was found to sometimes be large fluctuations between different two minute stretches of speech samples. Ranges also tended to grow wider towards the end of the day. However, maximum and minimum range measures were less random and showed a smaller variation for the repeated reading task than the free reading task. More importantly, it was found that the lower frequency range limit was generally less variable than the upper one. This was explained by it being closer to the physiological limits of the laryngeal vibratory mechanism. It also seemed independent of the type of speech task (Barry et al, 1990 b).

Frequency range measurements have been found to vary greatly between groups of speakers of the same language (Snidecor, 1943, Mysak,1959, McGlone and Hollien, 1963, Linke, 1973, Graddol, 1986). Baken (1987) warns against using a 'Total Range' measure as it is likely to include extreme values that are due to what he calls 'slips of the larynx'. Instead, range values tend to be expressed as the '90 % range' including only those values within which 90 % of the frequency distribution falls. These were used by Barry et al (1990 a and b) and will also be used in the current study.

Norms for speaking fundamental frequency measurements need take into account not only the subjects' age, sex, speech task and language background. Other variables that have been found to affect measures are regional accent and social class (Trudgill, 1974, Graddol and Swann, 1983), race (Hudson and Holbrook, 1981, 1982), length of speech sample (Fitch and Holbrook, 1970, Horii, 1975), time of day (Garrett and Healey, 1987, Barry et al, 1990 and 1991, Hirson and Roe, 1993) and even the emotional content of speech (Williams and Stevens, 1981, Sallinen-Kuparinen, 1985).

Baken (1987) suggests that the Modal frequency of a voice has the advantage over the Mean in that it is the most common frequency value and thereby most closely approximates the person's 'habitual pitch'. This was found in the study by Barry et al (1990 b, 1991) who found modal frequency values less task sensitive and more stable than the means. Neither a monologue nor a dialogue showed consistently lower values than the reading tasks. They still caution against using either as sole measure of personal voice characterisation, however.

Garrett and Healey (1987) found systematic fluctuations in individual speakers recorded at three different times of the day. Barry et al (1990 b, 1991) reported three of their four speakers showing a marked increase in pitch from the morning (9.40 am) to the midday (12.10 pm) recording and then a decrease from midday to the afternoon (3.30 pm).

The one speaker who did not show the increase did show a significant decrease from midday to afternoon.

An interesting attempt at sampling and measuring fundamental frequency parameters of 'natural speech' during a full working day is reported by Ohlsson (1988). A 'Voice accumulator' was developed, which was capable of measuring pitch and phonation time during a 12 hour period. The voice signal was recorded from the front of the neck using a contact microphone, which was worn throughout the day. A microcomputer program was used to analyse the accumulated data to compare average fundamental frequency used during the day, or any part of the day; fundamental frequency range expressed in standard deviation of the mean Fo and phonation time expressed as a percentage of a given time interval.

Two groups of professionals were compared over two days, 10 nurses and 10 Speech and Language (SL) therapists. There were no significant differences found in accumulated phonation time between the two groups of speakers, although the SLT group showed higher values for both days. A significant difference was, however, found between their average fundamental frequency, which was higher for the nurses than for the SL therapists.

No systematic variation or pattern was observed in fundamental frequency use through the day, but SL therapists tended to adjust their speaking fundamental frequency according to the communicative situation. They used systematically lower fundamental frequency during voice therapy sessions than e.g. during coffee breaks. Nurses did not show any such variation according to situation. There was found to be considerable variation in fundamental frequency and phonation time throughout the day, both within and between speakers.

The reason given for the differences between the groups are related to the SL therapists' having had voice training, which has been found to result in lowered fundamental frequency (Fex, 1987). The SL therapists reported fewer voice symptoms e.g. hoarseness, temporary or lasting, and 'throat pain' (without an infection) after speaking or reading loudly, than nurses.

Coleman and Markham (1991) investigated speaking fundamental frequency variation (SFF) using recordings and a 'Voice Identification Voice Monitor' for analysis in one group of 11 female graduate students reading the second paragraph of the 'Rainbow passage' (Fairbanks, 1960, Appendix 7) at least every three days for one month; another group of six 'model' speakers were televison newsreaders, who were recorded reading the news a minimum of six times over two weeks; the recorded sermons over a seven year period of one individual minister and similarly, a folk story teller recorded over a ten year period, were also analysed. In the latter cases two minute samples of text was used for the analysis.

Frequency values were converted to semitones to enable comparison of data from speakers of different sex and age. It was found that speakers could be expected to be within 3 semitones of their average speaking fundamental frequency for any single sample at least 90% of the time. *"Clinical experience suggests that a +/- 3 ST range of variation from repeated measures of SFF is a realistic value....a patient's SFF that varies by as much as 18% from one day to another is still within normal limits."*

Coleman and Markham (1991) caution against reading too much into reports of significant differences of less than 10-15% within individuals. They do however seem to consider SFF a stable enough measure that varies relatively little over time periods of a month, a year or even several years, to be useful for intra-individual comparisons: *"When a clinically significant change (here estimated as a 3 ST shift in SFF from an average value) does occur, it may be considered to be important with respect to the individual's emotional or physical health."*

vii Fundamental frequency measures in pathological voices.

The interest and effort spent in developing normative data for objective voice parameters is fuelled by the need to develop means of measuring and recording change in such parameters in pathological laryngeal states before and after treatment. As shown above there are considerable difficulties in developing such norms in view of the variation found within individuals, let alone between individuals without laryngeal pathology.

There are few studies which compare frequency measures of normal and pathological voices over longer stretches of speech or reading. Until recently, speech technology did not allow analysis of more than a few seconds of phonation, which accounts for the great number of studies that report measurements taken from spoken or repeated sentences.

Murry (1978) extracted the third sentence from a reading of the first paragraph of the "Rainbow Passage" and found that the mean speaking fundamental frequency, the standard deviation and the semitone range of the voices of patients with vocal fold palsy were significantly reduced compared to a sample of normal speakers. However, Mean speaking fundamental frequency failed to separate normal speakers from two other groups of pathological speakers; a group with benign mass lesions and a group with cancer of the larynx. Hecker and Kreul (1971) found that patients with laryngeal cancer had more restricted pitch range than normal speakers reading the second sentence of the "Rainbow Passage".

Lehman et al (1988) found no significant difference in mean fundamental frequency between a group of normal speakers and a group of patients irradiated for T1 laryngeal carcinoma. The tasks performed were a two minute conversation and a reading of the "Rainbow Passage", and also sustained phonation on [ɑ], [i] and [u].

An extensive study of differences between 'normal' and pathological speakers' phonation parameters is offered by Hirano et al (1991). They measured eight parameters of phonatory function in 40 normal speakers and compared them to measures obtained from 1563 voice patients presenting with 22 different laryngeal pathologies. The measurements were obtained with the subjects 'phonating into a mouth piece'. It is assumed, but the authors do not actually state this, that the phonatory task consisted of a sustained vowel.

Findings indicated that habitual pitch, Fo_{HAB}, was significantly higher in male patients with sulcus glottidis, glottic and supraglottic carcinoma. This is assumed to be a result of an increase in stiffness in the vocal fold cover. Mutational disorders in males also showed an increase in Fo_{HAB}. No pathological groups among females showed higher than normal Fo_{HAB}.

Lower than normal Fo_{HAB} was found in males with acute laryngitis and Reinke's oedema. This is attributed to an increase in mass and a decrease in stiffness of the vocal fold cover. It was also found in females with acute and chronic laryngitis, nodules, polyps, Reinke's oedema, sulcus, cyst, granuloma, carcinoma, vocal fold paralysis, hyper and hypo-functional voice disorders.

Fo_{H} refers to 'highest physiological tone' and was significantly lower for almost all disease groups than for normals, for both sexes.

Fo_{RG}, or 'physiological range of phonation' was significantly reduced for most pathologies among the male speakers and in females for Reinke's oedema, granuloma, hyperplasia, paralysis, trauma and hyperfunctional dysphonia.

"Lowest physiological tone", Fo_{L} was significantly higher than normal in males with sulcus glottidis, granuloma, papilloma, glottic and supraglottic carcinoma and mutational disorders. The lowest Fo_{L} was found in males with Reinke's oedema.

Finally, "habitual loudness of phonation" SPL_{HAB}, was found to be significantly greater than normal in males with nodules and females with polyps indicating their tendency to phonate loudly.

In most pathological groups Hirano et al (1991) found reduced ability to phonate at loud intensity, and also the available range of intensity was significantly reduced. They attribute this to decreased amplitude of vocal fold vibration and possibly also to incomplete glottic closure. Fundamental frequency and intensity related parameters used in their study reflected the effects of treatment and are therefore recommended for use in evaluation of voice disorders before and after treatment.

Hillman, Holmberg, Perkell et al, (1989) evaluating objective assessment parameters in vocal hyperfunction, suggest there are different undelying mechanisms producing vocal nodules and polyps as opposed to contact ulcers. In the former they found abnormally high transglottal pressures, reflecting increased vocal fold stiffness. Contact ulcers on the other hand were associated with normal pressures but abnormally low fundamental frequencies which they suggest indirectly reflect decreased vocal fold cover stiffness.

viii Perturbation in vocal fold vibration

Perturbation is described by Colton and Casper (1990) as small rapid cycle to cycle changes of vocal fold period and amplitude of vibration, that occur during phonation as a result of slight differences in mass, tension and biomechanical properties of the folds. Verdolini-Marston, Sandage and Titze (1994) showed significant reduction in Fo perturbation (jitter) in a group of subjects undergoing a five day course of increased hydration compared to a placebo and a dehydration condition.

Baer (1979) has suggested that frequency perturbation or 'Jitter', may be caused by slight variations in the neural control of the vocal folds.

Hollien, Michel and Doherty (1973) devised a measure that they decribed as a 'jitter factor'. It is calculated as the deviation of one period of vibration from the period of an adjacent one, averaging the differences obtained by calculating a series of such differences and dividing by the average period. The 'jitter factor' is obtained by multiplying the result by 100. This way of calculating jitter shows a negative correlation between fundamental frequency and jitter in that large cycle to cycle differences are associated with long fundamental periods and thereby high perturbation measurements, the higher the fundamental frequency the smaller the perturbations and the lower the jitter (Lieberman, 1963, Beckett, 1969, Koike, 1973, Horii, 1979, 1980).

Hecker and Kreul (1971) and Sorensen and Horii (1984) avoid this bias by using what they call the Directional Perturbation Factor (DPF). This measure ignores the magnitude of period perturbation and only takes into account the number of times the differences change direction. DPF is expressed as a percentage of the total number of differences for which there is a change in algebraic sign.

To reduce the influence of slow Fo changes during sustained phonation on the jitter measure, Koike (1973) developed a measure of Relative Average Perturbation (RAP). This was calculated as the deviation of a period from the average of that period and its immediate neighbours. It is usually expressed as a percentage (Takahashi and Koike, 1975).

Several other jitter measures have been developed but none has been found to give a measure completely uninfluenced by mean fundamental frequency. Baken (1987) suggests that *"..clinicians should expect relative perturbation to be somewhat higher in high frequency voices, while absolute jitter magnitude should decrease with increasing fundamental frequency."*

Amplitude perturbation or 'shimmer' is calculated in the same way as the jitter factor, this time measuring the size of cycle to cycle amplitude changes in the vocal fold signal (Haji, Horiguchi, Baer and Gould, 1986). Shimmer is however quite a cumbersome measure to calculate. In recent years it has been made easier by the development and increasing use of computer programs (Rontal, Rontal, Jacob and Rolnick, 1983, Karnell, 1991).

During speech there are voluntary changes in fundamental frequency and loudness of the vocal fold signal that are linguistically and paralinguistically determined. To reduce the influence of these changes on the calculation of jitter, measurement of Fo perturbation is normally carried out on sustained vowel phonation at a comfortable fundamental frequency and loudness (Murry and Doherty, 1981, Sorensen and Horii, 1983, Rontal et al,1983,). Johnson and Michel (1969) found a tendency for high vowels to show more jitter than low ones. However, Horii (1980) found significantly more jitter in [ɑ] and [i] than in [u], so did Bradley (1985) and Lehman et al (1988).

At the onset and termination of phonation there is more perturbation than during middle sections of sustained phonation. Haji et al (1986) usig a Fourcin Electrolaryngograph, selected 50 cycles of sustained phonation on [ɑ] measured at a point 250 ms from the beginning of the sample. This was also the number of fundamental frequency periods used for perturbation and signal to noise ratio analysis used by Eskenazi,Childers and Hicks, (1990). Karnell (1992) found however that the minimum sample 'window' of vibratory cycles neccessary for calculation of representative jitter was 190 cycles and for shimmer 130 cycles. He suggested that pathological voices may need a longer analysis window than normal voices.

Orlikoff and Kahane (1991) consider the effect of different intensity levels on vocal jitter. Their experiment, using normal males phonating the vowel [ɑ] in modal voice, at three different sound pressure levels trying to keep the fundamental frequency equal during increase in loudness, confirmed their theory that jitter decreased with increase in sound pressure level. They showed a significant correlation of -0.87 ($p < 0.001$). There was a similar but less strong relationship between shimmer and sound pressure level, -0.54 ($p < 0.01$). Mean shimmer magnitude was significantly greater during soft phonation than in loud phonation.

Few studies report perturbation measures calculated on long samples of speech or reading. This may be due to the very complex, precise measurements required as described above.

In the early stages of this study there were no means included in the computer program of calculating perturbation. However, it offered 2nd and 3rd order fundamental frequency (Dx) plots. The recorded Fx data was 'cleaned' by requiring adjacent pairs and triplets of vocal fold vibrations, respectively, to fall into the same frequency 'bin' along the logarithmic frequency axis, for admission to the next order distribution. This will be explained in greater detail in Chapter VII. However, 2nd and 3rd order distributions do reflect the degree of vocal fold regularity of vibration.

Abberton (1976) using Electrolaryngography and comparing second and third order Dx plots of reading samples produced by young and middle aged women found a tendency of decreased regularity of vocal fold vibration with age.

The sample size data pertaining to 2nd and 3rd order distributions has been used in this study to calculate what will be called, %TS, a vocal fold 'Regularity of vibration' measure. It is calculated as the proportion of the Total Sample of recorded periods carried into second order Dx distributions.

%TS is a gross measure of regularity. Not only is it influenced by the predetermined Fx range of logarithmic frequency 'bins', but it is inevitably also affected by prosodically determined fundamental frequency transitions as well as variation resulting from vocal fold asymmeties as described above. It will, however, be shown to be an extremely useful objective voice quality measure, related to the degree of frequency perturbation, or rather, as %TS is a measure of regularity, the lack of it. The advantage is, however, that it is easy to calculate and is based on long recorded samples of speech or reading (Kramer, 1989, Carlson, 1993 a, b).

ix Jitter and shimmer in vocal fold pathology.

Degrees of aperiodicity or perturbation in vocal fold vibration give rise to perceptually 'Harsh' or 'rough' voice quality. Large perturbations may be the result of reduced control over the phonatory mechanism due to emerging pathology (Beckett, 1969, von Leden, Moore and Timcke, 1960).

It is generally agreed that asymmetrical structural changes of the vocal folds will result in increased frequency jitter or amplitude shimmer. A change in mass, stiffness or shape in either of the three layers of the vocal folds, the epithelium and the superficial layer of the lamina propria, the intermediate and deep layers of the lamina propria, and/or he Vocalis muscle (Body) will lead to perturbed vibratory pattern (Laver, Hiller and MacKenzie-Beck, 1992).

Wendahl (1963) carried out an experiment in which he asked a great number of listeners to judge the degree of perceived 'roughness' in synthetically produced sounds. He varied the fundamental frequency and the amount of variation in frequency between successive cycles, the jitter. His results showed, that the greater the deviation from the fundamental frequency, the greater the perceived roughness. Even as little deviation as +/- 1 Hz sounded rough, but the roughness seemed greater if the jitter was superimposed on a lower fundamental frequency than on a higher one. It confirmed Hess's (1959) findings that harshness in high pitched voices is deemed less severe than in low pitched voices. The perceived roughness is the same whether it is caused by jitter or shimmer (Wendahl, 1966).

Patients with different laryngeal pathology will demonstrate degrees of both frequency 'jitter' and amplitude 'shimmer' in the Fo signal. As described above, many attempts have been made at developing ways of calculating perturbation, not least for use as an objective measure of vocal pathology (Lieberman, 1961, 1963, Hecker and Kreul, 1971, Horii, 1979, Rontal et al, 1983, Murry and Doherty, 1981, Askenfelt and Hammarberg, 1980, 1986, Wolfe and Steinfatt, 1987, Karnell, 1991).

Takahashi and Koike (1975) found significant correlations between perceptual 'roughness' and both amplitude and frequency perturbation quotients. 'Breathiness' correlated only with amplitude perturbation.

Haji et al, (1986) using electrolaryngographic measurements, found significant correlations between mean frequency perturbation, mean amplitude perturbation and perceptual ratings of 'hoarseness'. Mean fundamental frequency perturbation (jitter) only differentiated moderately well severely hoarse voices. However, mean amplitude perturbation differentiated between slightly, moderately and severely hoarse voices.

Beckford, Mayo, Wilkinson and Tierney (1990) using electroglottography found a fivefold increase in jitter values in a group of women after endotracheal intubation for surgery compared to a group of women who did not undergo surgery.

Askenfelt and Hammarberg (1980, 1986) report seven different procedures for speech waveform perturbation analysis in running speech. They compared measures of perturbation to perceived degrees of deviant voice quality in a group of 41 patients before and after successful voice therapy. They found that all voices with improved voice quality showed significant decrease in perturbation measures.

Hammarberg, Fritzell, Gauffin and Sundberg (1986) found significant correlations between waveform perturbation and perceptual voice qualities such as instability, roughness, flutter, diplophonia and creakiness/vocal fry. It indicates that perturbation was a significant factor in producing these perceptual qualitites. Degrees of perturbation did not differentiate between them, however. They also found that a certain amount of perturbation is part of normal voice quality, particularly in men who use creaky voice. There were consequently some pathological male voices which did not differ significantly in the degree of jitter from normal male voices

Orlikoff and Kahane (1991) state an important fact *".... short term frequency and amplitude variability may be influenced by any number or combination of physiologic phenomena. Hence, the degree of measured perturbation merely provides an acoustic indication of the stability of the physiologic balance or of the disruption of one or more of the constituent forces."* This highlights the fact that neither perturbation measure needs to be the result of changes in only one physiological parameter and, apart from the effects of pathological states on aerodynamic and muscular forces, speakers may react and compensate in different ways for pathological changes in the larynx.

"..*the existence of a pathology in the larynx does not necessarily differentiate a pathologic voice from a normal voice*" (Koike, Takahashi and Calcaterra, 1977).

Orlikoff and Kahane (1991) suggest that the fact that perturbation measures are sensitive to sound pressure levels makes it imperative that vocal intensity is controlled and accounted for, if perturbation is used for clinical comparison. They question the usefulness of shimmer (amplitude perturbation) as an acoustic index of vocal pathology due to its high intra- and inter-speaker variability.

Chapter III will look in greater detail at attempts at finding acoustic, objective correlates of perceptual voice quality features.

CHAPTER III - VOICE QUALITY

i Vocal Registers

Voice researchers tend to agree on the existence of at least three, perceptually different vocal 'registers' or 'Phonation types' (Laver, 1980) characterised by different modes of vocal fold vibration and frequency ranges. Hollien (1974) describes Falsetto, Chest and Creaky register. Each register is defined by characteristic fundamental frequency ranges and by different relationships between the open and closed phase of vocal fold vibrations and how these are achieved. Typical frequency ranges for the three main registers are suggested for male American English speakers:

Falsetto	275-634 Hz	(Hollien and Michel, 1968)
Modal	94-287 Hz	- " -
Creak	24- 52 Hz	(Michel and Hollien, 1968)

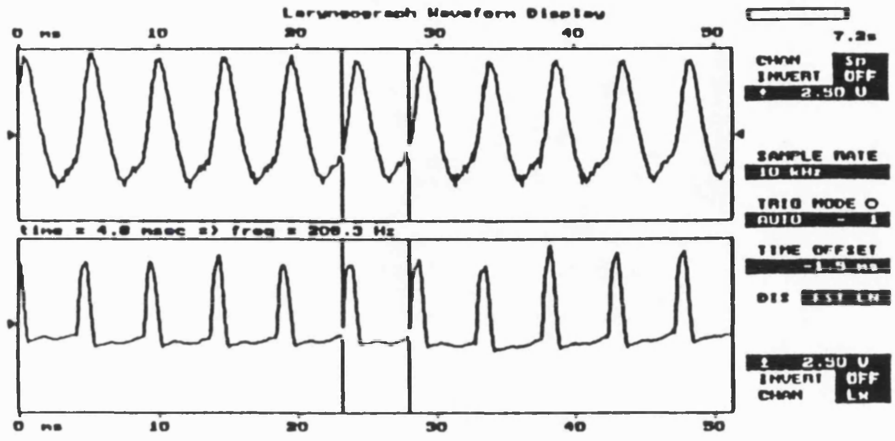
Fundamental frequency at the extremes of the register ranges tend to overlap. For instance high modal and low falsetto voice range and low modal and high creaky voice range.

Monsen and Engebretson (1977) found that creaky voices consistently showed fundamental frequencies below 100 Hz. Ranges varied between 30 and 90 Hz.

Perception of modal, falsetto or creaky voice is determined by the duration of the closing and opening phases of the vibratory cycle (Abberton, 1976, Fourcin, 1981 Carlson, 1993 b). In falsetto voice production these are of approximately the same short duration (Fig. 8 a). In modal voice the vocal fold closure occurs much faster than the opening and the duration of the closed phase is longer (Fig. 8 b).

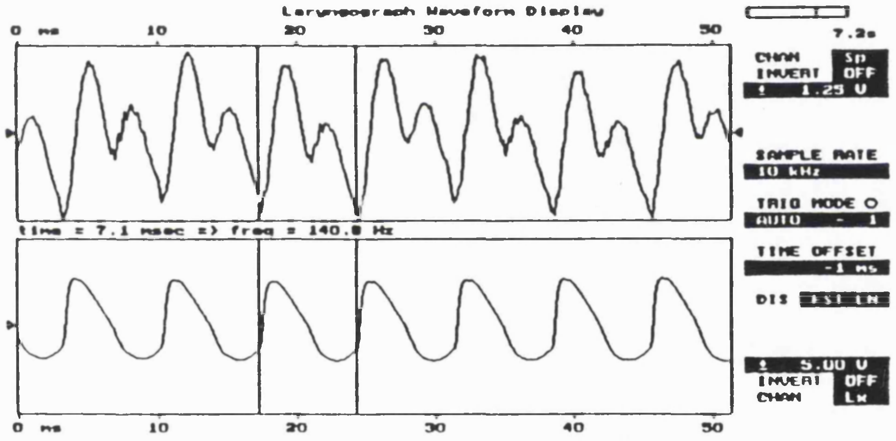
a)

High [i]



b)

High [i]

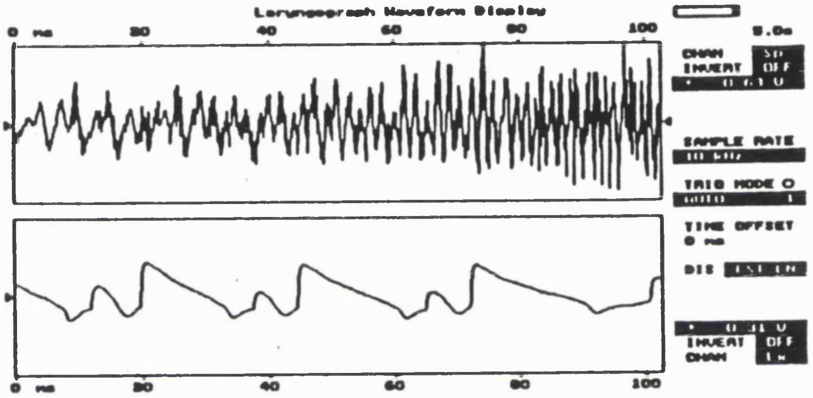


Lx and Speech waveforms

a) Falsetto voice
b) Modal voice

Figure 8

In creaky voice closing is rapid but there is a very long closed phase, a slow opening and an extremely short duration open phase (Fig. 9). Often periods of long and short duration alternate, creating what Monsen and Engebretson (1977) termed 'double pulse' phonation. This gives rise to highly irregular glottal waveforms This is illustrated in Fig. 9, which shows a typical Lx waveform pattern of 'double pulse' creaky phonation. Fourcin and Abberton (1971) also describe a single pulse creaky voice characterised by long closure and a very slow opening phase.



Lx waveform of creaky voice

Figure 9

The diagram in Fig. 3 (p. 34) illustrates the differential effect on the length and crossection of the vocal folds by the contraction of the intrinsic laryngeal muscles.

Modal Voice or chest voice is the 'physiological register of speech' (Hollien, 1974). Tension in the Vocalis muscles increases to attain higher fundamental frequency in chest register. This is compensated for by reduced passive tension in the vocal ligaments (Fig. 3, p. 34) (Laver, 1980). The folds are rounded in the frontal projection and *'Between their upper horizontal and subglottal oblique surfaces a vertical margin measuring several millimeters can be clearly made* *out' (Fig.5, p. 38). Vibration (occurs) with wide amplitudes and distinct phases of contact between the folds, causing complete vibratory closure of the glottis during about one third of the entire period'* (Kitzing, 1986).

Keidar, Hurtig and Titze (1987) found that trained listeners were able to distinguish perceptually with great consistency, pitch transition points between falsetto and chest (modal) register in different sung sequences of notes. They concluded that:

> *'the transition from chest to falsetto is primarily quality dependent...it is the shape of the waveform rather than its periodicity that is most salient in the perceived transition between the primary registers.'*

According to Laver (1980), modal voice is produced with moderate degrees of Adductive Tension, Medial Compression and Longitudinal Tension (Fig.4, p. 36). Increase in the Longitudinal Tension results in increased fundamental frequcny.

Complete glottal closure during the vibratory cycle is, however, not always the case during modal phonation in women. Recent studies suggest that the majority of women show a gap in the cartilaginous glottis during modal phonation (Biever and Bless, 1989, Södersten and Lindestad, 1990, Södersten, Hertegård and Hammarberg, 1994)

The closing phase of vibration in chest (modal) voice occurs extremely rapidly as the folds are sucked together in depth as a result of the Bernoulli effect (Fig. 7, p. 40). Kelman (1981) found closing time approximately the same within the same speaker irrespective of what vowel was sustained. Opening occurs more gradually from below upwards and from posterior to anterior. As fundamental frequency is increased, opening becomes more rapid and the open time shorter. (Abberton, 1976, Kelman, 1981). Durational relationships between the different phases of vocal fold vibration is well illustrated by Lx waveforms (Fig. 8 and 9). However, whether complete closure is achieved cannot be determined from the Lx waveform.

During phonation in modal voice, the mucosal wave, travelling up the medial surfaces of the folds, is most clearly seen because of the phase difference in the closing of the upper and lower edges, sometimes referred to as the upper and lower 'lips' (Fig.5, p. 38).

Falsetto voice is not a voice quality likely to occur in irradiated male voices, the subject of the present study, but is described by Van Den Berg (1968) as achieved by high Adductive Tension, large Medial Compression and High Longitudinal Tension in the Vocal Ligaments but not in the Vocalis muscles (Fig. 4, p. 36).

Activity in the Vocalis muscle (Fig. 3, p. 34) ensures the voice is in modal register. As soon as the vocalis muscles stop tensing the vocal folds, the voice changes into falsetto. (Sundberg, 1987)

Hollien, Moore, Wendahl and Michel (1966) describe the production of 'creaky voice' also sometimes referred to as 'vocal fry':

1. The vocal folds when adducted are relatively thick and apparently compressed.

2. The ventricular folds are somewhat adducted.

3. The inferior surfaces of the false cords actually come into contact with the superior surfaces of the true vocal folds.

Hollien et al (1966) suggest that the result of this is that *'an unusually thick, compact (but not necessarily tense) structure is created prior to the initiation of phonation.'* They also conclude that vocal fold vibration during creaky voice is sustained at low subglottal pressure. This is later substantiated by findings by McGlone and Shipp (1971) and Monsen and Engebretson (1977).

The approximation of the ventricular folds and their contact with the upper surfaces of the vocal folds during creaky phonation, seems to achieve a damping of vocal fold vibration. This damping may contribute to the occurrence, in some types of vocal fry, of 'double pulse' phonation (Monsen and Engebretson, 1977) (Fig. 9). This pattern was first detected by high speed filming of the glottis by Moore and Von Leden (1958).

Hollien et al (1966) describe laryngeal tension characteristics of creaky voice, and emphasise that although it is produced with strong adductive tension and medial compression, longitudinal tension is not great. The vocal folds are slack and subglottal airpressure and airflow low.

A study by McGlone and Shipp (1971) compared muscle activity, air pressure and airflow measurements in vocal fry and low modal phonation. Their results showed that 'in modal phonation there was greater airflow, greater cricothyroid and interarytenoid muscle activity than in vocal fry.' There was no difference in subglottal airpressure or posterior cricoarytenoid muscle activity. However, airflow during vocal fry was significantly lower than in modal phonation. This is explained by the increased duration of the closed time and extremely short open time of the cycle and results in the increased vocal tract resistance observed during vocal fry (Fig. 9).

Smitheran and Hixon, (1981) estimated laryngeal airway resistance during vowel production. It was calculated as the ratio between the translaryngeal pressure and translaryngeal flow, measured during production of the syllable [pi]. It becomes evident why vocal tract resistance increases during fry phonation, as the extremely short open time, mentioned above, dramatically reduces the translaryngeal flow.

McGlone and Shipp (1971) conclude that the increased vocal tract resistance during fry phonation is achieved at least partly through increased thickness of the vocal folds as a result of unopposed Thyroarytenoid muscle contraction. In modal phonation there was consistently greater Cricothyroid and Interarytenoid muscle activity, which lead to an increase in the stiffness and a thinning, rather than thickening, of the vocal folds. This also resulted in firmer approximation of the vibrating folds (Fig. 3, p. 34).

In an earlier study Keidar (1983) showed that vocal fry was mainly perceived as a function of fundamental frequency. Irrespective of waveshape (synthetic stimuli were used), frequencies below 70 Hz were judged as vocal fry those above 70 Hz as 'other than fry'.

Abberton (1976) found, however, in her study of how laryngeal information affects speaker identification by listeners, that creaky voice quality resulted in a middle aged woman speaker being assigned to a category of 'low voice', despite her mean fundamental frequency being the same as a speaker in a 'high voice' category.

Keidar et al (1987) consider the three labels fry, chest and falsetto perceptually validated as *'legitimate terms for the three vocal registers lying in succession within the frequency range of the human voice.'*

The different phase and timing relationships in vocal fold vibration in different registers can, as shown be observed in Lx waveforms (Fig. 8 a and b, and Fig. 9) (Abberton, 1976, Abberton, Howard and Fourcin, 1989, Carlson, 1993b)

In perceptual voice quality rating systems that will be described in the next section, and particularly in the Vocal Profile Analysis system used in this study, judgements are required regarding perceived register of phonation or 'phonation types' (Laver, 1980). As indicated above, they have been extensively researched and found to have acoustic and physiological correlates.

ii Perceptual evaluation of voice quality.

The normal speaking voice can be described as produced most of the time in modal register in terms of its fundamental frequency and vocal fold vibratory characteristics. However, the components of the complex acoustic waveform, harmonics, formants and noise components, that radiates at the lips during speech and other vocalisations, are affected by glottic and vocal tract conditions. These may vary and be under the speakers' control, or habitual long term 'settings' of the mechanism typical for an individual speaker (Laver, 1980). Acoustic features may also be affected by organic or musculo-sceletal changes in the glottis or the vocal tract.

Changes in the shape, stiffness, length and mucosal condition of the vocal tract and glottis have audible effects on the quality of the voice. Listeners perceive this as more than 'pitch' which is the perceptual equivalent of fundamental frequency resulting from the vibration of the vocal folds. Kreiman, Gerratt and Precoda (1993) suggest: *'Voice quality is fundamentally perceptual in nature'.... what brings patients to seek help is that they do not sound 'normal'.* They judge success of treatment on whether they sound better.

Numerous studies have been carried out to compare 'normal' and 'abnormal' voice quality using trained (Stoicheff et al, 1983, Fritzell, Hammarberg, Gauffin et al, 1986, Södersten and Lindestad, 1990, Eskenazi, Childers and Hicks, 1990, Sederholm, McAllister, Sundberg and Dalkvist, 1992, Kreiman et al, 1993) or untrained listeners (Wynter, 1974, Bassich and Ludlow, 1986, de Leeuw, 1991).

Eskenazi et al (1990) reported that listeners found it easier to agree on degrees of abnormal voice quality parameters than aspects of 'normal' voice quality. Kreiman et al (1993) found, however, that inter rater agreement was best both for near normal and for extremely 'rough' voices. Raters varied widely, however, in their ratings in the middle of a scale.

To be able to judge 'abnormality' of voice quality , it may be assumed that we all share a common understanding of what is 'normal'.

Moore (1971) suggests, however, that there are many 'normal' voices based on speakers' sex and age. Listeners judge any deviation from normalcy on the basis of their own cultural, linguistic, educational and environmental background. Eskenazi et al (1990) propose one simple definition of 'normal' voice as a voice that shows no evident pathology (organic or functional) and no unusual voice characteristics or habits. From what follows it will be shown, that to perceptually judge voice quality is no 'simple' task.

Listeners, even voice professionals, differ considerably in how they perceive and describe aspects of voice quality other than pitch (Kreiman et al, 1993). A study by Hammarberg, Fritzell, Gauffin et al (1980) investigated the extent to which experienced voice clinicians agreed in their understanding of the terminology used to describe different voice qualities. 28 out of 50 of the most common terms used to describe voice quality were selected and used for rating twenty voices on a five-point scale for each of the 28 parameters. The resulting correlation matrix of the mean ratings resulted in the emergence of five bi-polar factors which accounted for 85.3% of the variance. The factors were:

1. unstable - steady
2. breathy - overtight
3. hyperfunctional - hypofunctional
4. coarse - light
5. head register - chest register

(Hammarberg et al 1980)

Such a system is more useful for describing and communicating voice quality characteristics than the original 50 terms. The purpose of using perceptual rating scales of voice quality is to improve communication between clinicians and to allow them to differentiate consistently between and within voice quality parameters to show for example improvement in voice quality after treatment.

After several further refinements of Hammarberg's et al (1980) perceptual classification system (Hammarberg and Fritzell, 1986, Hammarberg and Askenfelt, 1986, Hammarberg, 1986), the most recent classification (Hammarberg, 1986) proposed contains twelve variables. It was found that experienced voice clinicians achieve good agreement in rating voice quality when using this terminology.

Voice Quality Parameters:

1. Aphonic/intermittent aphonic
2. Breathy
3. Hyperfunctional/tense
4. Hypofunctional/lax
5. Vocal fry/creaky
6. Rough
7. Grating
8. Diplophonic
9. Voice Breaks
10. Instability
11. Register
12. Pitch

(Hammarberg, 1986)

Prerecorded tapes illustrating the above parameters are used in training speech and language therapy students to use a better defined terminology (Hammarberg et al, 1986).

Most studies of perception of voice quality seem to choose a more limited number of perceptual items for rating. Many use five or seven point equal appearing interval scales. An extensive review of scales used in the literature is given by Kreiman et al (1993).

A recent study by Sederholm et al (1992) using an adapted version of Hammarberg's terminology, showed that the term 'hoarseness' achieved the best agreement out of 15 'voice quality' terms and pitch judgements in a test of interrater agreement of dysphonia judgement in children. They concluded that the term is an 'uncomplicated' one and still useful in clinical practice. It is certainly the most commonly used and understood term by patients seeking treatment for voice disorders and one used in the present study for subjects to rate themselves.

In clinical diagnosis however, 'hoarseness' is not sufficient for determining treatment approach as it is only descriptive of a number of acoustic consequences of malfunctioning in the larynx or the vocal tract. Factor analysis in Sederholm's et al (1992) study demonstrated what particular auditory features constituted 'hoarseness'. These were 'hyperfunctional', 'breathy' and 'rough' and they comment that *"...different combinations of these properties occur as symptoms of different pathological states and call for different types of treatment".*

Kreiman, Gerratt and Berke (1992) also suggest that 'breathiness' and 'roughness' are *'related multidimensional constructs'*. Listeners have a tendency to vary in their attention to various aspects of each quality, which in their view may account for inter-rater unreliability in voice quality ratings.

Another attempt at structuring perceptual evaluation of voice quality is reported from Japan (Isshiki, Yanagihara, Morimoto, 1966, Isshiki, Okamura, Tanabe and Morimoto, 1969). Isshiki et al (1969) identified four factors that characterised hoarseness, rough (R), breathy (B), asthenic (A) and 'degree' (D). The scale was further developed and reported by Hirano (1981) as the "GRBAS"-scale. It uses five parameters:

Grade (G) represents the degree of hoarseness or voice abnormality,

Rough (R) refers to irregularity of vocal fold vibration i.e. perturbation of both fundamental frequency and amplitude of the glottal source sound.

Breathy (B) represents the degree of air leakage through the glottis and is related to turbulence.

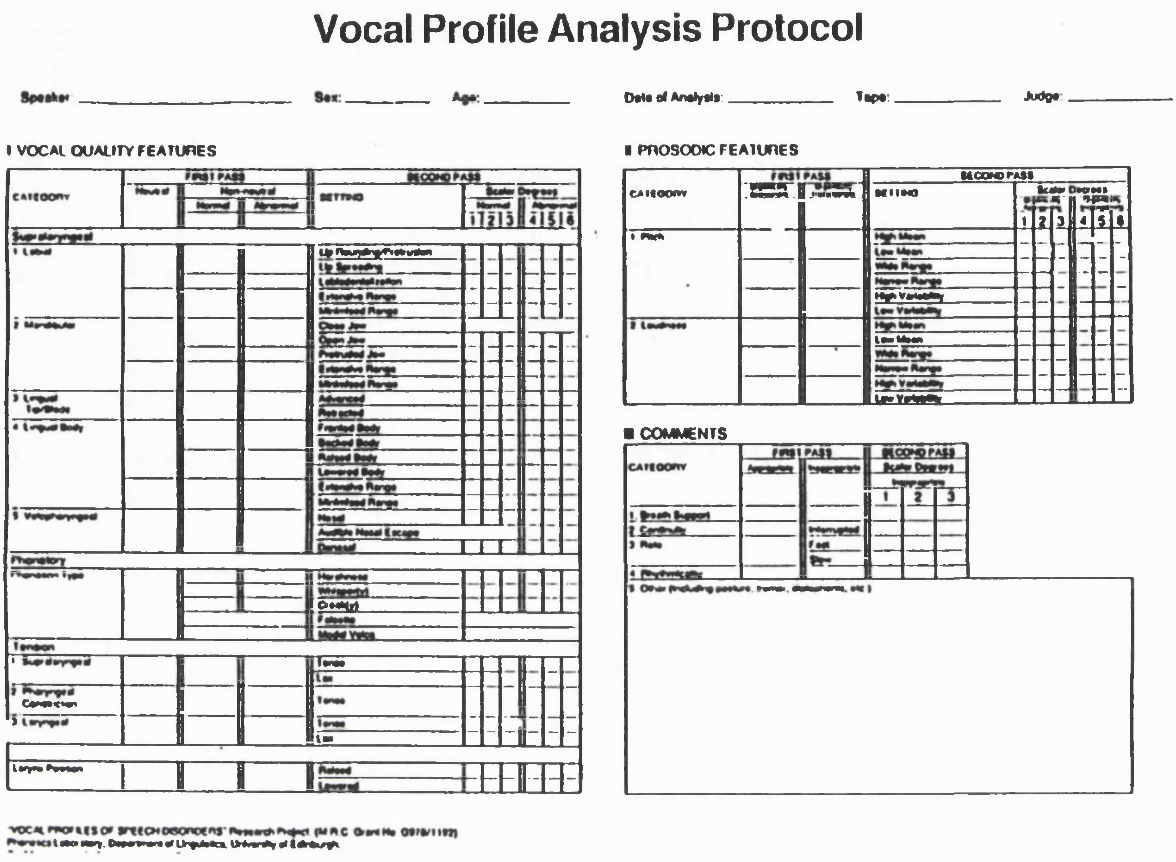
Asthenic (A) represents weakness or lack of power and is related to weak intensity of the glottal source sound and lack of higher harmonics.

Strained (S) refers to a hyperfunctional state of phonation and is related to abnormally high fundamental frequency, noise in the high frequency spectrum and/ or *'richness in high frequency harmonics.' (Hirano, 1981).*

Within each parameter judges are asked to rate voices using a four point scale, '0' denoting absence of the feature, '3' an extreme degree of the same. Training tapes have been produced but Hirano states that the method is in need of further improvement.

Sederholm et al (1992) discuss the problem of using numerical scales. They suggest that the number of points may be too small to allow experienced raters to discriminate between degrees of a given parameter, which would result in loss of potentially useful information. On the other hand, suggest Kreiman et al (1993), the use of a visual analogue scale, which Sederholm et al (1992) propose, may offer too fine a scale compared to judges' perceptual acuity in discriminating between degrees of voice quality features.

A problem with numerical scales may be that the rater's quantitative judgement becomes affected by a qualitative judgement which is by convention associated with some pathological state. An example of this is in the Vocal Profile Analysis system (VPA) (Laver, Wirz, MacKenzie and Hiller, 1981) where the convention is adopted of scalar points 1-3 denoting a feature being present but within 'normal' limits. Scalar points 4-6 indicating the presence of a feature to an 'abnormal' degree (Fig. 10). The authors are aware of the problem but suggest that the distinction between normal and abnormal, the boundary between 3/4 scalar degrees, should be used as an *'anchor for perceptual judgements by emphasising an approximate midpoint in the scale'*. They found that speech therapists trained in using the system showed better agreement about the 3/4 boundary than about any other boundaries between scalar degrees in the VPA.



Vocal Profile Analysis Protocol

Speaker: ______ Sex: ______ Age: ______ Date of Analysis: ______ Tape: ______ Judge: ______

I VOCAL QUALITY FEATURES

CATEGORY	FIRST PASS Neutral	FIRST PASS Non-neutral Normal	FIRST PASS Non-neutral Abnormal	SECOND PASS SETTING	Scalar Degrees Normal 1	2	3	Scalar Degrees Abnormal 4	5	6
Supralaryngeal										
1 Labial				Lip Rounding/Protrusion						
				Lip Spreading						
				Labiodentalization						
				Extensive Range						
				Minimized Range						
2 Mandibular				Close Jaw						
				Open Jaw						
				Protruded Jaw						
				Extensive Range						
				Minimized Range						
3 Lingual Tip/Blade				Advanced						
				Retracted						
4 Lingual Body				Fronted Body						
				Backed Body						
				Raised Body						
				Lowered Body						
				Extensive Range						
				Minimized Range						
5 Velopharyngeal				Nasal						
				Audible Nasal Escape						
				Denasal						
Phonatory										
Phonation Type				Harshness						
				Whisper(y)						
				Creak(y)						
				Falsetto						
				Modal Voice						
Tension										
1 Supralaryngeal				Tense						
				Lax						
2 Pharyngeal Constriction				Tense						
3 Laryngeal				Tense						
				Lax						
Larynx Position				Raised						
				Lowered						

II PROSODIC FEATURES

CATEGORY	FIRST PASS Appropriate	FIRST PASS Inappropriate	SECOND PASS SETTING	Scalar Degrees Appropriate 1	2	3	Scalar Degrees Inappropriate 4	5	6
1 Pitch			High Mean						
			Low Mean						
			Wide Range						
			Narrow Range						
			High Variability						
			Low Variability						
2 Loudness			High Mean						
			Low Mean						
			Wide Range						
			Narrow Range						
			High Variability						
			Low Variability						

III COMMENTS

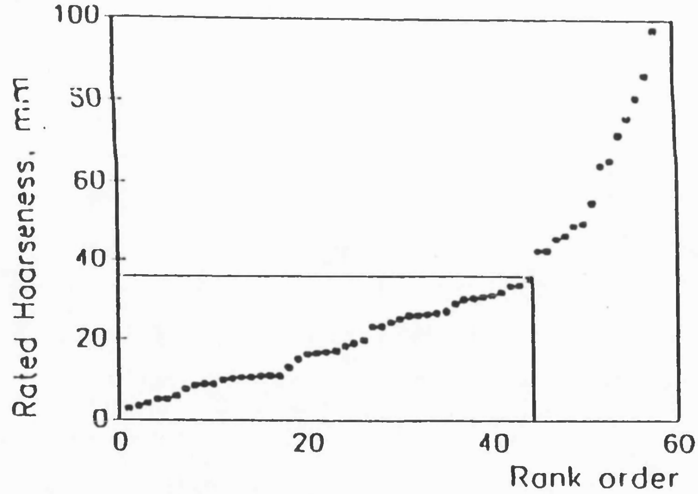
CATEGORY	FIRST PASS Appropriate	FIRST PASS Inappropriate	SECOND PASS Scalar Degrees Inappropriate 1	2	3
1. Breath Support					
2 Continuity		Interrupted			
3 Rate		Fast			
		Slow			
4 Rhythmicality					

5 Other (including posture, tremor, dislocations, etc.)

"VOCAL PROFILES OF SPEECH DISORDERS" Research Project (M.R.C. Grant No. G978/1192)
Phonetics Laboratory, Department of Linguistics, University of Edinburgh.

Vocal Profile Analysis (VPA) Assessment form.

Figure 10



Rank order plot of 58 voices according to degree of 'Hoarseness' rated on a visual analogue scale in millimeters (mm). Note the marked 'discontinuity' at 35 mm.

(From Sederholm et al, 1992)

Figure 11

Sederholm et al (1992) try to overcome these problems by using Visual Analogue Scales (Wewers and Lowe, 1990), where each voice quality parameter is rated along a 100 mm continuous line, the extremes of which correspond to absence versus 'extreme degree' of a feature. Pitch was rated along a 200 mm line as values both below (low pitch) and above (high pitch) a 'Neutral' point would be expected. The ratings (in mm) of seven experts were averaged for each of 58 voices and each of 14 parameters and the mean values plotted in rank order for each voice and for each parameter.

It was found that for some parameters this graph showed a distinct 'breakpoint' or 'knee', which could be used to separate 'normal' from 'atypical' in a particular population (Fig. 11). Sederholm et al (1992) suggest that this discontinuity indicates *"a demarcation line between different systems."* As opposed to the 'normal' versus 'abnormal', 3/4 scalar point boundary in the VPA, the 'discontinuity' in some graphs (indicating 'normal' vs 'atypical') described by Sederholm et al (1992) are at different points along the 100 mm scale depending on the feature judged.

Wynter (1974) asked 25 judges, some trained some untrained to rate subjects' voice quality on a four point scale, '0' denoting 'normal'. The parameters chosen were the degree of nasality, pitch deviation (appropriate pitch for age and sex), hoarseness, breathiness and stridency.

The results showed reasonable agreement between raters of pitch deviation, hoarseness and stridency. Agreement was less for breathiness and very poor for the degree of hyper- and hyponasality.

Stoicheff et al (1983) used 8 experienced judges to rate the degree of 'dysphonia' and the predominant voice quality during reading of the first paragraph of the 'Rainbow Passage'. Subjects were a group of 'normal' speakers and a group of patients before, and one year after radiotherapy for laryngeal carcinoma. They used a seven point equal appearing interval scale to rate voices along the 'dysphonia' dimension, and also whether voices sounded 'Rough', 'Breathy', 'Hoarse', 'Strained' and 'Normal'.

Results showed statistically significant differences in the degree of dysphonia between normal speakers and patients both before and after radiotherapy. The type of dysphonia in patients changed from being predominantly 'strained', 'hoarse' and 'breathy' before radiotherapy to being 'hoarse' and 'rough' one year later. It seemed the radiotherapy treatment reduced the perceived 'Breathy' and 'Strained' quality of the pre treatment voices. Normal, i.e non-irradiated, voices were mostly described as being 'normal' with a degree of 'rough' quality.

iii Acoustic correlates of 'hoarseness'.

As shown above there is a large body of data using perceptual descriptive terms for evaluation of voice quality. In an attempt to reduce terminological ambiguity and confusion and help communication about 'voice quality', much research has also been devoted to finding objective acoustic correlates of perceptual voice quality parameters.

These studies tend to use sustained vowel phonations as it has been found that acoustic bases of voice quality are found primarily in the vowel spectrum (Fairbanks, 1960). A major reason must however also be that early instrumentation did not allow analysis of longer stretches of speech. Recent computer based instrumentation has enabled researchers and clinicians to record and analyse considerably longer samples of speech and reading approximating more 'natural' habitual voice production (Hammarberg et al, 1980, 1984, 1986 a,b, Hammarberg, 1981, 1986, Ohlsson et al, 1987 a, b, c, 1988, Ohlsson, 1988, Barry et al, 1990 a,b, 1991, Carlson, 1993 a, b).

However, an important, early, acoustic study of hoarseness by Yanagihara (1967) classified degrees of 'hoarseness' in four different groups according to the effect on the narrow band analysis of spectra of sustained vowel sounds:

> Type 1 'The regular harmonic components are mixed with the noise component chiefly in the formant region of the vowels.'
>
> Type 2 'The noise components in the second formants of [ɛ] and [i] dominate over the harmonic components and slight additional noise components appear in the high frequency region above 3000 Hz in the vowels [ɛ] and [i].'
>
> Type 3 'The second formants of [ɛ] and [i] are totally replaced by noise components, and the additional noise components above 3000 Hz further intensify their energy and expand their range'.
>
> Type 4 'The second formants of [ɑ], [ɛ] and [i] are replaced by noise components, and even the first formants of all the vowels often lose their periodic components which are [supplanted] by noise components. In addition, more intensified high frequency additional noise components are seen.'

Yanagihara (1967) found that the noise element in the spectra was more evident in the vowels [ɑ], [ɛ] and [i] than in [u] and [ɔ]. He concluded that the noise components may be the result of turbulent airflow through an incomplete closure of the glottis during phonation, or due to irregular vibrations of the vocal folds: *'the quality of hoarse voice is not entirely dependent upon the noise components and changes in harmonic structure... [but also to] the aperiodicity of the fundamental frequency.'* He also found that Spectrographic analysis added useful objective information that correlated well with subjective perceptual evaluations of hoarse voice.

Several more recent studies of hoarseness using spectrography, attempt to specify and quantify the relationship between spectral noise and harmonic components in perceived hoarse voices. They all confirm that a characteristic feature of hoarseness is the replacement of harmonics by noise energy in the spectrum (Isshiki et al,1969, Murry, Singh and Sargent, 1977, Schaedel, 1979, Kim, Kakita and Hirano, 1982, Yumoto, Gould and Baer, 1982, Klich, 1982, Yumoto, Sasaki and Okamura, 1984, Bradley, 1985, Wolfe and Steinfatt, 1987).

Sasaki et al (1991) used a digital sound spectrograph, which offers great flexibility in the choice of band filters and frequency ranges for analysis purposes. Sustained [ɑ] spoken by 34 normal speakers and 54 with laryngeal pathology was recorded and stored digitally. They calculated a) the area under the line connecting the peaks of the fundamental frequency and the harmonic frequencies, which they termed the 'total acoustic energy' (V) and b) the area under the line connecting the troughs of the harmonic components or the 'noise energy' (N). The ratio N/V correlated significantly with severity of hoarseness as judged by experienced listneres on a 4 point scale.

Breathy voice quality has attracted particular attention. It gives rise to a characteristic loss of high frequency components in the spectrum as a result of the increased open phase of the glottal cycle. Klich (1982) found that energy in all three spectral frequency ranges 100-500 Hz, 1.5-2.5 kHz and 3.5-4.5 kHz was related to listener ratings of breathiness. Inspection of narrowband spectrograms showed that as breathiness increased the number of discernible harmonics decreased. Fritzell et al (1986) used inverse filtering of the glottal waveform and found good correlation between the maximum/minimum airflow quotients and ratings of breathiness. Their results indicate a quantitative relationship beteen breathy voice quality and glottal insufficiency.

Further extensive research has been carried out in attempts to relate perceptual terms of voice quality to their acoustic correlates by Murry, Singh and Sargent (1977), Hammarberg et al (1986), Hammarberg and Askenfelt (1986), Hammarberg (1986).

Hammarberg et al (1986 a) used trained voice clinicians to rate the voices of ten voice patients before and after treatment, six patients who received no treatment and six 'normal' subjects. The subjects were recorded reading a passage as the assumption was that *'vocal function is more realistically reflected in continuous speech than in sustained vowels'*. Acoustic analyses, Longtime average spectrum analysis (LTAS), fundamental frequency distribution analysis (FFDA) and frequency perturbation analysis were applied to the recorded reading samples and factor analysis to the perceptual ratings for each of 25 voice quality parameters on a 5 point scale.

The factor analysis of the pathological speakers' voice ratings revealed five bipolar factors. Three of these correlated significantly with acoustic data extracted from LTAS, FFDA and frequency perturbation. A large standard deviation of the fundamental frequency and high degree of frequency perturbation corresponded to 'Unstable voice quality' including perceptual factors such as diplophonia and voice breaks. Frequency perturbation was found to decrease significantly after voice therapy. Typical LTAS profiles were obtained for voices described as 'steady/sonorous', 'breathy/hyperfunctional' and 'breathy/hypofunctional'.

Wolfe and Steinfatt (1987) explored acoustic correlates of vocal severity among 51 patients with a variety of voice disorders. Sustained vowels [ɑ] and [i] were recorded and judged on a 7 point scale according to:

a) overall severity
b) whether they were perceived as hoarse, breathy, strained or normal.

Spectrographic analysis was carried out on the vowel samples. Yanagihara's (1967) categories, described above, were employed for visual analysis of the spectrograms with one added category to enable judging of a pattern without spectral noise. Period measurements were also carried out and fundamental frequency and jitter measurements were among the acoustic parameters calculated for the voice samples.

Overall, spectrographic noise showed the highest single correlation with severity of disorder. Adding the 'natural log of the period standard deviation' (LNPSD) improved the predictability. For the 4 individual perceptual 'voice types' identified, the acoustic predictors of severity varied. For instance in 'breathy' voice the LNPSD was the best predictor of severity. For the 'hoarse' voice type, spectrographic noise was the only significant predictor.

The explanation given by Wolfe and Steinfatt (1987) of the spectrographic difference in predictability between 'breathy' and 'hoarse' voice, was that laryngeal irregularities contribute to produce the turbulent airflow in the hoarse voice. This gives rise to widespread and strong noise components in the spectrum as well as loss of harmonic structure. The relative lack of turbulence in the breathy voice only reduces the harmonic structure and does not necessarily give rise to noise.

Takahashi and Koike (1975) offer support for the notion that hoarseness is the result of more complex laryngeal activity than breathiness. They found that 'roughness' correlated both with amplitude and frequency perturbation quotients, breathiness only with amplitude perturbation.

The third 'voice type' in Wolfe and Steinfatt's study was 'strained' voice, where the best severity predictor was the 'period coefficient of variation' (PCV) described as the Period standard deviation divided by mean period. Prediction improved when LNPSD, signal to noise ratio (SN) and fundamental frequency were added. Finally the 'unclassified' voice type severity was best predicted by SN. Prediction improved when LNPSD was added.

Wolfe and Steinfatt (1987) suggest that their findings show that different predictor variables are needed to determine severity of different voice qualities and they conclude their exhaustive investigation.... *"Current knowledge suggests that, like perceptual correlates of voice disorders, acoustic predictors of voice types appear to be multidimensional."*

iv The Vocal Profile Analysis Scheme (VPA).

The 'Vocal Profile Analysis Scheme', the 'VPA' (Laver, Wirz, McKenzie and Hiller, 1981, 1986) has become widely known in Britain and used for perceptual evaluation of voice quality, as the number of trained users has increased. It offers a formalised, and extremely detailed, system for evaluation of phonatory and articulatory behaviour. It takes into account both laryngeal and supralaryngeal contributions to overall perceived voice quality and the interaction between laryngeal and articulatory 'settings' (Laver, 1980).

According to Laver and Hanson (1981) *"...virtually no activity can take place anywhere in the (vocal) apparatus without repercussions in varying degrees elsewhere"*.

The most recent version of the VPA (Fig. 10, p. 83) gives a profile of an individual's articulatory and phonatory characteristics described in a mixture of assumed physiological and auditory perceptual terms. Laver chooses to define 'voice quality' very broadly and not only in terms of the result of laryngeal activity, which is the case in other voice quality rating systems. The VPA user is invited to rate 'Supralaryngeal settings'; labial, mandibular, lingual and velopharyngeal; 'Phonatory settings'; Harshness, Whisper, Creak, Modal and Falsetto voice (see comments on 'registers' pp. 71-77); 'Tension settings' Supralaryngeal, Pharyngeal and Laryngeal tension or laxity; 'Larynx position' - Raised or lowered. 'Prosodic features': pitch level, range and variability, and 'Loudness'. Comments are invited on 'Breath support', 'Continuity', 'Rate' and 'Rythmicality' and the presence of e.g. diplophonia, tremor and other observations.

Supralaryngeal, Phonatory and Tension 'Settings' are defined in terms of their relationship to a postulated 'Neutral' reference setting producing the acoustic characteristics of an adult male vocal tract, 17 cm in length and of approximately equal crossection along its full length and with a closed velopharyngeal port (Laver, 1980, p. 14).

A 'Setting' is described as *'...a tendency for the vocal apparatus to be subjected to a particular longterm muscular adjustment'*. An example of a 'setting' would be a tendency for an individual to speak with 'non-neutral' amount of e.g.'lip-rounding', which may be the result of both his individual anatomy and to '*the use to which he puts his vocal apparatus'*. In this case it would result in a lengthening of the vocal tract and the acoustic effect of lowering all formant frequencies (Fant, 1960). All settings are described in Laver (1980) in terms of muscle systems involved and in terms of the acoustic effect they rise to.

Voice qualities are described as products of composite settings, which interact in various ways. The anatomical coupling between the resonators and the vibratory voice source means that different supralaryngeal settings may be associated with different patterns of laryngeal vibration.

'Compatibility' refers to the idea that some settings are physiologically 'incompatible' e.g. falsetto voice with 'neutral' voice, because they are produced by different amounts of longitudinal and adductive tension and medial compression resulting in different vocal fold contact patterns. Others are 'compatible' in the sense that *'the action of one setting can modify another'* e.g. 'Whispery' voice and 'Modal' voice can co-exist in that an increase in medial compression to produce 'Whispery' voice quality will not neccessarily change the mode of vocal fold vibration.

Laver also accounts for the fact that certain vowels and consonants differ in how sensitive they are to the influence of different 'settings' by what he calls segmental 'susceptibility'. This would include the minimal physiological and acoustic effect of e.g 'velarization' on a close back vowel as opposed to the maximal effect on an open front one. *Intermediate vowels will be affected by the setting in a hierarchy of susceptibility'*.

When using the VPA, the trained judge in the 'First Pass' listening task decides whether the perceived Supralaryngeal, Phonatory and

Tension settings are within 'Neutral', 'Normal' or 'Abnormal' limits. These are defined in Laver and Hanson (1981) and Laver et al (1981). 'Neutral' settings are described as 'standard reference settings' produced by the postulated adult male vocal tract described earlier. If aspects of a voice are deemed 'Non-neutral', the 'Second Pass' listening requires the judge to rate the 'degrees' of non-neutral settings, whether they are within 'Normal' or 'Abnormal' limits.

Laver et al (1981) are aware of the difficulties involved in the Normal/Abnormal decision and tentatively suggest as a rough rule of thumb that settings in the 'abnormal' range are those that are deemed to require treatment. They admit this does not stand up to scrutiny, however, as *'Even when the vocal profile is taken into account alongside other factors such as diagnosis of pathology, the patient's own assessment of voice etc. it is seldom single settings, but rather...constellations of settings which cause the vocal profile as a whole to indicate the need for treatment..'*. They suggest these settings may well be within 'normal' limits but in unusual combinations.

Laver et al (1981) also point out regarding the acoustic consequences of postulated physiological settings that *'perceptually equivalent qualitites may be produced by physiologically different mechanisms and in pathological speech the auditory quality-physiology relationship may become very unclear. The continuum implicit in the (VPA) form is therefore an auditory one only'.*

In the VPA qualitative judgements are made within 'normal' and 'abnormal' limits and are given numerical scalar degrees 1-3 and 4-6 respectively (Fig. 10, p. 83). This is a somewhat artificial association between quantitative scale points and qualitative value judgements and imposes restrictions on judges which may distort the scale particularly for judgements of supralaryngeal settings. There is a great range of 'normal' regional variations in articulatory settings. Some lingual settings would be within 'normal' limits in e.g. Scotland and Ireland but considered 'abnormal' in the south of

England only because they do not belong in a Southern British accent. Definitions of what is 'normal' and 'abnormal' will consequently vary and norms would need to be developed for different regional varieties of the same language and for different languages. The 'normal'/'abnormal' divide is controversial and should be redefined or removed.

Laver (1980) also refers to studies which have demonstrated speakers using sociolinguistically determined typical vocal tract 'settings' e.g. working class speakers in Norwich used different supralaryngeal and phonatory settings and pitch characteristics compared to middle class speakers (Trudgill, 1974). Esling (1978) showed that in Edinburgh higher social status correlated with the use of creaky phonation; lower status with greater whisperiness and harhshness. This seems further evidence for the inadvisability of applying the 'normal'/'abnormal' dichotomy.

v Phonation types

In this study we choose, however, to use the VPA to judge 'voice quality' in a 'narrow' sense; only rating the perceptual effect of laryngeal events described in Laver's terms as 'Phonation types' and 'Laryngeal Tension' settings.

Laver (1980) uses only a small number of terms to describe voice quality: Modal, Falsetto, Whisper, Creak, Harshness and Breathiness. He suggests that some of these can combine to form 'Compound Phonation types'. The current VPA offers five 'Phonation types', Harshness, Whisper and Creak and a judgement of Modal or Falsetto voice. All, except 'Harshness' can appear as 'simple phonation types' on their own. Modal and Falsetto cannot combine, suggests Laver, in this case as a result of 'incompatibility' of mode of vocal fold vibration. They can, however, combine with all or some of the first three giving rise to for instance 'Whispery (modal) voice', Whispery Falsetto', 'Creaky (modal) voice', 'Creaky falsetto', 'Harsh, whispery, falsetto', 'Whispery, creaky, voice' etc.

The origins of the VPA can be found in Laver's (1980) 'The Phonetic Description of Voice Quality'. As it is the perceptual evaluation system used in this study, this warrants some explanation of the terminology used.

Laver describes the 'Neutral' mode of phonation as being *'regularly periodic, efficient in producing vibration without audible friction'*. This description corresponds closely to Hollien's (1971) 'Modal voice' which, however, Hollien also defines in terms of fundamental frequency range. Laver suggests:

> *Modal voice has a moderate degree of medial compression while every other type , falsetto, whisper, creak, and harshness, has a high or very high degree. Also, modal voice has only moderate adductive tension, where falsetto, creak and harshness have high or very high degrees. It is true that whisper is similar to breathiness in having low adductive tension, but whisperiness and breathiness have been defined here as complementary actions of the same scale, so that their potential combination is excluded by definition.*

In the first version of the VPA (1981) Laver distinguished between 'Whispery' and 'Breathy' phonation types. Although he recognised there was a close auditory relationship between them, and they could be considered auditorily at different points of a continuum, he suggested the physiological differences in how they were produced warranted a differentiation between them.

In 'Whispery' voice, phonation would be produced with a greater degree of laryngeal effort, more specifically, with increased Medial Compression (Fig. 4, p. 36), to achieve a greater amount of glottal friction, which would be more prominent than in 'Breathy' voice. Breathy voice would be produced with minimal adductive tension, extremely weak medial compression and low longitudinal tension (Laver,1980). The modal, periodic component of the voice would be more prominent in breathy voice than in whispery voice, which would show either equal degrees of audible friction and modal voice, or more friction.

Whereas we would be able to perceive 'Compound Phonation Types' such as 'Whispery voice', 'Whispery Creak', And 'Whispery falsetto' because they are all produced with high degrees of medial compression. 'Breathy' voice could not combine with any other phonation type except modal voice, i.e. not with falsetto nor 'creak' as it is produced with very low medial compression (Laver, 1980).

This difference in definition between 'whispery' and 'breathy' voice quality may be illustrative of the physiological correlates of the commonly used gross dichotomy in describing abberrant voice quality as 'Hyperfunctional' or 'Hypofunctional' (Hammarberg, 1986, Hillman et al, 1989, Rammage et al, 1992). 'Whispery' voice would be associated with a degree of effort (Hyperfunction) with increased medial compression giving rise to a posterior glottic chink and audible turbulence. 'Breathy' voice would be associated with low degrees of adductory tension and medial compression (Hypofunction) with or without a posterior glottic chink.

The most common physiological correlate of 'breathy' voice is a persistent posterior glottal chink between the arytenoids. This has been found to be common especially in women, without being perceptually abnormal. (Södersten and Lindestad, 1990, Södersten et al, 1994, Biever and Bless, 1989, Rammage et al, 1992)

'...several sources of glottic and supraglottic factors influence perceptual voice characteristics, including those of 'breathiness'. The relative contribution of turbulent airflow through a posterior glottic chink may be small.' (Rammage et al, 1992).

In the current version of the 'Vocal Profile Analysis Protocol' (Edinburgh, 1986), (Fig. 10, p. 83), the breathy and whispery phonation types have, however, been merged and only appear as 'Whispery'. 'Whispery' and 'Breathy' are now considered on a continuum of medial compression, whisper having high, breathy voice low, but both having low adductive tension giving rise to degrees of posterior glottic chinks.

Using one term to describe breathy phonation seems to make sense. The rating of degrees of laryngeal tension or laxity in the VPA, help differentiate ('Hyperfunctional' and 'Hypofunctional') voices with degrees of glottal air leakage as described above. Whether 'Whispery' is the term that best describes this feature, and is most widely understood, is debatable. It highlights the need however, for objective measurements to complement perceptual evaluation and reduce the number of impressionistic labels used.

Laver (1980) describes Harsh voice as a variety of Modal voice. The pitch may be normal (Michel, 1964) or close to the bottom of the modal range and there is a greater amount of aperiodicity of vocal fold vibrations (Wendahl, 1963, Zemlin, 1964, Michel, 1964). A physiological correlate of harshness is high laryngeal tension level (Laver, 1980, Milisen, 1957, Van Riper and Irwin, 1958, Kaplan, 1960 and Zemlin, 1964). Harsh voice is produced with greater medial compression and adductive tension than modal voice.

'Creak' has previously been described in detail in the section about 'Vocal Registers'. Laver (1980) recognises the existence of registers but due to ambiguity in the interpretation and understanding of 'register', he prefers to describe falsetto, modal and creaky phonation as 'Phonation types'. He avoids defining phonation types in terms of fundamental frequency because of the previously suggested overlap between the extremes of the low falsetto/high modal voice range and the low modal/high creaky voice range.

Creak or 'vocal fry' tends to be defined in terms of low fundamental frequency. Laver, however, suggests that 'creak' can also be perceived in a falsetto voice and it is a common feature in some speakers of American English and Italian, where it is interspersed throughout the vocal range and used more or less continuously as part of habitual voice production. This was also found by Abberton (1976). This would be described as 'Creaky (modal) voice' as opposed to 'Creak', often used in English at the end of a falling intonation contour to signal the end of a statement.

'Tension settings' are described as *'settings of overall tension, which exercise their effect throughout the vocal system'*. Two broad categories are defined, 'Tense voice' and 'Lax voice'.

Modification in degrees of laryngeal tension in the form of Medial Compression, Adductive and/or Longitudinal tension, (Fig. 4, p. 36) will result in the emergence of perceptions of harshness and/or whisperiness (Laver, 1980).

Laver goes into a great deal of detail describing the effect of tension in the laryngeal, pharyngeal and supralaryngeal vocal tract and describes, for instance, 'tense voice':

> *A tense voice will tend to...ligamental, harsh or ventricular phonation which will sound comparatively louder and higher pitched; higher subglottal air pressure; slightly raised larynx; constriction of the upper larynx and lower pharynx, and possibly of the faucal pillars; a tensed velum; vigorous and extensive radial movements of the convex - surfaced tongue in segmental articulation; vigorous activity of the lips; and a more mobile jaw.*

A 'lax voice' will tend to the opposite and, particularly considering laryngeal and pharyngeal tension, the lax voice will demonstrate...

> *...breathy or whispery phonation which will sound comparatively softer and lower pitched; lower sub-glottal air pressure; a slightly lowered larynx and unconstricted pharynx.....*

The main acoustic effect of different tension levels in the larynx and vocal tract, is on the relative amount of harmonic energy in the upper part of the spectrum. A low level of tension results in absorption of acoustic energy in the vocal tract and low energy levels in the spectrum; a tense voice would result in strong upper partials.

The present study will report VPA ratings of 'Phonatory features' and 'Laryngeal Tension' characteristics (Fig. 10, p. 83).

vi <u>Reliability of perceptual voice ratings.</u>

Auditory perceptual scales for evaluation of voice quality will always be subjective, however stringent have been the attempts at defining terms and scalar points, and however thorough the training of users. The more variables that are introduced the greater the number of potential disagreements between judges and, it may be expected, the lower the reliability of the system.

Bassich and Ludlow (1986) found poor degrees of agreement between their untrained judges asked to rate voices along 13 dimensions. Both within and between judges there was poor agreement. They concluded that the task of rating voices perceptually is difficult and requires substantial training

The 'GRBAS' system (Hirano, 1981) looks immediately attractive and easy to use. With such relatively few aspects of voice quality to rate on a four point scale, inter-rater reliability may be assumed high. The drawback may be that it does not offer enough descriptive labels and/or scalar points.

On the other hand the VPA invites assessment of a large number of variables, and inter-rater reliability may be low or at least vary, depending on the parameter being judged. An attempt at establishing reliability is referred to in Laver, Wirz, McKenzie and Hiller (1981).

The degree of agreement among trained judges in assessing six test-voices using the VPA was measured. Error scores (defined in Laver et al, 1981) for each of the 21 parameters are given for two groups of judges (Laver et al, 1981). One group trained 'intensively' over 2½ days, the other group was trained once a week in 1½ hour sessions for 8 weeks.

An analysis of Average Error Scores shows that the majority of parameters (11) reached a satisfactory degree of agreement between judges in the 'intensive' group. Seven parameters needed some further training and three substantial further training. The figures for the eight week training program show seven parameters reaching acceptable agreement, eight needing some, four considerable further training.

Four parameters, Supralaryngeal Tension, Larynx Position, Fronted-Backed Tongue Body and Raised-Lowered Tongue Body are the most difficult for judges to agree on. They all achieve the highest error-scores in both groups.

A reflection by this VPA user after training and practice on the VPA is that it is difficult to keep in mind perceptual norms for scalar points for such a great number of variables. Regular refresher courses are needed to 're-calibrate' and ensure agreement between judges. Maybe this difficulty is reflected in the group of judges above trained over eight weeks. They have somewhat higher average Error Scores (mean 1.28, median 1.58) and need further training on more variables than the 'intensive' group (Mean Average Error Score 1.03, Median 0.98).

Hammarberg (1986) investigated the test-re-test reliability for the minimum number of perceptual variables necessary for the description of aberrant voice quality established through factor analysis. Experienced voice clinicians were judges. *"Correlation coefficients of r = 0.82 - 0.97 were obtained for repeated ratings for eight (of the twelve) parameters."* All voice qualities except "grating" achieved a satifactory inter-rater reliability coefficient. Hammarberg et al (1986) in a similar study found good inter- rater reliability and conclude *'The results indicate that perceptual evaluation by clinically well trained listeners is reliable and reproducible and can be used for systematic evaluation purposes if handled with caution.'*

Wolfe and Steinfatt (1987) found individual test-retest agreement for their eight listeners between 83-95%; overall it was 89%. Interjudge agreement was also adequate.

Klich (1982) found high reliability among listeners rating breathiness. Rammage et al (1992), however, found poor agreement between experienced judges in their judgement of breathiness and suggest this may have been due to poor agreement on the perceptual and operational definition of breathiness. However judges agreed closely in their ratings of degree of overall dysphonia severity.

Sederholm et al (1992) found satisfactory inter-rater reliability for 14 out of 15 parameters judged. The highest one was for for 'hoarseness' which was .92, the lowest for 'voice breaks', 0.36.

Kreiman et al (1993) in an extensive review of 57 randomly drawn studies reporting perceptual evaluation of voice quality conclude, however, that most of them suffer from one or more methodological problems. They cite: '*...failure to evaluate listener reliabiltiy, inadequate samples of raters, too few repeated trials for intra-rater reliability estimates, failure to report confidence intervals and use of inappropriate statistics*'.

They further suggest that '*no clear relationships between methods and reliability/agreement has emerged and no model of voice perception or for the rating process has been developed.*'

In the present study we will confirm some of the difficulties encountered in other perceptual studies referred to by Kreiman et al (1993).

CHAPTER IV - DIRECT AND INDIRECT EXAMINATION OF LARYNGEAL FUNCTION.

1 Observation of vocal fold function.

In cases of emerging laryngeal pathology, an early diagnosis of the cause of vocal symptoms is important. The larynx, due to its prime function as a protector of the airway, and its anatomical position, is a notoriously difficult structure to examine and visualize without causing some degree of distress and discomfort. In the context of examination of vocal dynamics, malfunction and remediation, this is a severe limitation.

Visualisation of the larynx and the vibrating vocal folds may be achieved through Indirect Laryngoscopy (IDL). In recent years, the increased use of Nasopharyngoscopy, where a flexible fibreoptic bundle is passed via the nose, behind the soft palate and suspended above the laryngeal vestibule, allows direct observation of the vocal folds during a greater variety of phonatory conditions with reduced discomfort for the patient. Like IDL, it can be combined with Videorecording or Videostroboscopy for later viewing and comparison with previous recordings. A rigid fibrescope, introduced orally is also commonly employed and gives a superior, enlarged and more detailed view of the structure and function of the vocal folds.

Direct Laryngoscopy (DL), or Microlaryngoscopy with the patient under general anaesthesia, may be undertaken to allow close examination of laryngeal structures under a microscope. This allows observation of morphological changes and biopsies to be taken of laryngeal tissues and removal of benign or malignant growths. This is always performed on patients with mass lesions of the vocal folds and related structures.

All such methods are inevitably invasive. They require the insertion of instruments via the oral and/or nasal cavities which may lead to atypical modes of articulation and phonation due to interference with the tongue, lips and soft palate. The close proximity and muscular relationship of these structures with the larynx, may give rise to atypical laryngeal gestures during phonation (Södersten and Lindestad, 1992). Besides, except in flexible nasopharayngoscopy, subjects can only sustain phonation on a single vowel, which will not necessarily give a true picture of habitual phonation.

The above mentioned methods are essential for diagnosis of disease, but can not be used for monitoring or rehabilitation of vocal behaviour.

Various other techniques have been employed for direct observation of the folds during phonation such as ultra- high speed filming (Childers and Larar, 1984, Childers, Smith and Moore, 1984, Childers, Hicks, Moore and Alsaka, 1986), ultra-short pulse X-radiography (Noscoe, Fourcin, Brown and Berry, 1983) and Laryngostroboscopy (Fourcin, 1974, Lecluse, 1975, Fog-Pedersen, 1977, Hirano, 1981, Anastaplo and Karnell, 1988).

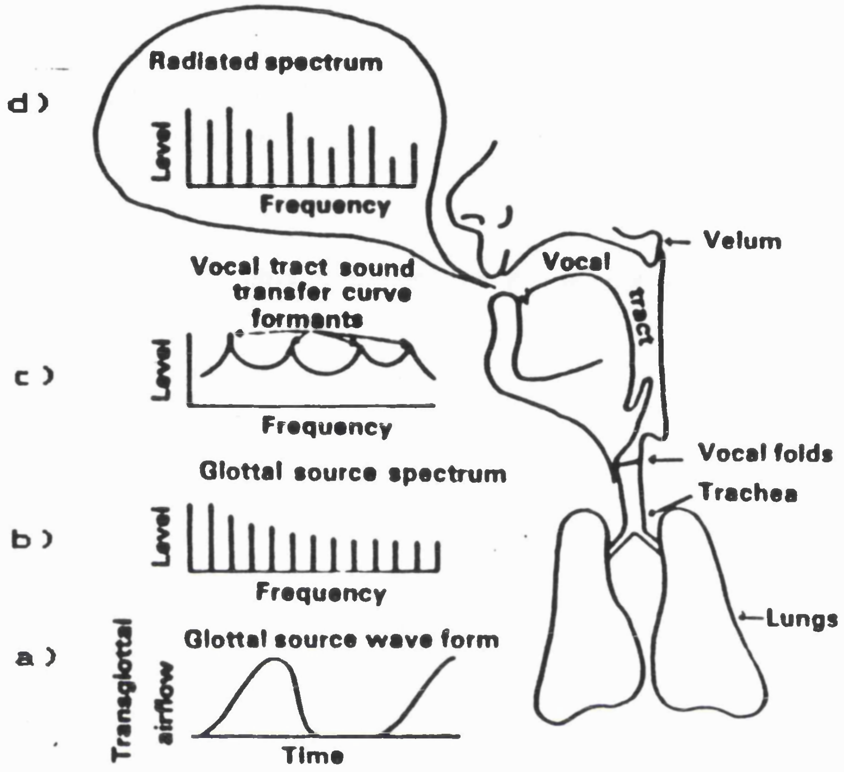
Some of these studies will be described in detail below. Such techniques have helped further our understanding of the complexity of laryngeal aerodynamics during phonation. They have been used to record a variety of vocal activities and allowed observation of subtle adjustments in laryngeal and supralaryngeal structures, that affect the mode of vibration of the vocal folds.

However, methods exist for noninvasive, indirect examination of vocal function by analysis of the complex acoustic waveform that radiates at the lips during speech. Some of these methods will be described below.

A ACOUSTIC MEASUREMENT TECHNIQUES.

ii Spectrography.

Spectrography is totally noninvasive and capable of analysing many aspects of speech and voice as it is based on acoustic recordings of speech.



Schematic illustration of the generation of voice.
a) Vocal fold vibration results from a sequence of air flow pulses
b) These give rise to a spectrum of harmonic overtones, the amplitudes of which decrease monotonically with frequency.
c) The spectrum is filtered according to the sound transfer characteristics of the vocal tract with its peaks, the formants and the valleys between them.
d) In the spectrum radiated from the lip opening the formants are depicted as peaks, because the partials closest to a formant frequency reach higher amplitudes than the neighbouring partials. (From: Sundberg, 1987)

Figure 12

In Fig. 12 the relationship is illustrated between a) the voice source, which constitutes the pulses of air that pass between the vocal folds in synchrony with their frequency of vibration, b) the resulting spectrum of harmonic overtones, filtered and modified in terms of their amplitudes as they pass up the vocal tract giving rise to c) formants at different frequencies, according to which sound is produced, and ultimately d) the spectrum of this sound radiating at the lips (Sundberg, 1987).

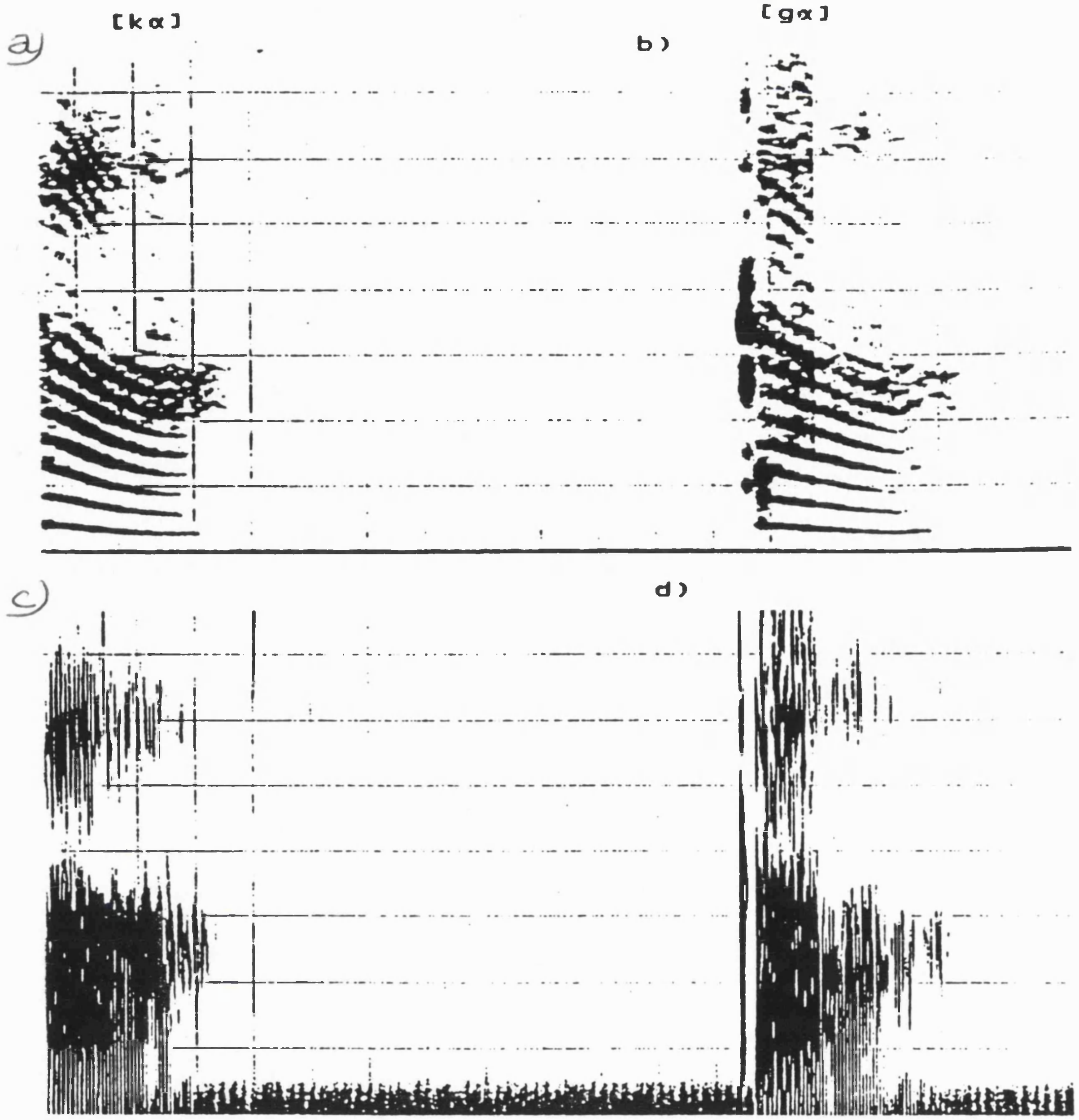
The radiated spectrum (d) illustrates the frequencies at which there are harmonics and how strong these harmonics are. They are strongest in the vicinity of the formants. The complex acoustic waveform radiating at the lips, is the result of the summation of the harmonics and their respective amplitudes (Fig. 12).

The Spectrograph translates the complex acoustic pressure waves into electrical impulses, which illustrate the most basic components of acoustic waves in the form of a spectrogram. This shows frequency plotted against time with the intensity of the input components shown as degrees of blackness of the trace (Figs. 13 a and 13 b) (Fry, 1979, Baken, 1987).

Depending on which aspect of the speech sample is the object of investigation, wide or narrow band analysis of the acoustic input is used. If the duration, onset or off-set of a particular feature is of interest, wide band analysis is used (Fig. 13 c and d). This shows the timing of individual vocal fold vibrations in the form of vertical striations along the horizontal axis. It also allows the identification of the frequency and intensity of formants along the vertical frequency axis.

If the frequency domain is the object of investigation, narrow band spectrographic analysis is used (Fig. 13 a and b) which clearly shows the harmonic structure of a complex periodic tone.

The Spectrograph input consists of both voiced and voiceless speech sounds, and it does not only analyse periodic, voiced sounds, but



Spectrograms

Narrow band, a and b Wide band, c and d.

Figure 13

also aperiodic 'noise' as produced by voiceless sounds. The spectrogram shows such noise, in the case of voiceless fricative or plosive bursts (Fig. 13 a and c), as random traces across a very wide range of frequencies without evidence of striations in the wide band analysis (Fig. 13 c), and without identifiable, horizontal harmonics in the narrow band spectrogram (Fig. 13 a).

In the case of voiceless plosives, a characteristic 'blank', 'silence' appears in the wide band spectrum during the momentary obstruction of the vocal tract. This is followed by a burst of high frequency 'noise' of very short duration (Fig. 13 c).

Spectrography has allowed the study and analysis of speech in a great deal of detail. Formant structure and transitions and knowledge of the manner and place of articulation of speech sounds, allows the skilled observer to 'read' a spectrogram and discover abnormalities of the voice source resulting in abnormal patterns of noise components, striations and harmonic structure (Isshiki, Yanagihara and Morimoto, 1966, Yanagihara, 1967, Cooper, 1974, Schaedel, 1979, Bradley, 1985).

Several more recent studies of hoarseness using spectrography attempt to specify and quantify the relationship between spectral noise and harmonic components in perceived hoarse voices. They all confirm that a characteristic feature of hoarseness is the replacemant of harmonics by noise energy in the spectrum (Kim, Kakita and Hirano, 1982, Yumoto, Gould and Baer, 1982, Yumoto, Sasaki and Okamura, 1984, Bradley, 1985, Wolfe and Steinfatt, 1987).

Sasaki, Okamura and Yumoto (1991) report a method of calculating the ratio between noise energy (N) and total acoustic energy (V) from a section display of a spectrogram. They found significant correlations between the N/V ratio and ratings of the severity of hoarseness judged by experienced listeners. The sample rated was sustained phonation on the vowel /a/ at comfortable pitch and volume.

Spectrographic analysis is complex, the interpretation of results is time consuming and the equipment is expensive. Available instruments are only able to analyse very short speech segments, a few seconds in duration. This limits its usefulness in a clinical setting.

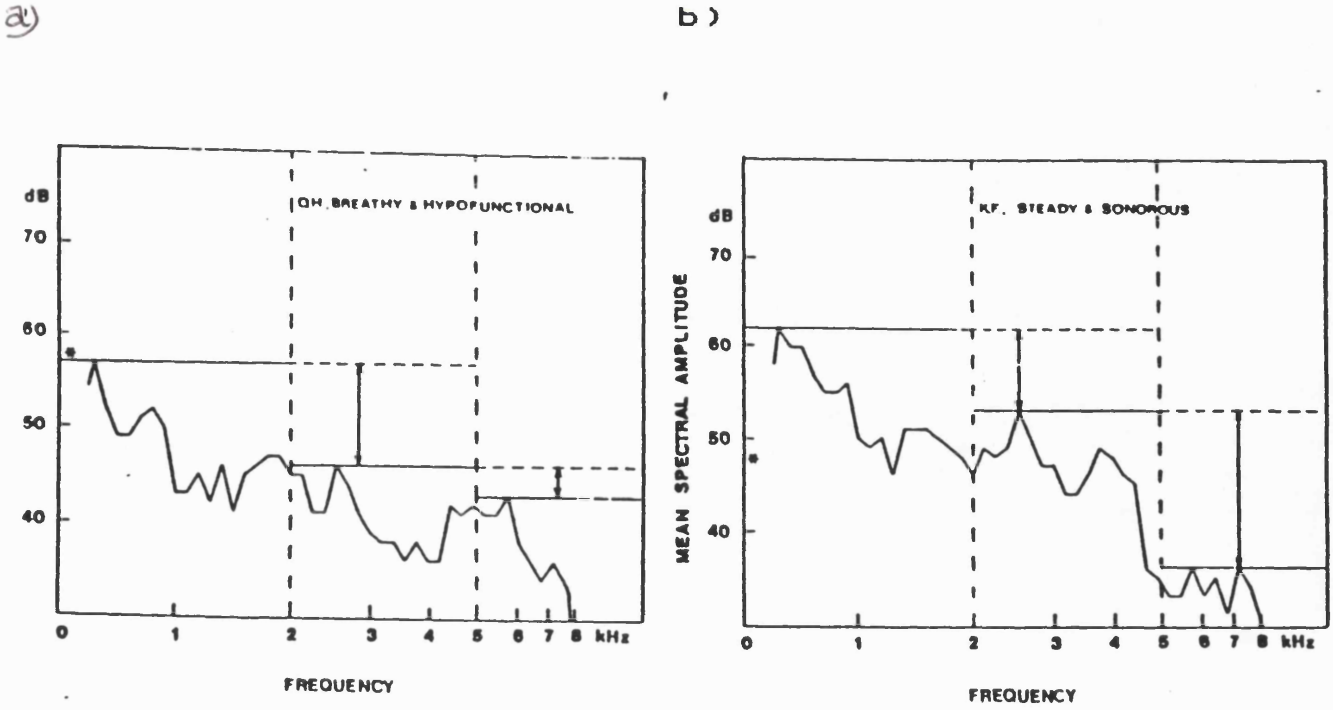
iii Long Time Average Spectrum Analysis.

Long Time Average Spectrum Analysis (LTAS) is a computerized method for spectrographic analysis of speech samples of up to 40 seconds duration, the equivalent of a reading passage of approximately 90 words. It gives information about the mean spectral energy distribution in the sample by band-pass filtering the speech signal, averaging the intensity of each channel and plotting the data in the form of a frequency-intensity diagram (Fig.14) (Hammarberg, 1986).

The method has been used by Frøkjær-Jensen and Prytz (1974) and by Prytz (1977) in an attempt at differentiating three different voice disorders from normal voice by means of LTAS. Prytz concluded however, that LTAS could not be used as a diagnostic instrument as no characteristic curves were identified which differentiated the normal from the disordered voice; nor ones that differentiated between different disorders.

Hammarberg (1986) points out that studies which try to find LTAS attributes of different pathologies, disregard the fact that the perceptual qualities of the same disorder may vary considerably and not all features are necessarily reflected in LTAS parameters. This is also emphasised by Wendler, Rauhut and Krüger (1985).

Nolan (1983) used LTAS to classify 'laryngeal settings' in Laver's (1980) terminology. He found that it was possible to differentiate between falsetto, whispery falsetto and breathy voice, which contrasted in their LTAS with harsh, ventricular and creaky voice quality.



Long Time Average Spectra (LTAS)
a) Breathy voice
b) Steady/sonorous voice

(From: Hammarberg et al, 1986)

Figure 14

Hammarberg, Fritzell, Gauffin, Sundberg and Wedin (1980), Hammarberg, Fritzell and Schiratzki (1984), Hammarberg, Fritzell, Gauffin and Sundberg (1986) and Hammarberg (1986, 1988 and 1992) used LTAS as one of several acoustic measures for correlation with perceptual terms used in voice assessment. In their studies, the instrument was capable of separating out voiceless sounds which otherwise might have interfered with the high frequency noise components from the voice source.

For a 'breathy' voice, LTAS showed a lack of spectral energy in the formant region 0.4-4 kHz and a high spectral level above 5 kHz relative to the level below this point (Fig. 14 a). What Hammarberg termed a 'Steady/sonorous' voice quality, was characterised by a high spectral level below 5 kHz, a low level above 5 kHz and a low level of the fundamental (Fig. 14 b). Another study using LTAS measures for voice assessment found a high spectral level above 7 kHz in LTAS in patients with paralytic dysphonia (Hurme and Sonninen, 1985).

Hammarberg (1986) summarises findings after very extensive studies and concludes that *'Systematic changes in the LTAS agree well with changes in perceptual voice qualities'* e.g 'steady vs sonorous', 'breathy vs hyperfunectional'. She puts this down to the fact that LTAS is sensitive to the contributions of the voice source itself to 'voice quality'. This confirms what has earlier been shown in the section on Spectrography.

Baken (1987) points out, however, that LTAS requires a considerable amount of sophisticated hardware and computer support and is therefore not widely available in a clinical setting even if it has shown promise as an objective means of assessing voice quality.

B GLOTTOGRAPHIC TECHNIQUES.

Other indirect methods of laryngeal examination have been developed, which attempt to record and analyse the function of the voice source itself without the need for filtering and processing of complex acoustic information. Examples of such methods are Electroglottography or EGG.

i Inverse filtering and the Flow glottogram.

Inverse filtering is based on our knowledge of the formant frequecies of different vowel sounds. Fig. 12 c shows the effect of the filtering of the harmonic rich glottal signal in the vocal tract. This curve is called the 'Vocal tract transfer function'. Those harmonics that fall at or close to the resonance peaks, i.e. the formant frequecies, in the vocal tract transfer function, are enhanced. Those that are further away from the resonance peaks are attenuated.

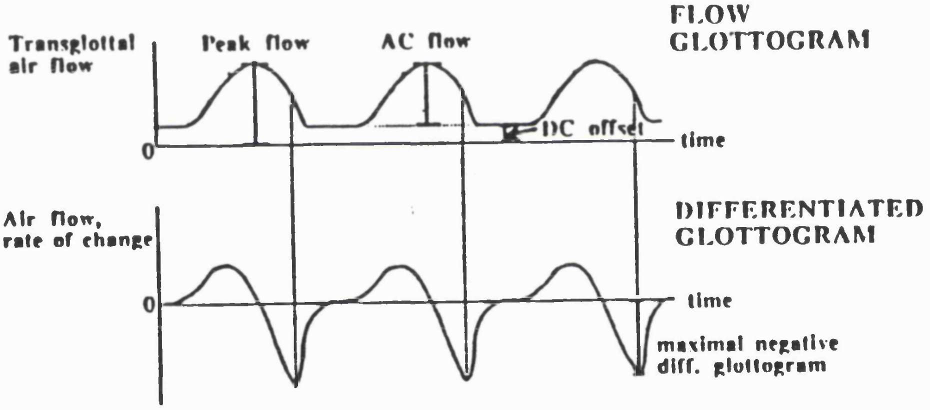
The position of the formants change (i.e. the 'vocal tract transfer function changes) with the changes in the vocal tract as different vowels are produced.

Sundberg (1987) describes an Inverse Filter as a *'series of filters, ideally one for each formant, the transfer function of which is the negation of the transfer function of a formant.'* The filters are tuned so as to neutralize or cancel out formant frequencies of vowel sounds and eliminate the influence of those formants on the spectrum. What is left is a signal with a spectrum corresponding to the voice source (Fig. 12 b). This gives information about the harmonics or partials that constitute the complex tone, before the filtering effect of the vocal tract.

Frequency modulated (FM) recordings allow the analysis of the waveform of the voice source, in the form of *'flow glottograms'* (Fig. 12 a).

To arrive at the pure voice source waveform illustrated by the flow glottogram (Fig. 12 a), which reflects only the behaviour of the vocal folds, Rothenberg (1973) devised a mask vented round its circumference, the vents covered with a fine mesh. The mask contains a microphone which measures pressure differences across the mesh. These correspond to the variation in air flow as the vocal folds open and close. The subject puts the mask over his mouth and nose while repeating syllables usually [pæ]. The choice of the syllable [pæ] ensures an effective velopharyngeal closure and no airflow before the vowel sound which is important for the detection of the glottal flow waveform during vowel production.

The resulting inverse filtered waveform, illustrates the transglottal air flow over time (Fig. 15). It shows vocal fold vibration in terms of the duration of the open and closed phases and the volume of air flowing through the glottis at any one time during each cycle. If the waveform does not reach zero during the closed phase, this indicates lack of complete vocal fold closure in the vibrational cycle. It is often referred to as the 'DC offset' (Holmberg, Hillman and Perkell, 1987).



Flow Glottogram DC 'offset'.
(From: Hertegard and Gauffin, 1991)

Figure 15

Sundberg and Gauffin (1979) and Gauffin and Sundberg (1980) related flow glottograms (FGG) to acoustic voice parameters and found that the higher the amplitude of the glottogram, the stronger was the voice source fundamental (cf.Fig.12 b). Fant (1979) showed that the faster the closing rate of the flow glottogram waveform (i.e. the steeper the falling off of the transglottal airflow curve on vocal fold closure Fig. 15) the stronger were the amplitudes of higher frequency partials.

The flow glottogram amplitude, the rate of vocal fold closure and whether the waveform reaches the baseline during the closed phase, are all important features reflecting the mode and efficiency of vocal fold vibration. The DC 'offset' of the waveform from zero during the closed phase is measured in ml/sec and used to quantify 'glottal insufficiency' (Fig. 15) (Hertegård et al 1992)

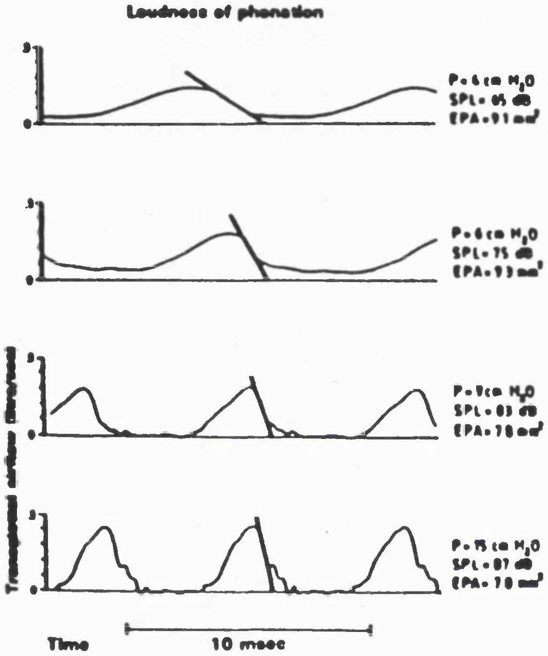
Sundberg and Gauffin (1979) contrast what they call 'Flow phonation' (Fig. 16 c) with extreme 'Pressed' phonation (Fig. 16 a) which would combine high adductive tension with high subglottal pressure and lead to minimum sound level achieved with maximum effort.

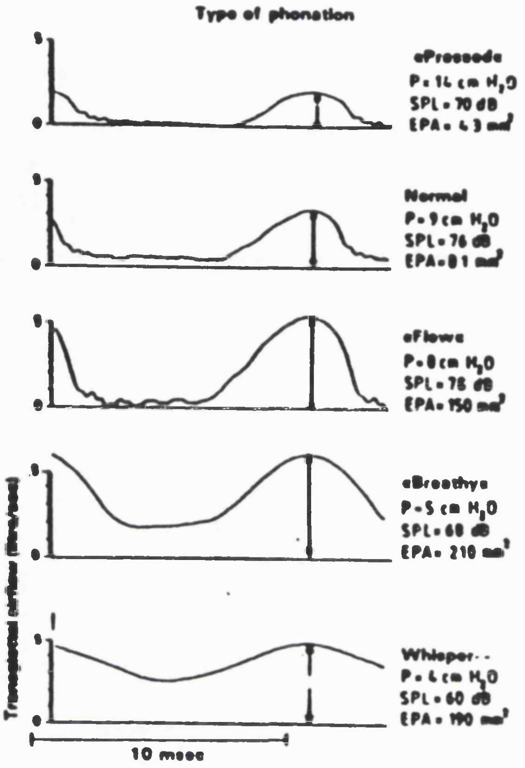
Flow phonation is characterized by a high flow glottogram amplitude and a fast closing rate, with definite evidence of zero flow during the closed phase. Sundberg and Gauffin (1979) suggest that the Bernoulli effect plays a more important part in closure during flow phonation than in pressed phonation, where muscular adduction forces would play a greater part in closure of the vocal folds.

The effect on the flow glottogram of different types of phonation, from 'Pressed' to Whisper, in terms of a) the amplitude of the waveform b) the speed of vocal fold closure indicated by the steepness of the negative slope c) whether complete closure is achieved and d) the duration of closure, is shown in Fig. 16.

Sundberg (1987) suggests that the position of phonation types along the 'Pressed' - 'Breathy' continuum, is signalled by the amplitude of the flow glottogram, which illustrates the amplitude of the source spectrum fundamental.

The effect of varying degrees of loudness expressed as Sound Pressure Level (SPL) is shown in Figure 17 a-d (Sundberg, 1987).





Typical glottogram changes obtained when loudness or mode of phonation is changed. As loudness of phonation is raised, the closing part of the curve becomes steeper. When phonation is 'Pressed' the glottogram amplitude is low and increases as adduction force is reduced. The least adducted, but not 'leaky' phonation is called flow phonation.
To the right of the curves are given subglottic pressure (P), sound pressure level (SPL) at 0.5 m, and estimated glottal peak area (EPA). In pressed phonation pressure is high, flow amplitude is low and sound level low, while in flow phonation pressure is moderate, flow amplitude high and sound level is high. (From: Sundberg, 1987)

Loudness of phonation, effect on flow glottogram.

Figure 16

Type of phonation, effect on flow glottogram.

Figure 17

Perceived vocal loudness is related to the closing rate (steepness and tilt to the right) of the flow glottogram.

Such aspects of glottal airflow waveforms and ratios of modulated (AC) and unmodulated (DC) flow (Fig. 15) have been used by Hillman et al (1989) for objective description of vocal hyperfunction. They found that AC flow and maximum flow declination rate, indicating speed of vocal fold closure, were abnormally high in their subjects with organic manifestations of adductive hyperfunction e.g. polyps and nodules, as opposed to subjects with non-addyuctive hyperfunction and without organic lesions.

The above illustrates the relationships between physical modes of vibration of the vocal folds, glottal aerodynamics, acoustic spectral parameters and the perceptual attribute of loudness.

Inverse filtering and analysis of flow glottograms in normal speakers has been reported by Holmberg, Hillman and Perkell (1987) and Hertegård, Gauffin and Karlsson (1992) and in vocal pathology by Hammarberg and her co-workers (1980 and 1986), Fritzell et al (1986) and Hillman, Holmberg et al (1989). Hammarberg (1988) suggests that the voice qualities that are best illustrated by Inverse filtering are features such as 'breathy', 'hyper- or hypofunctional', vocal fry (creak) and the difference between modal and falsetto register. These voice qualities are differentiated by the degrees of vocal fold closure during phonation, which can be estimated from the flow glottogram waveform (DC) offset (Fig. 15 and 16).

Holmberg et al (1988) found that almost all their normal subjects had DC offsets in their glottal waveforms, assumed to be related to a posterior glottic chink.

ii Photoglottography.

Photoglottography (PGG) is a "semiinvasive" means of observing glottal activity. It requires a powerful lightsource to be applied outside the neck just below the cricoid cartilage. A lightsensitive probe is inserted via the nose into the laryngeal vestibule where it registers variations in light through the changes in glottal area during phonation. These variations are translated into waveforms which reflect the glottal area variations (Fant, Ondraçkova, Lindkvist, and Sonesson, 1967, Dejonkere, 1981, Kitzing, 1982, Kitzing, Carlborg and Löfqvist, 1982, Baer, Titze and Yoshioka, 1983 b, Cranen, 1991, Gerrat, Hanson, Berke and Precoda, 1991).

iii Ultrasonic glottography.

Kaneko, Uchida et al, (1981) have been at the forefront of developing Ultrasonic glottography for observation of glottal activity.

An Ultrasonic beam is generated and received by a transducer placed on the skin over the thyroid cartilage. *'Ultrasonic pulses pass from the skin through prelaryngeal muscles, thyroid cartilage and vocal muscles, and they reach the free margin of the vocal fold. When they reach the margin, most of the ultrasonic energy is reflected at the boundary surface and returns the same way to the transducer. The pulses are received by the same probe and are represented as an echo of the vocal fold.'*(Kaneko et al, 1981).

Kaneko and co-workers (1981) have used Ultrasound to study a wide variety of laryngeal function e.g. horizontal and vertical components of vibration of the vocal folds, phase differences between the upper and lower margins of the vocal folds, influences of vocal tract shape, pitch and loudness on the glottal movement and changes of contact area between the vocal folds during one vibratory cycle. They also report that ultrasound can give information about the submucous extension of laryngeal tumours.

Constant ultrasonic waves applied to the neck with a receiver applied on the other side of the thyroid cartilage from the transducer, will be interrupted every time the glottis opens. On the basis of this, 'Continuous Wave glottography' (CW) was developed, which shows open-closed phases corresponding to those obtained by Electroglottographic methods (Hamlet, 1971, Holmer and Rundqvist, 1975).

The CW amplitude also changes in proportion to vocal fold contact area (Hamlet and Palmer, 1974) and narrow beam CW ultrasound has been used to demonstrate vertical larynx height (Hamlet, 1980).

Ultrasonic glottography however, does not seem to add any information that other, more accessible methods are already offering. It is expensive, and obtaining and interpreting the glottograms requires considerable training (Baken, 1987).

CHAPTER V - ELECTROGLOTTOGRAPHY (EGG) AND ELECTROLARYNGOGRAPHY (ELG).

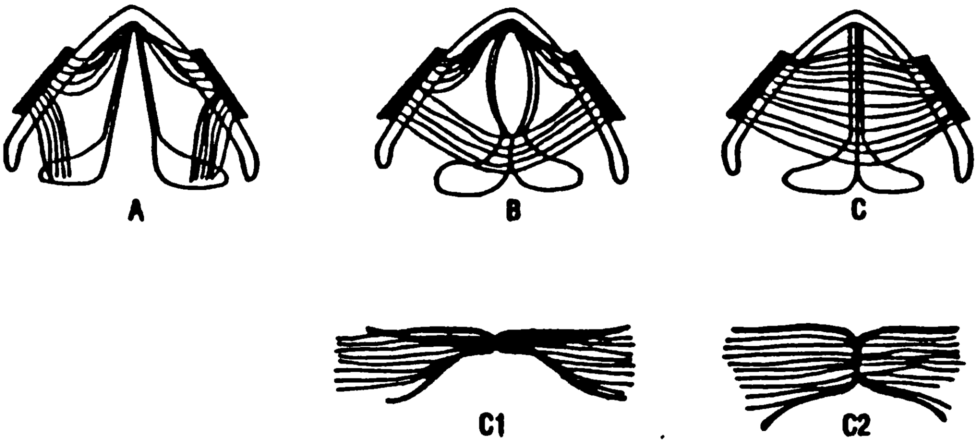
1 The development of EGG and ELG.

Electroglottographic (EGG) waveforms reflect the electrical resistance across the glottis during each cycle of vibration by means of application of superficial electrodes either side of the thyroid cartilage at the level of the vocal folds (Fig. 18) (Kitzing, 1979, 1986, Titze and Talkien 1979, Titze, 1984, 1989c, 1990, Childers et al 1983, 1984, 1985, 1986 a,b, 1990, Orlikoff, 1991 a, Rothenberg, 1992) It is, like Ultrasound, a completely noninvasive means of observing vocal fold function.

The origins of Electroglottography (EGG) stem from work done by Fabre (1957, 1958), who applied a technique previously used to record arterial bloodflow, to measure impedance changes across the glottis during phonation. Fabre interpreted the waveforms from his Electroglottograph as indicative of variations in glottal area.

Much research and many developments later of this technique, which showed promise due to its being entirely non-invasive, have shown that the most important information gained from EGG waveforms is gained from the very rapid drop in impedance on vocal fold contact (Fant, Ondráckova, Lindkvist and Sonesson, 1966, Frøkjær-Jensen and Thorvaldsen,1968, Fourcin and Abberton, 1971, 1976, Fourcin 1974, 1981, 1982, Lecluse, Brocaar and Vershuure, 1975, Kelman, 1981, Gilbert, Potter and Hoodin, 1984). This is the reason why Fourcin calls his device for measuring and displaying laryngeal impedance variations a 'Laryngograph', and the method using the Fourcin electroglottograph will from now on be referred to as ***Electrolaryngography or ELG.***

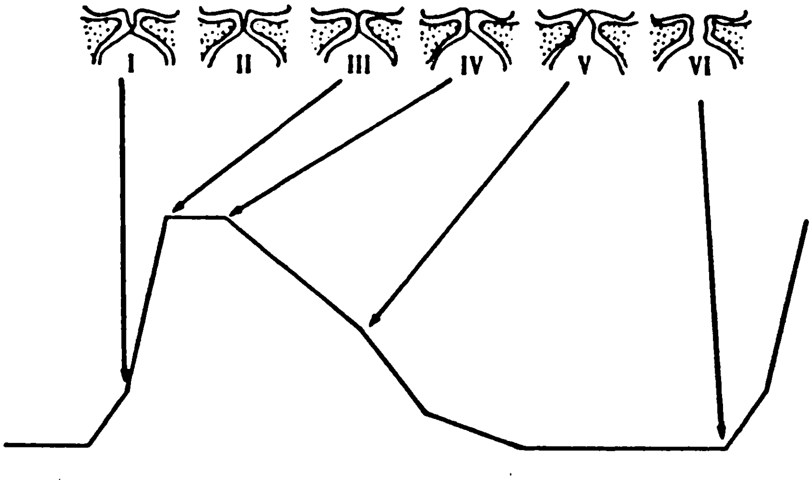
Fourcin made considerable changes to the Fabre glottograph, which turned out to have a very unstable output with great variation in output for the same speaker, let alone between speakers.



Current flow between superficial laryngeal electrodes. (From Baken, 1987)

A) Ventilatory posititon of the folds
B) Open phase of the phonatory cycle.
C) Closed phase of the phonatory cycle.
C1 and C2 different degrees of vocal fold contact during the closed phase in coronal section.

Figure 18



ELG (Lx) waveform relationship to vocal fold contact (From: Baken, 1987).

Figure 19

The current version of the Laryngograph (Abberton, Howard and Fourcin, 1989), monitors the variation of electrical impedance between two goldplated electrodes, applied either side of the thyroid cartilage (Fig. 18). A constant voltage is applied, and the current flow between the electrodes varies in synchrony with the vibration of the vocal folds. Both the input and output from the neck are detected in a low impedance circuit and the electrodes have guard rings at earth potential to reduce surface conduction across the skin (Fourcin and Abberton, 1971).

The current flow will be at is maximum during contact and at a minimum when the folds are apart. Figure 19 shows the relationship between points on the Lx waveform and degrees of contact between the vocal folds (Rothenberg, 1981, MacCurtain and Fourcin, 1982).

The modifications have been made to reduce the intersubject variation in output inherent in the Fabre glottograph.

Another modification by Fourcin is the inverted phase of the displayed output waveform. The Larynx excitation - Lx waveform - is positive going for increasing vocal fold contact (Fig. 19) and shows a left tilt, due to the abruptness of closure compared to the more gradual opening phase. This is the opposite of the Flow glottogram waveform shown in Fig. 15 which is tilted to the right. The degree of this tilt is related to loudness of phonation as shown by Sundberg (1987) and illustrated in Fig. 16.

Titze (1984) in his computer modeling of EGG waveforms, which will be described in detail in a later section, suggests that the left tilt of the EGG/ELG waveform (later also referred to as 'contact area' waveform) (cf. Lx) is due to a convergent glottal crossection and vertical phase differences on vocal fold closure, whereas the right tilt of the flow glottogram waveform results from the effects of air pressure and flow on the vocal folds, what he calls vocal tract loading. This seems to agree with Sundberg's explanation in terms of an increased Bernouilly effect in his continuum of voice qualities (Fig. 17).

The difference between the flow glottogram (FGG) described in the previous chapter, and the EGG and ELG (Lx) waveform is that the former gives information about variation in airflow between the vocal folds as a function of time. i.e about the opening and open phases of the cycle. It will also reveal if there is any glottal air leakage during the closed phase in the DC offset (Fig. 15).

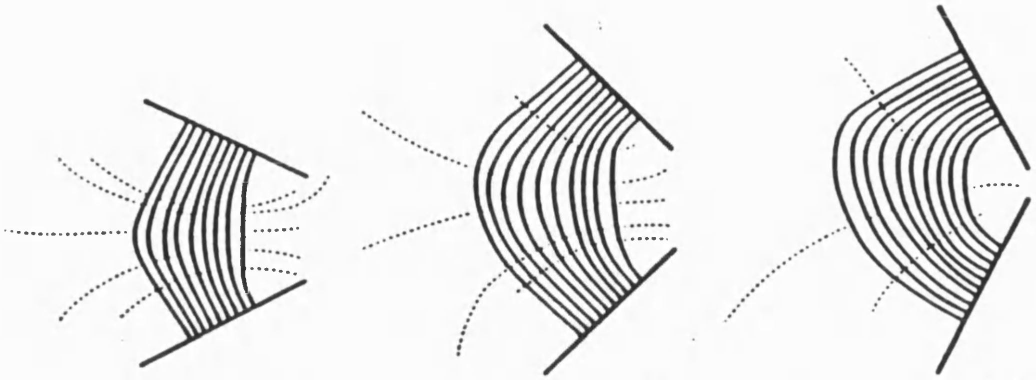
Lx, and other EGG waveforms, reflect the timing of the start of vocal fold contact and of maximum contact, contact area and the condition of the current path between the electrodes (Titze and Talkin, 1981, Childers , Hicks, Moore and Alsaka, 1986). The steepness of the closing (rising) slope of Lx will also reflect the speed of closure of the vocal folds, which determines the acoustic excitation of the vocal tract. Methods used for verification and interpretation of EGG and ELG waveforms will be described in a later section.

ii Interpretation of amplitude variations in the EGG/ELG waveform.

In clinical practise, an important feature of the EGG/Lx waveforms is the relative amplitude, as it is related to the degree of vocal fold contact. The explanation for this lies in the understanding of electrical conductivity through different media. To quote Childers, Hicks, Moore and Alsaka (1986) '*...the resistance of a substance is inversely proportional to the crossectional area of the substance.*'

They highlight the effect of vocal fold contact on the amplitude of the EGG waveform by the explanation that '*...total impedance is a function of the tissue path length as well as the tissue crossectional area and composition...impedance increases as the folds separate because the current paths become fewer and less direct having to pass around the anterior and posterior regions of the larynx.*' (Fig. 18 and 20)

It then follows, that when the folds are in maximum vertical *and horizontal* contact, this creates the minimum impedance (resulting in minimum of the EGG waveform = maximum of the ELG (Lx) waveform), because in this instance there are many parallel conductive paths

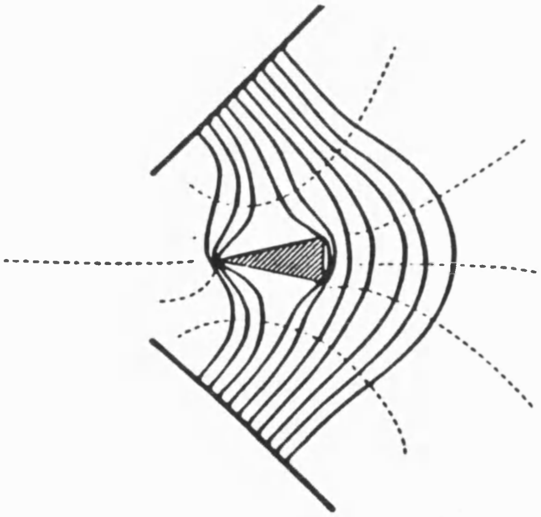


100 kHz, round electrodes, 60°
(a)

100 kHz, round electrodes, 90°
(b)

100 kHz, round electrodes, 120°
(c)

Equipotential lines (dashed) and electric field lines (solid) between round electrodes placed at an angle of a) 60° b) 90° and c) 120°. Electrode diqameter 1.5 cm. Frequency 100 kHz. (From: Titze, 1990).



Equipotential lines and electric field lines between the same round electrodes placed at an angle of 90° with an acrylic (non conducting) wedge included to simulate the glottis.

Figure 20

between the laryngeal electrodes (Fig. 19). *'The combined total parallel impedance is less than the impedance of any one path.'* (Childers, Hicks, Moore and Alsaka, 1986).

Gilbert, Potter and Hoodin (1984) performed an experiment where a high resistance polymer strip was inserted between the vocal folds and gradually withdrawn, on sustained phonation. The strip was inserted to the complete vertical depth of the vocal folds and covered 5 mm of glottal length. The Lx waveform was present but reduced while the strip was in place. The Lx amplitude was shown to increase as the polymer strip was withdrawn allowing the area of vocal fold contact to increase.

The EGG/Lx waveform does not differentiate between the degree of vertical and horizontal closure, but gives an integrated measure of the total current flow between the electrodes. Anastaplo and Karnell (1988), however, tried to find out if the discontinuity or 'knee' in the opening phase (Fig. 19), which sometimes appears in the waveform, corresponded to the change in vocal fold opening from the vertical to the horizontal plane. They found that it did not always appear in the waveforms even of the same speaker, but where it did, it did correspond to such a gesture.

Anastaplo and Karnell (1988) used four adult speakers and found some degree of posterior glottic chink in all their speakers, greater in women than in men. Incomplete posterior vocal fold closure, sometimes involving part of the membranous portion has been found to be common in women without giving rise to abnormal voice quality (Holmberg, Hillman and Perkell 1987, Biever and Bless, 1989, Södersten and Lindestad, 1990). The extent of incomplete closure will not be possible to determine from the Lx waveform, except in a calculation of open and closed quotients (Fourcin 1989 a and b, Howard Lindsey and Allen, 1990, Howard, Lindsey and Palmer, 1991), but it is likely to affect the relative Lx amplitude.

From this it would follow that vocal fold closure e.g. in a small size larynx and/or production of higher pitch as in female speakers or children, would reduce the relative amplitude of EGG or Lx, compared to the amplitude produced by a larger male or female larynx (Titze, 1990). This is what is found in clinical practise. Phonation at low intensity has also been found to be performed with incomplete closure of the glottis and is reflected in EGG/Lx amplitude (Biever and Bless, 1989, Södersten and Lindestad, 1990).

Low EGG/Lx amplitude might also reflect limited closure due to bowing vocal folds or due to an anterior web, which bridges the vocal folds at the anterior commissure, as in one of the irradiated speakers who will be described in this study.

Alternating high and low Lx peaks, i.e. Lx amplitude perturbation, giving rise to perceived roughness, are not infrequently produced by pathological voices and will be illustrated in our analysis of irradiated speakers' Lx waveforms.

Because of the above it is important, when interpreting the EGG/Lx waveform, not to assume 'complete closure' at the peak of the waveform (Lx), but to consider this the 'point of maximum contact' during any one cycle, as suggested by Orlikoff (1991)

iii Problems in the interpretation of EGG waveforms.

A great deal of research has been carried out in an attempt to verify and interpret the shape and phases of EGG and ELG waveforms. Electroglottographic methods for studying vocal fold vibration are simple to apply and the waveforms seem easy to interpret, compared to for example drawing conclusions about vocal fold behaviour from the interpretation of acoustic spectrograms or speech waveforms. The danger is that too much may be inferred about vocal fold vibratory behaviour from EGG/ELG waveforms alone (Childers and Krishnamurthy ,1985, Rothenberg, 1984, Baken, 1987, Colton and Conture, 1990) and there are several problems inherent in the methods of data collection, exacerbated by the recording conditions.

Baken (1987) summarises the pitfalls inherent in the recording and interpretation of EGG. More recently Colton and Conture (1990) give a very extensive historic overview and critical evaluation of EGG methods.They account for three potential sources of difficulties in the interpretation of EGG waveforms, instrumental, 'procedural' and 'subject concerns'. They base this on their extensive experience of clinical application of EGG and routine evaluation of over 100 patients with a variety of vocal and other communication disorders.

Instrumental problems arise from the fact that the EGG is a composite signal resulting from, not only the fast changes in electrical impedance with variation in vocal fold contact, but also from slow movement in other structures during speech production e.g the tongue, epiglottis, ventricular folds and the laryngopharynx. Fourcin (1981) deals with this problem by high pass filtering and amplifying the original waveform, Gx, to arrive at the Lx waveform (Fig. 19), where such low frequency variations are not much in evidence.

Automatic gain control (AGC) incorporated by manufacturers of EGG devices to control for different neck resistance among speakers is another problem that may introduce artefacts into the shape of the open phase of the EGG waveform. The high pass filtering performed may also distort the shape of the waveform according to Colton and Conture (1990). This problem is lessened by the use of a linear phase high pass filter as described by Rothenberg and Mahshie (1988) and they suggest that manufacturers reveal the type of filter that has been used to help users interpret EGG waveforms.

The procedure used for recording EGG is another potential source of problems, including finding the correct placement of the electrodes and keeping them in the right place during recording. Colton and Conture (1990) found that asking the subject to hold the electrodes against the neck solved the problem of the larynx moving out of the way of the current during speech, which, they suggest, may be more likely if a neck band holding them in place is being used. It may cause variable neck-electrode contact, however.

What Colton and Conture (1990) call 'subject concerns' relate to the fact that women and children tend on the whole to produce poorer EGG signals because of their smaller vocal folds than adult males with large protruding 'Adams apples'. They do not offer such a large area for the current to pass through and therefore they do not produce the amount of current variation reflected in the EGG amplitude (Fig. 18). Other physical difference are the angle of the thyroid cartilage (Fig. 20) which is wider in women and children (Kahane, 1978, Titze, 1990), and a tendency for women to have more adipose tissue covering the larynx, providing a longer tissue path for the current, thereby greater resistance resulting in a weaker EGG signal.

Rothenberg and Rothenberg (1989) studied the effect on the EGG of different neck circumferences and 'composition of the current path' whether of adipose or muscular tissue. They found that the neck circumference showed a moderate correlation with total neck impedance. The peak to peak change in impedance was greater in average size men and women with low total neck impedance than in subjects with large total neck impedance i.e with 'fat' necks, who produced very small peak to peak changes. They found the total neck impedance was greater in subjects with adipose rather than muscular necks.

Baken (1987) brings up a problem that arises with the taperecording of EGG waveforms for later analysis. If the recording is done on an ordinary amplitude (AM) modulated taperecorder, the play back will introduce shape and phase distortions of the waveform due to loss of very low frequencies reflected in the flat portion of the EGG waveform. Dealing with AM taperecording problems, Judd (1983) described a method for correcting low-frequency phase distortion in Lx recordings on playback. FM recording is otherwise recommended if played back waveforms are going to be used for interpretaton of vocal fold behaviour. Waveforms have also been printed, or photographed from an oscilloscope screen, with the subject phonating 'live', which also overcomes this difficulty (Painter, 1981, Carlson, 1988, 1993). The recent wider availability of digital audio recorders (DAT), however, ensures that the waveform shape and phase is retained for future playback without distortion.

Despite all these potential sources of difficulties Colton and Conture (1990) have found only 15 % of all their EGG waveforms impossible to interpret.

Trying to overcome some of the above mentioned difficulties in the recording of EGG waveforms, Titze (1990) studied the effect on the electrical field patterns between rectangular and round electrodes, different size electrodes and the angle and distance between them (Fig. 20 a, p. 121). Figure 20 b shows a simulation of the glottis by introducing a non-conductive wedge between the electrodes. He concluded from his experiment that, ideally, the electrode dimensions should be comparable to those of the examined glottis. This, he suggests, may be accomplished by having interchangeable sets of electrodes.

Secondly, the distance and the angle between the electrodes should be small. The angle should allow the electrodes to be as close to parallel as possible. This is due to the signal to noise ratio being determined by the total number of electric field lines that are shunted behind and in front of the larynx compared to the total number of lines between the electrodes (Fig. 20 b, p. 121).

To ensure the correct placement of the electrodes and avoid voice synchronous noise in the signal, Rothenberg (1992) describes a 'multichannel glottograph' with a vertical arrangement of electrodes.

Baken (1987) expresses concern about the risk of skin-electrode resistance introducing artefacts into the EGG waveform. Colton and Conture (1990) suggest cleaning the subject's neck with alcohol to remove as much surface skin oil as possible before applying the electrodes and also to use electrode gel to maintain a low resistance interface between the skin and the electrodes. The skin-electrode resistance is reduced in the Fourcin Laryngograph by providing a guard ring at earth potential around the electrodes (Fourcin and Abberton, 1971).

iv Verification of EGG/ELG waveforms.

The relationship between electroglottographic waveforms and vocal fold modes of contact and vibration has been studied with simultaneous, synchronised stroboscopic filming by Fourcin, (1974), Lecluse, Brocaar and Verschuure, (1975), Fog-Pedersen (1977), Sopko, (1986), Anastaplo and Karnell, (1988), Hertegård, Gauffin and Karlsson, (1992); by simultaneous ultra high speed filming by Van Michael, Pfister and Luchsinger (1970) and by Childers with a number of co-workers between 1977 and 1990; using ultra-short pulse X-ray by Noscoe, Fourcin, Brown and Berry, 1983, using Photoglottography by Fant, Ondrackova, Lindqvist and Sonesson (1967), Dejonckere (1981), Kitzing, 1982, Kitzing , Carlborg and Löfqvist (1982), Baer, Lofqvist and McGarr (1983 a), by Baer, Titze and Yoshioka (1983 b) and by Carding and Murty (1991).

Titze and Talkin (1981) compared EGG and flow glottogram waveforms and confirmed that the EGG, indicating the area of vocal fold contact provided more information during the closed phase, whereas volume flow provided more information during the open phase of the glottal cycle.

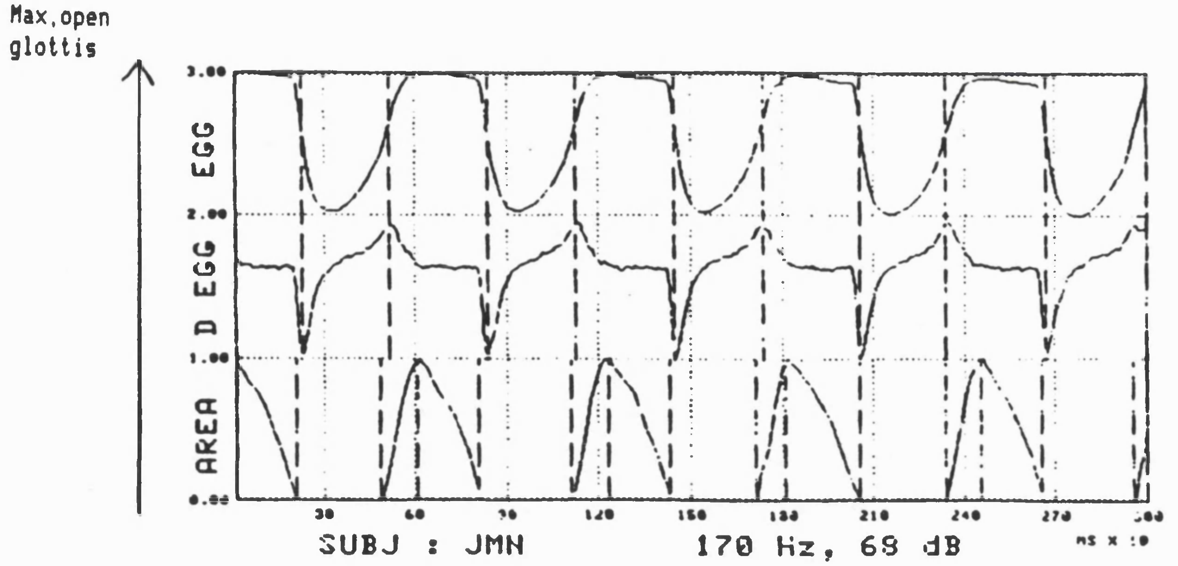
Kitzing (1982) used Photoglottography (PGG) and Electroglottography (EGG) to study register bound changes in vocal fold vibratory patterns in a trained and an untrained singer. There was close correspondence between a distinct closed segment in the PGG and the time of decreased impedance in the EGG. The closed segments had an approximate duration of 40% of the total cycle and an Open Qoutient (OQ = Open Time/Total period time) of 0.61.

Experiments where direct ultrahigh speed filming has been used in conjuction with Photoglottography, Electroglottography and/or analysis of the speech waveform, offer an opportunity to observe real glottal activity in different dimensions.

Baer et al (1983 a) found that measurements from photoglottography (PGG) and high speed filming, gave much the same information regarding peak glottal opening and glottal closure. They feel confident that the EGG signal reliably indicates vocal fold contact. They further maintain that PGG and EGG together provide much of the information obtained from filming and besides, potentially show horizontal phase differences during opening and closing. In their male subject the EGG showed a slope discontinuity at the instant of glottal opening, as identified on the laryngeal film, whereas the female EGG showed a gradual glottal opening with large observed horizontal phase differences.

Childers, Smith and Moore (1984) used simultaneous recordings of EGG, and acoustic signals with synchronous ultrahigh speed filming of the vocal folds during production of a sustained [i]. Their periodicity and jitter analysis of the acoustic and EGG signals gave the same results, and there was a high degree of correlation between events in the EGG signal and high speed film of vocal fold vibratory events.

Childers and Krishnamurthy (1985) give an extensive and detailed overview of the state of the art of Electroglottography until 1985, and include a detailed account and interpretation of EGG based on their own recorded data. This consists of ultrahigh speed laryngeal films synchronised with EGG and acoustic voice recordings of subjects performing nine phonatory tasks, varying voice intensity and



Comparison of the timing of maximum glottal opening and closure in EGG and area waveforms.
(From; Childers and Krishnamurthy, 1985)

Figure 21

fundamental frequency.Their EGG waveforms are shown with decreasing vocal fold contact in a positive direction (Fig. 21)

They identify the vocal fold ***closed phase*** as beginning at the point of maximum glottal contact as indicated by the glottal area curve decreasing to zero (c) (Fig. 21). The figure shows how this instant is indicated in the EGG by the rapid fall in the waveform (c). However, the minimum in the EGG corresponding to maximum lateral fold contact, occurs slightly after this moment. Childers and Krishnamurthy (1985) suggest:

> *'..this minimum corresponds to vocal fold extension inferiorly and superiorly as a result of the elastic collision of the tissue. This elastic collision causes the rounding of the EGG waveform at its minimum extension. Occasionally the EGG will be nearly flat instead of rounded at this interval, which we feel reflects the fact that the depth of contact is constant or the area of contact is constant... The EGG begins to increase from its minimum while still in the closed phase, reflecting the separation of the folds along the inferior surfaces toward the upper margins. '*

They define the glottal ***opening phase*** as the time from the instant of glottal opening shown in the glottal **area waveform** (a) (Fig. 21) to the maximum value of the glottal area (b). Laryngeal films show a gradual initial opening with large horizontal phase differences during this phase.

> *'First the EGG increases monotonically reflecting the decreasing lateral contact between the folds as they start the opening phase. During this interval the EGG is concave downward. Once the folds have separated, the EGG remains constant while the folds pull further apart. '*

The interval between the maximum (b) to zero glottal area (c) in the area waveform is termed the ***closing phase***. It is found to be of the same duration or slightly longer than that for opening (d-a, Fig. 21).

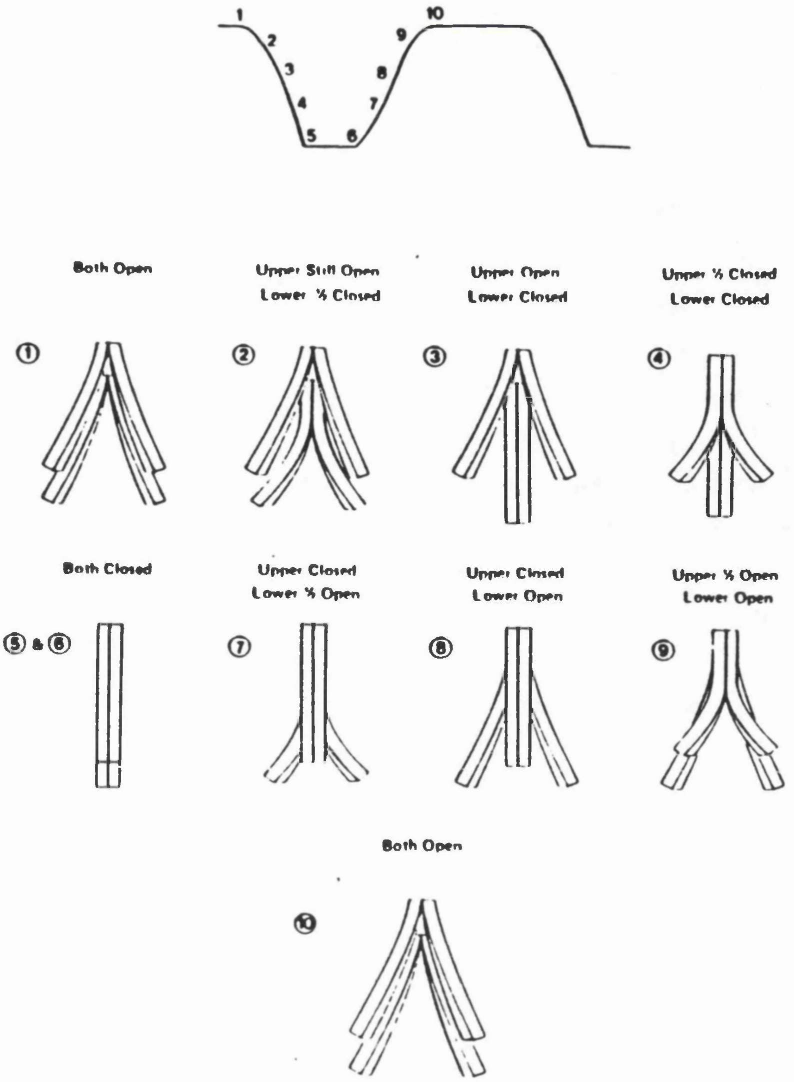
Childers and Krishnamurthy (1985) make interesting observations of two different types of closure. One type is described as occurring in the following manner:

> *Over a large portion of the closing phase, the vocal folds adduct towards their medial position with little or no change in the length of contact along the midsagittal line. Just prior to closure, the vocal folds are almost parallel with a narrow opening along their entire length. Closure occurs almost simultaneously along the entire mid sagittal line....while the glottal area does not reflect this fact, the glottal closure is an abrupt phenomenon. This type of closure is typically seen as the pitch is raised.* '

The second type of closure is described as occurring with the vocal folds first closing anteriorly, remaining nearly fully open posteriorly:

> '*Initial fold contact is made inferiorly along the lower margin. Closure occurs gradually along the midsagittal line from anterior to posterior in a zipperlike manner. Simultaneously the folds roll together from the lower margin toward the upper margin.*'

This latter type of closure seems to be the one on which Childers et al (1986 a) base their elastic two-mass model (Fig. 22). In this type of closure the EGG shows two stages. In the first stage the EGG drops at a moderately steep slope, then the slope is suddenly very steep. Childers and Krishnamurthy (1985) feel this is a result of the gradual vocal fold contact anterior to posterior, followed by a more rapid lateral closure.



Elastic two-mass model of the vocal folds and the EGG waveform. (From Childers et al, 1986)

Figure 22

Anastaplo and Karnell (1988) used synchronised EGG and videostroboscopy to document the often observed 'knee' in the opening stage of the EGG waveform (Fig. 19, p. 118). This is thought to reflect vocal fold separation changing from vertical to horizontal along the superior margin (Fourcin, 1981, Rothenberg, 1988, Childers and Krishnamurthy,1985). They found that in the cases where an obvious 'knee' was visible in the EGG waveform, this did coincide with vocal fold opening along the superior border. They also found that this usually proceeded from posterior to anterior in the normal larynges they investigated.

However, complete glottal closure just before the 'knee' appeared, was only observed in one male subject and then only inconsistently. The relative size of the pre-knee opening was greater for the female than for the male subjects. In the light of these findings Anastaplo and Karnell (1988) reiterate warnings against inferences about vocal fold physiology from EGG waveforms without additional validation.

The vertical depth of closure is small just before opening. Baer (1981) reported the depth of closure as almost negligible immediately before the glottis opened. The lower impedance observed must then be the result of higher conductivity of the contacting surfaces e.g. through free mucus bridging the vocal folds on opening. Colton and Conture (1990) illustrate an example from a patient where mucus is assumed to have caused an additional current path just before complete opening of the vocal folds.

The effect of mucus on the Lx waveform is noted by Fourcin (1982). He suggests, however, that the first contact between the vocal folds on closure, is by a mucus bridge between the approaching epithelia. This would cause the very rapid rise in the Lx waveform (Fig. 19, p. 118) due to increased conductance. Fourcin suggests, regarding the interpretation of the ELG (Lx) waveform, that it only gives reliable information about glottal closure. It is not recommended for drawing conclusions about glottal opening configurations. Colton and Conture (1990) agree.

In summary, because of the dramatic effect on the waveform, with tissue contact at onset and during closure, this phase of the cycle is relatively easy to identify. The effect on the electrical current path of the gradual separation of the folds, in the horizontal and vertical plane, during opening, is more diffuse and may be further blurred by the existence of a mucus strand bridging the folds. The presence of oedema or growths may also have the effect of delaying separation (Childers and Krishnamurthy, 1985, Colton and Contour, 1990).

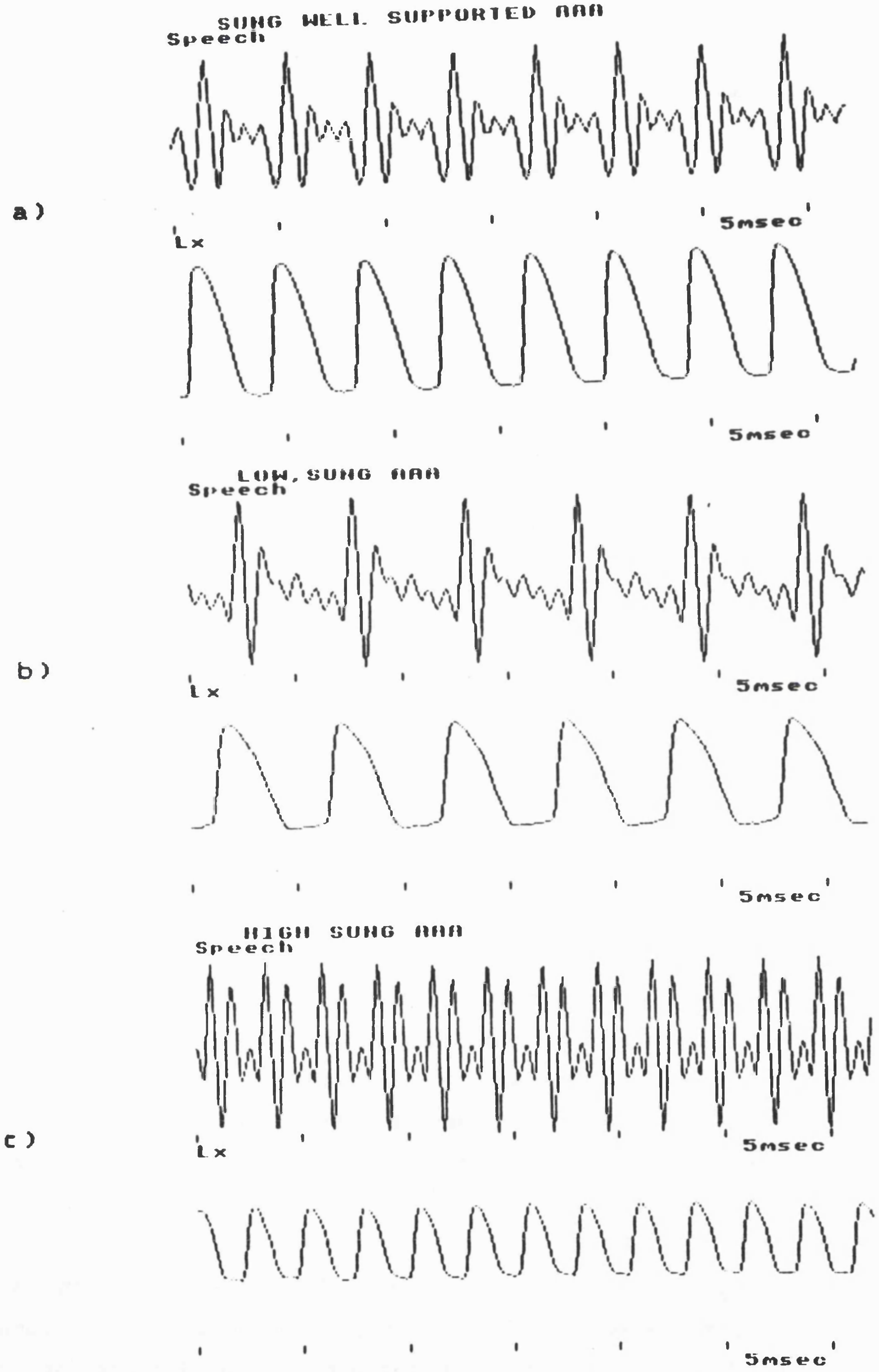
For the purpose of confirmation and verification of EGG waveforms, most researchers, recommend the use of additional means of observing glottal events.

v Combining Speech and Lx waveforms.

As shown above, there is some general agreement regarding the identification of the moment and speed of closure from EGG/ELG waveforms (Lecluse, 1975, Kelman, 1981, Dejonkere and Lebacq, 1985), just as there is a general lack of confidence about the identification of the moment, speed and duration of the opening of the vocal folds from the waveform (Lecluse, 1977, Pedersen, 1977, Dejonkere and Lebacq, 1985, Painter, 1988, Anastaplo and Karnell, 1988, Colton and Conture, 1990).

Fourcin (1982) quotes uniform Lx peaks, sharply defined Lx contact, long closure duration and regular contact periodicity as features of the Lx waveform that may be used to judge the efficiency of vocal fold closure. These are illustrated in Fig. 23 a-c where Lx waveforms are shown with the corresponding speech pressure waveforms for the vowel /α/ sustained by a normal female voice at comfortable volume and middle, low and high pitch.

Fig. 23 a-c shows each Lx cycle corresponding to one cycle in the Speech waveform. The latter is considerably more complex as a result of variation in the numbers and amplitudes of its constituent components and illustrates the damping that occurs during the glottal open phase (cf. Lx). This results from acoustic energy being



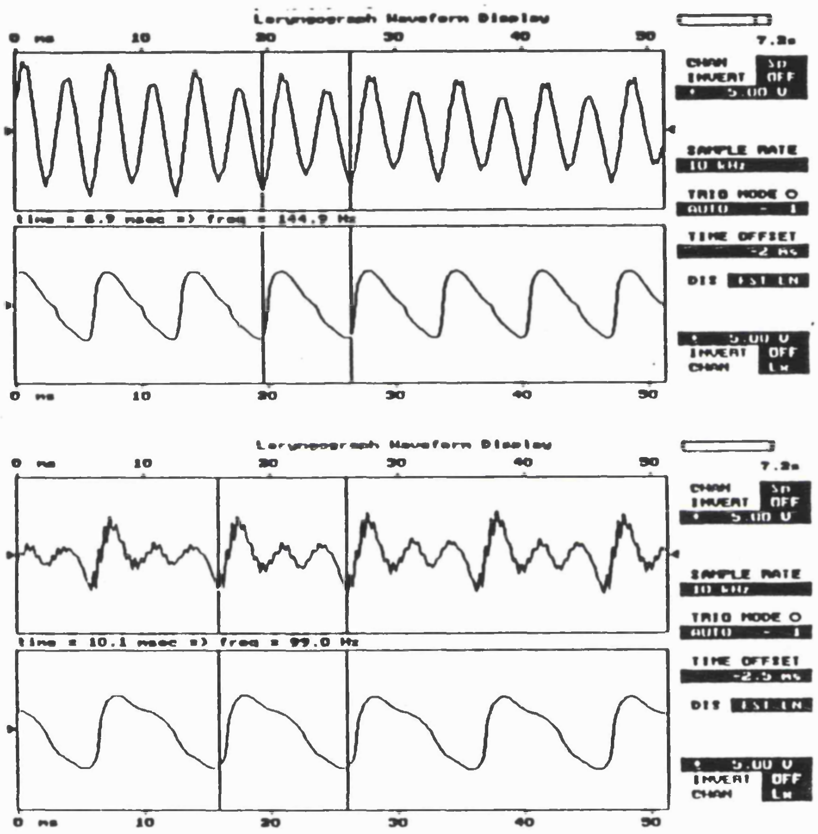
Lx and Speech pressure waveforms produced by normal female voice.

Figure 23

lost in the subglottis through the coupling of the sub- and suprgalottal spaces while the vocal folds are open (Fourcin, 1989).

The Lx waveform closed phase roughly corresponds to the major positive peaks in the Speech waveform. (Fig. 23 a-c). The moment of greatest acoustic excitation of the vocal tract closely follows the onset of vocal fold closure as evidenced by Lx. There is a slight delay in the onset of the major positive peaks in the Speech waveforms, compared to the moment of closure indicated on Lx. This is due to the delay in the acoustic signal reaching the microphone, relative to the registration of closure on Lx via the laryngeal electrodes. This problem is addressed in a recent version of the PCLx software, where an optional 'offset' facility enables the user to align the onset of Lx closure with the beginning of the major peak in the speech waveform (Fig. 24).

Combining Speech and Lx waveforms in this manner offers a means of verifying the onset, duration and efficiency of closure of the vocal folds as indicated on Lx. The speech waveform gives additional information about the response of the vocal tract to glottal excitation and broadly reflects supralaryngeal contributions to voice quality other than fundamental frequency.



time = 6.9 msec => freq = 144.9 Hz

CHAN Sp
INVERT OFF
± 5.00 V
SAMPLE RATE
10 kHz
TRIG MODE O
AUTO - 1
TIME OFFSET
-2 ms
DIS 1ST LN
± 5.00 V
INVERT OFF
CHAN Lx

time = 10.1 msec => freq = 99.0 Hz

CHAN Sp
INVERT OFF
± 5.00 V
SAMPLE RATE
10 kHz
TRIG MODE O
AUTO - 1
TIME OFFSET
-2.5 ms
DIS 1ST LN
± 5.00 V
INVERT OFF
CHAN Lx

Lx and Speech waveforms offset against each other in the PCLx 'Wave' program.

Figure 24

vi <u>Computer simulation of EGG waveforms.</u>

Detailed experimentation and modelling of EGG waveforms carried out by Childers and co-workers (1986 a,b, 1990) and by Titze (1984, 1989c, 1990) has verified many major features of EGG waveforms and contributed greatly to our understanding of the detail and complexity of vocal fold vibration.

The reasons for the difficulty in identifying timing details in the opening and open phase of the EGG waveform become obvious in the modeling experiments.

Childers, Hicks, Moore and Alsaka's (1986) 'elastic' two mass model of vocal fold vibration (Fig. 22, p. 132), is based on observations of ultrahigh speed laryngeal film. On simulating vocal fold contact area, (this time showing increasing contact in the more conventional negative going direction) the model takes into account the following:

1. The horizontal phase delay or time lag in vocal fold closure. (The 'zipperlike' closure and opening of the vocal folds).
2. The different angles on opening and closing, measured from a reference line through the mid-sagittal plane of the glottis, and the medial edge of the vocal folds (Fig.23). Variations in this angle parameter (θ), is shown to have a differential impact on segments of the EGG waveform.
3. The vertical phase delay between the upper and lower edges of the vocal folds. The lower edge always leads over the upper edge, both on opening and closing (Fig. 22).

Childers et al (1986) use this model to calculate the contact area and to predict the EGG waveform. Their findings show that varying the opening and closing angles and the vertical phase delay, the simulated EGG corresponds quite closely to measured EGG of male voices at modal pitch. They were able to model the effect on the EGG of a mucus strand bridging the vocal folds on opening, and also the effect of a polyp or nodule. Both these events would, in this model, tend to affect the rising (opening) portion of the waveform (Fig. 22, p. 132).

Childers et al (1986) point out the potentially distorting effect of a mucus strand on the EGG. It provides a highly conductive current path and may give a false impression of the amount of vertical fold contact. *'The moment the mucus strand breaks, there is an immediate flat portion of the EGG, indicating complete opening'*.

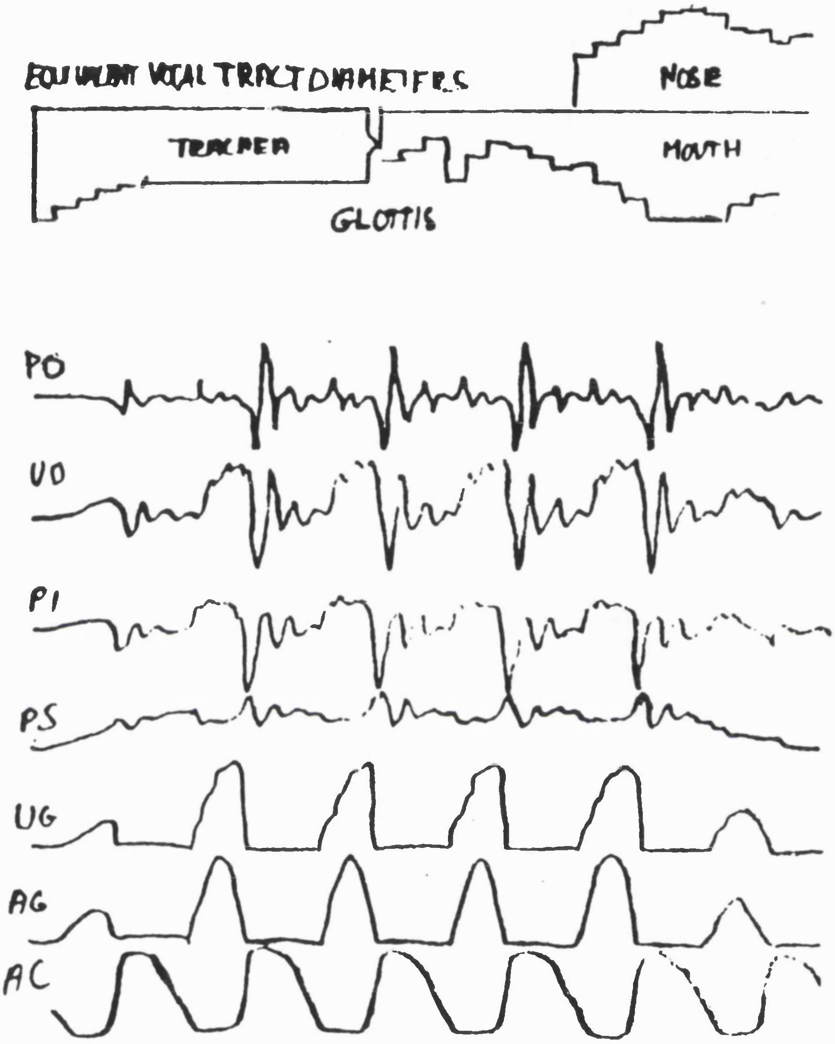
Nodules and polyps, suggest Childers et al (1986), would have the effect of allowing the folds to begin to peel apart and then possibly, when the point in opening has been reached where the mass interferes, give rise to a plateau in the rising EGG, indicationg a delay in opening, as the protruding mass is staying in contact longer with the opposite fold.

The simulations above seem to lend credence to observations made of horizontal and vertical phase differences and different angles of opening and closing of the vocal folds during vibration. The waveforms replicate quite closely measured ones under certain phonatory conditions.

Titze (1984) devised a computer model to predict the effect on glottographic waveforms, of varying a number of parameters known to be important in vocal fold vibration e.g. vocal fold length, thickness, depth, surface contours, adduction, vibrational amplitude and phase differences in movement of inferior and superior edges of the vocal folds.

With this model, he hoped to *'take a step towards relating phonatory acoustics to a complete kinetic description of air and tissue movement.'*

The model allowed the controlled variation of aerodynamic parameters to do with subglottal pressure, glottal resistance, glottal inertance, vocal tract inertance and maximum or average flow (Fig. 25). It was based on earlier studies where Titze and Talkin (1981) showed that the most important factors in the establishment of self-oscillation of the vocal folds were:



Simulation of glottal waveforms and vocal tract waveforms for the vowel [α]. Vocal tract configuration for trachea, supraglottal system and nasal cavity is also shown. (From: Titze, 1984)

PO = Pressure out at mouth   UO = volume flow out
PI = Pressure into vocal tract   PS = Subglottal pressure
UG = Glottal flow   AG = Glottal area   AC = Glottal contact area

Figure 25

1. proper adduction of the arytenoid processes.
2. vertical pre-phonatory shape of the glottis (converging, rectangular or diverging).
3. a phase difference between movement of upper and lower portions of the vocal folds.
4. loading conditions on the vocal folds by the vocal tract.

In his model Titze (1984) describes each of these adjustments with a specific parameter. Fig. 25 shows the simulated glottal and vocal tract waveforms after entering a number of parameters into the computer which relate to vibrational amplitude, fundamental frequency, abduction, shape and phase quotients and lung pressure. Titze points out a number of important features:

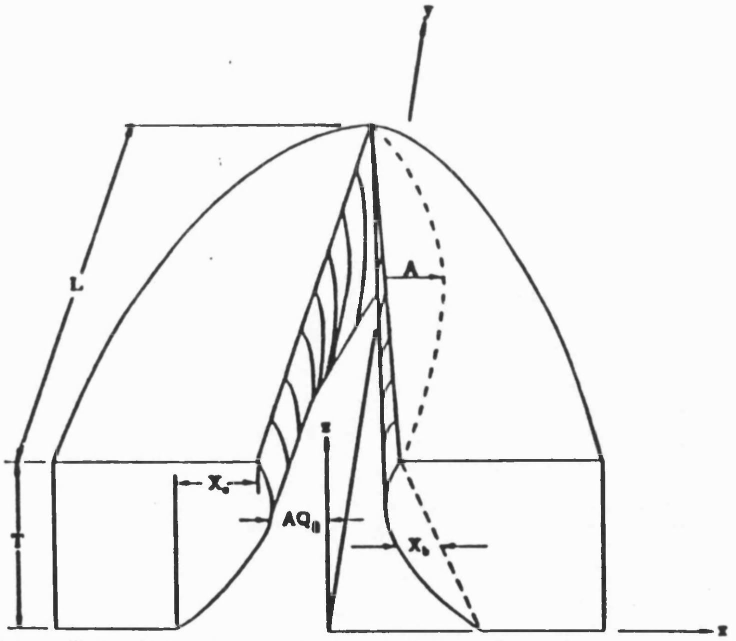
> *As a general rule, contact area (AC) is skewed to the left, glottal area (AG) is nearly symmetric and volume velocity (UG) is skewed to the right. The reverse symmetry properties of contact area and volume velocity are based on very different physical properties. The contact area is skewed on the basis of a convergent glottal geometry and vertical phase difference in movement........The volume velocity....is skewed on the basis of vocal tract loading.'* (Titze, 1984)

Another important observation was that during rapidly changing portions of the glottal area (AG) and flow (UG), there is a flat portion in the contact area waveform (AC) and vice versa (Fig. 25). Titze (1984) shows how a combination of waveforms provides more information about 'glottal kinematics' than any one waveform on its own. He concluded that the model brings us closer to the possibility of predicting glottal configurations from given EGG waveforms, rather than from direct observation. He adds however, that the difficulties encounterd with the model are due to factors that it does not take into account, for instance the different amplitudes of the top and bottom edges of the folds, parameters regarding the curvature of the medial surfaces of the vocal folds and, also, the posterior chink between the arytenoid cartilages, would need to be handled independently of the vocal processes.

Both Titze (1984) and Childers et al (1986) admit to limitations in their models, however, due to factors, which had not been taken into account. This restricted the range of possible glottal vibratory configurations that could be predicted.

In attempts at simulating more closely vocal fold vibration by means of computer models, Titze (1989, 1990) added yet more time varying parameters e.g. to do with horizontal and vertical phase differences during opening and closure; degrees of convergence between the approximating folds; medial surface bulging; variation in glottal halfwidth at the vocal processes and amplitude of vibration (Fig. 26 and 27).

By 1989 he presented what he describes as his 'Four parameter model of the glottis' (Fig. 26). By varying four parameters which take into account the three dimensionality of the glottis as well as adding the 'medial surface bulging' parameter, Titze arrived at simulated waveforms which closely resemble those produced by live subjects in a clinical setting (Fig. 27).



A model of the vocal folds, with the abduction quotient Q_0, the amplitude of vibration A, the net convergence x_c, and the medial surface bulging x_b controlling the shape of the glottis. L is the length, and T is the thickness of the vocal folds.

Four parameter model of the vocal folds.
(From: Titze, 1990)

Figure 26

Nominal waveform resembling male EGG

(a) Nominal ($Q_a = 0.3$, $Q_p = 0.25$, $x_a = 0.4$ cm, $x_b = 0.2$ cm)

No surface bulging,
Resembles female EGG

(b) No surface bulging ($x_b = 0$)

More abduction, cf. Breathy voic

(c) More abduction ($Q_a = 2.0$)

Less abduction, cf. Pressed voi

(d) Less abduction ($Q_a = -1.0$)

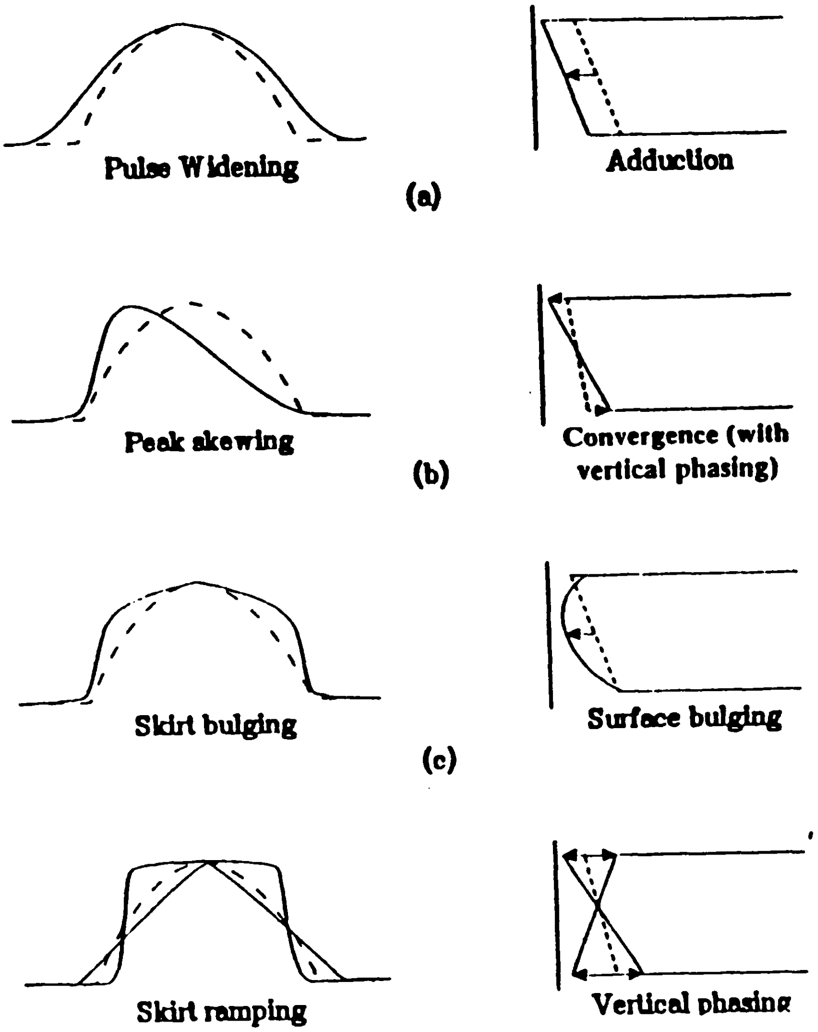
(e) Less convergence, no bulging ($x_a = 0.2$, $x_b = 0$)

(f) No convergence ($x_a = 0$)

Simulated EGG waveforms corresponding to 'live' EGG. (From: Titze, 1990)

Combinations of waveform features as simulated by Titze's model in Fig. 26.
a) nominal configuration b-f) one or two parameters altered, with all other parameters remaining (or returning) to the nominal state.

Figure 27



EGG waveform features (left) and corresponding vocal fold frontal section (right). (From: Titze, 1990)

a) Pulse widening resulting from increased adduction.
b) Peak skewing resulting from convergence and vertical phasing
c) Skirt bulging resulting from medial surface bulging
d) Skirt ramping resulting from increased vertical phasing.
Arrows indicate tissue displacements of the right vocal fold.

Figure 28

Simulating EGG waveforms Titze (1990) emphasises that skewness to the left of the waveform (Fig. 28 b) requires a convergent prephonatory glottis and a vertical phase difference *'..when the glottis closes, the vocal fold must square up to produce an abrupt increase in contact over the full thickness. On the other hand when the glottis opens, the vocal fold must become more wedge shaped to produce a gradual release of contact.'*.

Fig. 27 e shows a waveform illustrating how skewness is affected and increased by reducing convergence and simulating a highly wedge shaped vocal fold on opening, producing a *'nearly squared up vocal fold at closure'*

Titze (1990) further explains increased triangularity or 'ramping' of the EGG waveform (Fig. 28 d) as a result of increased vertical phasing and illustrates this by a waveform that he suggests is similar to one produced in a 'breathy voice' (Fig. 27 c).

This is borne out by Orlikoff (1991) who finds that 'ramping' of the EGG waveform appears to vary with the amount of energy in the acoustic signal in a 'live' subject.

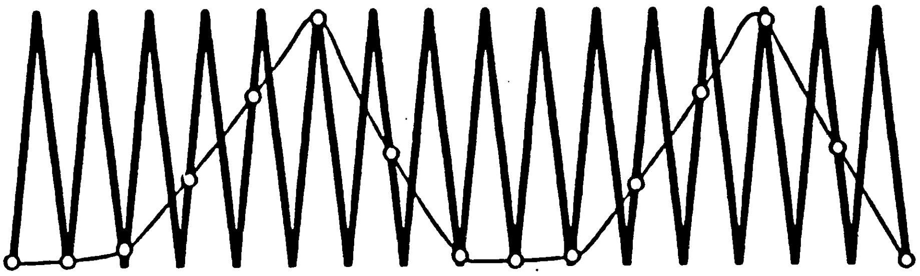
Fig. 27 d illustrates a simulated waveform shape that is often observed in 'Pressed' voice. The result is a waveform with considerably increased pulse width and a much reduced open phase.

Titze finally also illustrates simulated waveforms that he describes as typical for male and female EGG signals (Fig. 27 a and b). The difference lies in the prescence of 'medial surface bulging' in the case of the male waveform (Fig. 27 a). This results in a more sudden release of contact on opening than in the female waveform (Fig. 27 b) where opening shows a 'flatter and more constant decreasing slope' because of the lack of medial surface bulging (Titze, 1990).

vii Quantification of the open and closed phases of the vibration cycle.

Detailed observation of vocal fold vibration became possible with the advent of stroboscopy and high-speed filming, where the stages of vocal fold vibration were clearly visible.

The stroboscopic image is composed of successive vocal fold vibrations, filmed at different points of each cycle and combined into the final image observed (Fig. 29). Stroboscopy is therefore not suitable for actual measurement purposes, but has been very crucial in the discovery of the existence of the mucosal wave and the vertical dimension of vocal fold vibration. However, as described by Anastaplo and Karnell (1988), synchronised stroboscopy and EGG was used for determination of what causes the 'knee' in the opening phase of some EGG waveforms.



Principle of stroboscopy - When a rapidly moving object (represented by the high frequency waveform) is strobed by flashes at a lower frequency (curve with open circles) the rapidly moving object appears to move more slowly. (From: Colton and Casper, 1990)

Figure 29

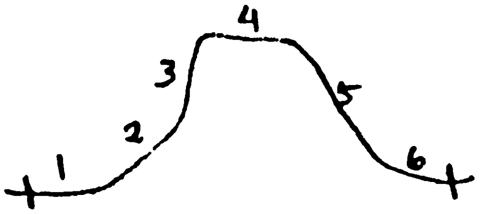
Lecluse, Brocaar and Vershuure (1975) used stroboscopy to verify different EGG and ELG waveforms on an oscilloscope screen. They were also interested in finding the optimum placement of the electrodes (Fig. 18, p. 118). They confirmed a close relationship between what could be observed of the vibratory cycle in terms of degrees of contact, maximum contact, opening and maximum open phase in the strobscopic images, and points on the EGG waveform (Fig. 19, p. 118). They emphasised, however, that the exact moment of vocal fold closure was impossible to judge from the waveforms. Similarly, the moment of complete opening was impossible to define.

The folds closed slightly before the peak of the waveform was reached and stayed closed for a short while after the maximum had been reached, as a result of the changing vertical degree of closure and elastic tissue compression which causes a rounded peak of the waveform. This has since been confirmed by Childers and Krishnamurthy (1985) (Fig. 21, p. 129).

Lecluse et al (1975) also studied ELG waveforms for different vowels, as they had been found to be unexpectedly different. They found, however, very similar waveforms produced by different vowels and concluded that vocal fold vibration must also be very similar.

Painter (1988) describes EGG waveforms as sequences of 'stages' (Fig. 30 and 31). His is a formalised but still qualitative description of EGG. His addition of the percentage closed phase allows quantification of individual waveforms. The duration of the closed phase is approximated by including *'half of the duration of stage 3, all of stage 4 and half of stage 5'*. According to Painter all waveforms will have a closing stage '3' and an opening stage '5'. Using his model Painter arrived at 16 possible 'waveforms types' by adding or subtracting each of stages 1, 2, 4 and 6 (Fig. 31). Some of these have been experimentally observed.

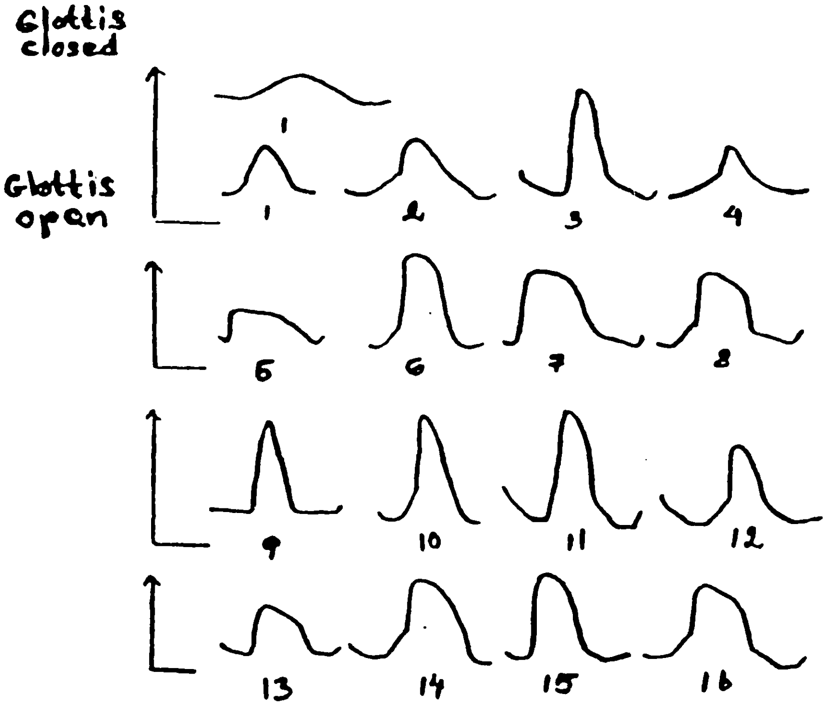
The value of the closed phase percentage would reflect waveshapes to some extent but looking at the 16 possible waveform types in Fig. 31, it seems some of these calculations may be extremely arbitrary, especially the basis on which decisions are made of where stages 3 and 5 begin and end.



EGG waveform stages (From: Painter, 1988)

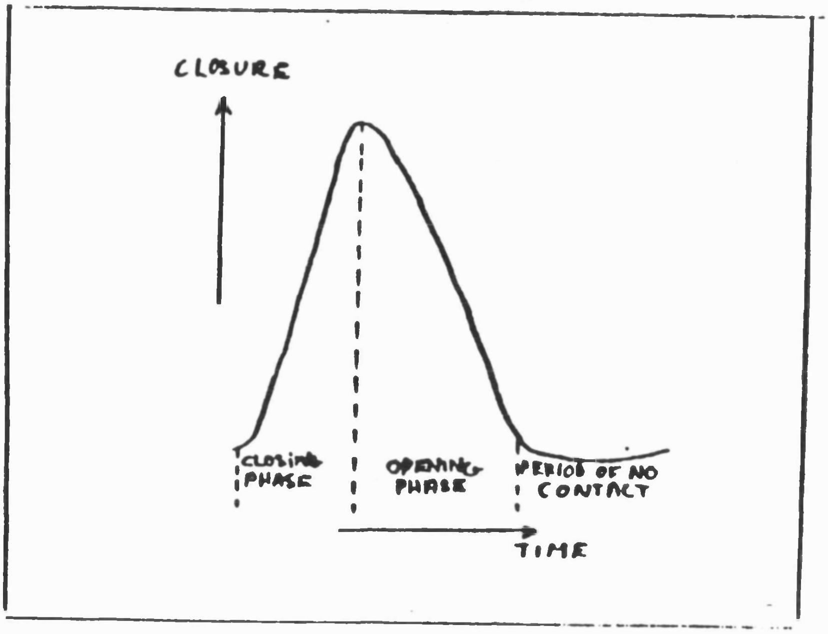
1. Fully open
2. Initial closing
3. Further closing
4. Fully closed
5. Initial opening
6. Further opening

Figure 30



EGG waveform types according to Painter (1988)

Figure 31



Idealised Lx waveform. (From: Kelman, 1981)

Figure 32

Kelman (1981) used the Fourcin Laryngograph (Fourcin and Abberton, 1971, Fourcin 1974, 1979, 1981) to record clinically normal subjects phonating the vowels [i], [e], [ɛ], [ɑ], [o] and [u] at habitual pitch for 5-10 seconds. He measured on the Lx waveform the Open Time, Closing Time, and Opening Time (Fig. 32).The most important part of the vocal fold vibratory cycle was the closing time, which constituted the shortest phase of the cycle and corresponded to the moment of maximum excitation of the vocal tract as suggested by Fourcin 1974).

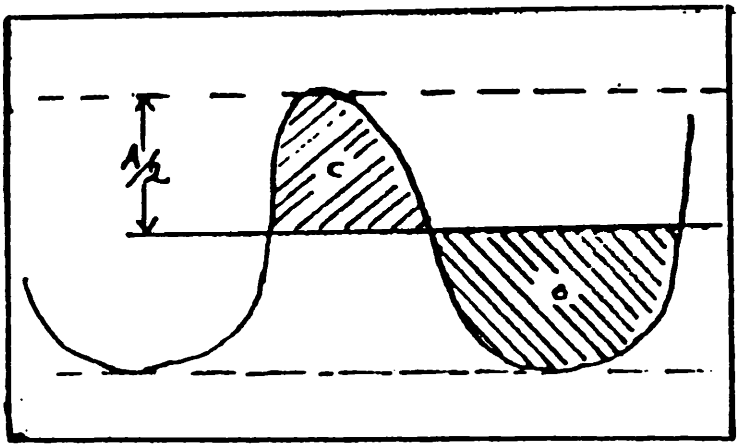
Kelman also measured the proportion of the total vibratory cycle taken up by the closing time and found it occupied between 4 and 9% of the period in male subjects and between 10 and 19 % of the period in female subjects. For male subjects the open time was always shorter than the opening time (Fig. 32), but in females these were roughtly of the same duration. This is illustrated in Fig 23 a-c, where the Lx waveforms for sustained [ɑ] produced by a normal female voice at mid, low and high pitch, is displayed with the corresponding acoustic waveform (Speech).

Kelman (1981) found no significant differences between the closing time, opening time and open time for the different vowels, and the values were fairly constant for each subject.

There was no relationship between fundamental frequency (Fo) and the closing time, but variations in Fo correlated with variation in opening and open time, which are shorter at higher fundamental frequencies. This is clearly illustrated in Fig. 23 (p. 135) where the duration of the open phase is considerably longer at low [α] than at mid or high [α] (Fig. 23 a-c).

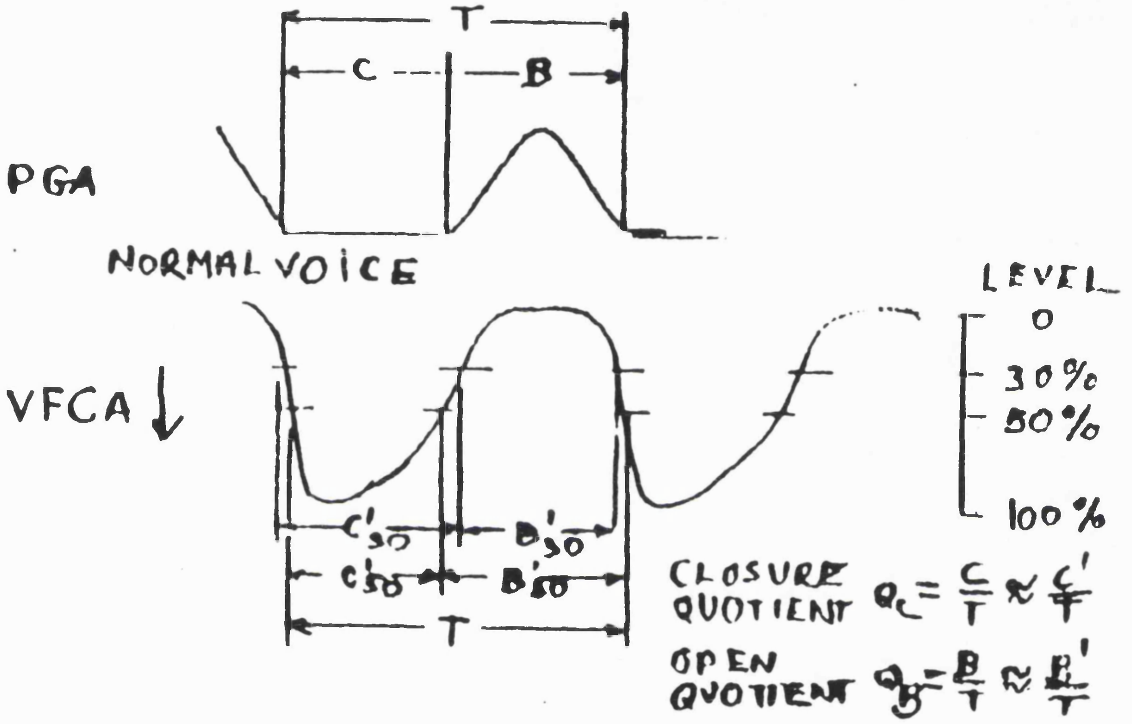
Colton and Conture (1990) measure the 'closing time' or 'closing slope' of the EGG waveform (cf. Kelman, 1981). They measure the distance from the start of the closing phase to the point of maximum contact (Fig. 32). They choose this measure as it is relying on features of the EGG waveform that are clearly evident in the waveforms of patients whose waveforms are possible to analyse. In their sample of 100 subjects, they found only about 15% of the waveforms impossible to analyse.

Dejonckere and Lebacq (1985) attempted to differentiate Lx waveforms produced by normal young females phonating on [α] at 220 Hz and 70 dB, from the waveforms produced by age matched females with vocal nodules. By drawing a line through the waveforms at the 50% level of maximum closure they derived a measure, which they call the S-quotient (Fig. 33). They calculated the area under the waveforms at closure and divided this by the area over the waveform during opening. An S-quotient (= C/O) was calculated. They deemed this quotient as *'somehow related to the importance of the contact phase relative to the open phase'*, and considered it as combining information about both the relative duration of closure and the surface of contat in each cycle. The S-quotient was significantly smaller, i.e. the open phase was longer, among subjects with nodules.



Schematic EGG waveform with definition of the C(losed) and O(pen) areas. Illustrating how the S-quotient is calculated as C/O. A/2 = half of the maximum amplitude. Maximum conductance above. (From: DeJonckere and Lebacq, 1985).

Figure 33



Inverted vocal fold contact area (VFCA) and Photoglottographic waveforms (PGA) for normal voice showing influence of different criterion levels chosen for calculation of open and closed quotients.
(From: Rothenberg and Mahshie, 1988)

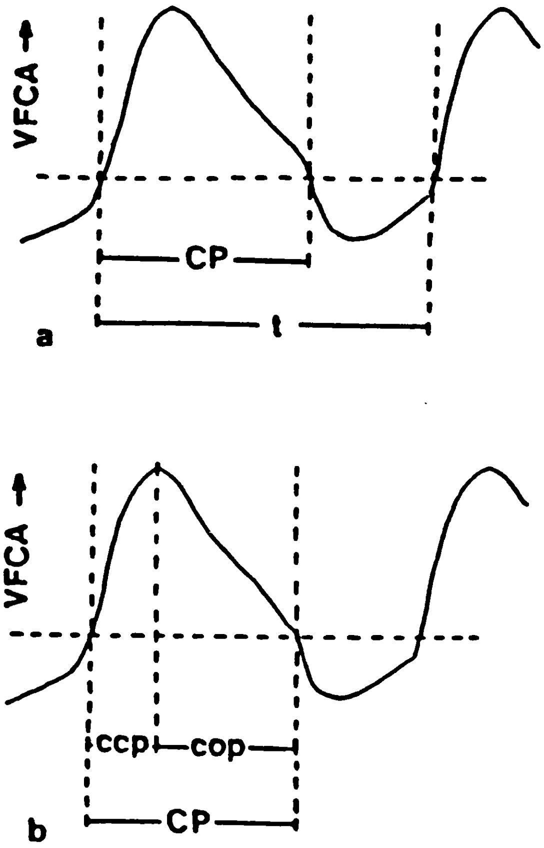
Figure 34

Dejockere and Lebacq (1985) chose to measure areas under the EGG waveforms to avoid the arbitrariness involved in trying to determine where the open phase of the glottal cycle begins. The level chosen i.e. the midpoint between the maximum closure and opening, corresponds roughly to Rothenberg and Mashie's (1988) 50 % 'criterion level' (Fig. 34). The duration of the closed or open phase is considerably simpler to measure than the areas under a curve, therefore, the glottal 'duty cycle' as defined by Rothenberg and Mashie and described below, seems the more practical parameter to use. Dejonckere and Lebaq, however, supply some interesting data to confirm the usefulness of such a measure.

Rothenberg and Mashie (1988) describe a method where a linear phase high pass filtered EGG waveform is used for calculation of 'glottal duty cycle'. The duty cycle is defined as consisting of either the closed or the open quotient, calculated as shown in Fig. 34. The waveform is high pass filtered to remove low frequency noise and movement artifacts. They measure the duration of the glottal pulse from the points at which the EGG waveform crosses a predetermined criterion level (CL). The maximum and minimum amplitudes of the waveform to be measured are used to define the extremes between which a percentage level is set to either 50 % or 30 % (measured from most open glottis) (Fig. 34).

They conclude that ' *given a reasonably strong signal, the output of an EGG can be used for monitoring the degree of abduction or adduction of the vocal folds at least diuring voiced speech without strong supraglottal constriction.* ' Rothenberg and Mashie (1988).

Orlikoff (1991) used the Fourcin Laryngograph and the same orientation in space of the Lx waveform, and Rothenberg and Mashie's (1988) suggestion of baseline crossings at different level to mark the EGG **Contact Phase (CP)**(Fig. 35). He chose the 25% level from the most open phase from which to calculate what he calls the **Contact quotient (CQ)** and a **Contact Index (CI)**. He describes CQ as related to the degree of approximation and relative compression of the vocal



$$CQ = \frac{CP}{t}$$

$$CI = \frac{ccp - cop}{CP}$$

Schematic representation of electroglottographic landmarks and intervals used to derive a) Contact quotient (CQ) and b) Contact Index (CI). The horizontal dashed line represents 25% of peak to peak EGG amplitude; VFCA = vocal fold contact area; CP = contact phase; ccp = contact closing phase, cop = contact opening phase. (From: Orlikoff, 1991)

Figure 35

folds in the horizontal plane. It is calculated as CQ = Contact Phase/ Period time.

Orlikof avoids talking about an open phase, no contact or loss of contact as he feels that without other simultaneous measures, it is not possible to draw conclusions about this from the Lx waveform. He prefers to specify the trough in the waveform as showing the duration of 'minimal vocal fold contact'. therefore, in calculating his Contact Index, he does not describe 'closing' or 'opening' time (cf. Kelman, 1981, Colton and Conture, 1991), but a Contact Index = contact closing-contact opening/contact phase (CI = ccp-cop/cp) (Fig. 35).

His subjects were ten males with normal larynges and voices, aged 26-37. Each subject sustained three samples of the vowel [ɑ] in comfortable modal voice at three different intensity levles, soft, moderate and loud. The average CQ for all phonations was 0.57 +/- 0.07 ranging from 0.37 to 0.68. There was a significant difference and increase in the mean CQ with increasing loudness. This confirms that increased intensity of phonation is achieved by increased degrees of vocal fold approximation.

He suggests the Contact Index reflects vocal fold tonus and may be sensitive to mucosal dynamics in the vertical plane, as described by Hirano (1974), Rothenberg (1981) and Titze and Talkin (1979).

The mean CI for all phonations was -0.52 +/- 0.08 ranging from -0.68 to -0.37. The 'contact closing' duration represented about 24 % of the entire contact phase, which was consist ent with observations of increased skewing of the EGG with increasing intensity (Fourcin 1981, Kelman, 1981) The CI was significantly different between soft and moderate intensity conditions but not between moderate and loud intensity.

Orlikoff (1991) concludes:

> *'Sound pressure and EGG data indicate that both the slope of increasing EGG contact and EGG duty cycle were significantly related to the amplitude of the acoustic signal. These results suggest that quantitative electroglottography may provide powerful insights into the control and regulation of normal phonation and into the detction and characterisation of pathology.'*

Howard, Lindsey and Palmer (1991) calculated Closed Quotients of Lx waveforms by measuring the end of the Closed Phase (CP)*from the point where the negative Lx waveform crosses a fixed ratio 7:3 of the current cycle's amplitude'* (CQ=CP/Tx). They found that untrained singers displayed falling CQs with increasing pitch, whereas trained singers increased their CQ with increase in pitch.

In another study Howard Lindsey and Allen (1990) showed how trained singers maintained high CQ values (> 50%) during reading aloud. Untrained singers maintained consistently lower CQ than trained singers throughout both reading and singing. In singing, some even showed lower CQ than they did in reading. The studies show how the CQ calculated on Lx waveforms can be utilised to evaluate vocal efficiency in normal and pathological speakers. This is also described by Fourcin (1989).

Orlikoff (1991) summarises the usefulness of EGG by saying *EGG can provide a relatively simple and non-invasive means of assessing the vibratory behaviour of the vocal folds, providing powerful insights into the regulation, maintenence and quality of phonation.'*

In this study, Fourcin's Electrolaryngograph is used both for observation of changes in the Lx waveform as a function of time after radiotherapy and as a tool in measuring voice fundamental frequency and vocal fold regularity of vibration after radiotherapy for glottic carcinoma.

viii The effect of glottal pathology on the EGG waveform.

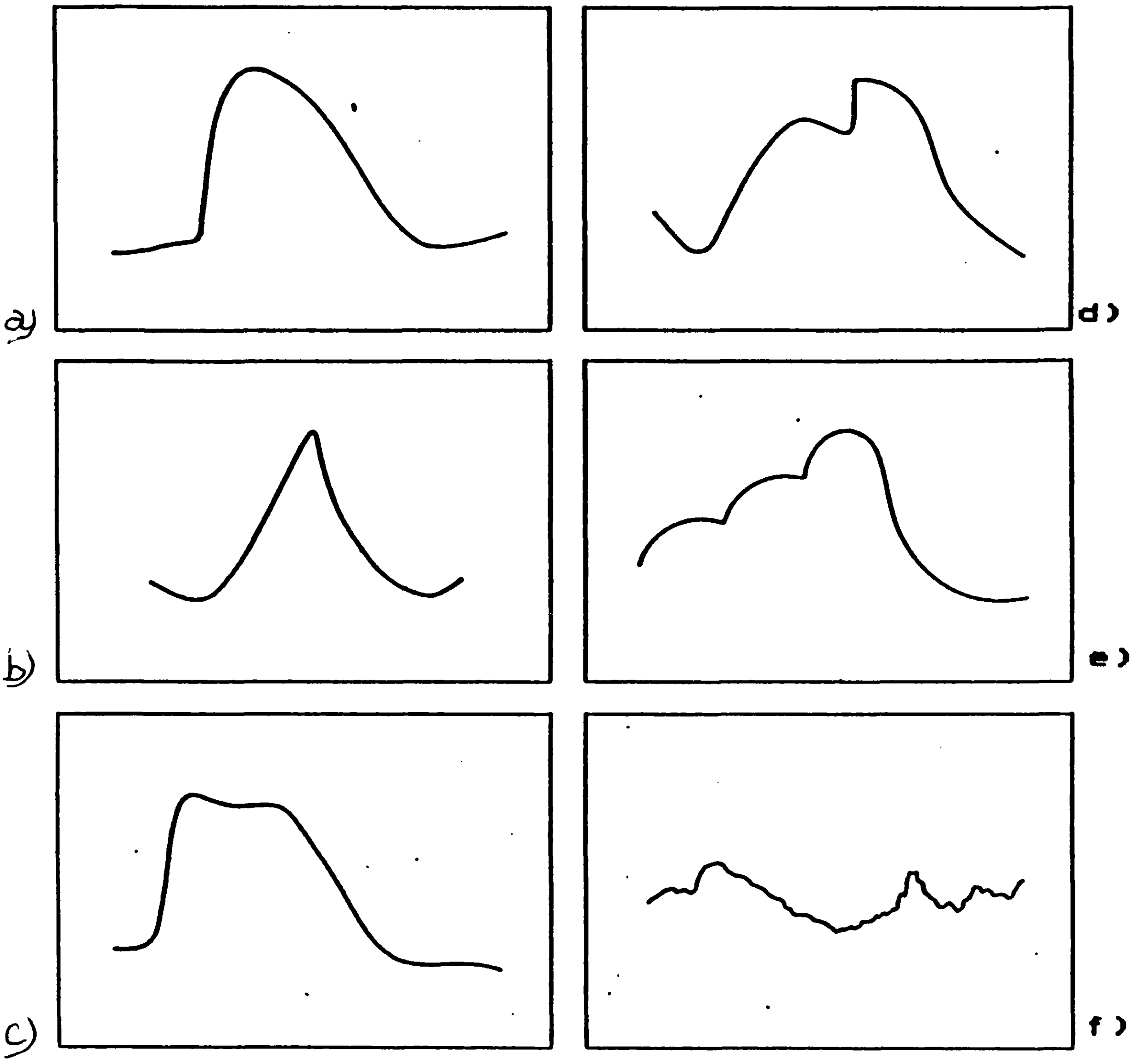
As the closing and closed phases of the vibration cycle are associated with the moment of maximum excitation of the vocal tract, Abberton and Fourcin (1984) suggest that disorders of phonation are likely to be associated with vocal fold closure problems. Asymmetry or abnormal tension of the vocal folds may affect the Lx waveform (Fourcin 1981).

The first observer of abnormal EGG waveforms in dysphonic patients was Van Michel (1967), who noted a 'notch' in the closing portion of the wave in patients with vocal nodules. EGG/Lx waveforms have been used for assessment of pathological voice problems in a number of studies since (Wechsler, 1975, Fourcin and Abberton, 1976, Carlson, 1986, 1988 a, b, c, 1993 a and b, Beckford, Mayo, Wilkinson and Tierney, 1990, Motta, Cesari, Iengo and Motta, 1990).

Beckford et al (1990) studied Lx waveforms in two groups of women pre- and post - endotracheal intubation. They found what they describe as a 'Broader Peak' or longer closed phase after surgery as a result of reactive vocal fold oedema. This seems to confirm Titze's (1990) prediction of 'pulse widening' as a result of increased adduction in his modeling experiments (Fig. 28 a, p. 144).

Motta et al (1990) report the most extensive study of 432 dysphonic patients' EGG waveforms compared to those of 50 'normal' speakers'. Subjects were recorded sustaining [i] at comfortable pitch and loudness. Table 4 summarises the most common types of waveforms. The table shows the breakdown of frequency of observation of certain waveform features and the diagnoses that most commonly displayed these features in Motta et al's study. Examples of waveform types are shown in Fig. 36.

They also found that patients after surgery for removal of organic lesions sometimes did not produce normal looking waveforms. This was achieved only after a period of voice therapy.



EGG waveform types produced by dysphonic patients. (From: Motta et al, 1990)

a) Normal voice,
b) Hypokinetic dysphonia
c) Hyperkinetic dysphonia
d) Vocal nodules
e) Vocal fold polyp
f) Reinke's oedema

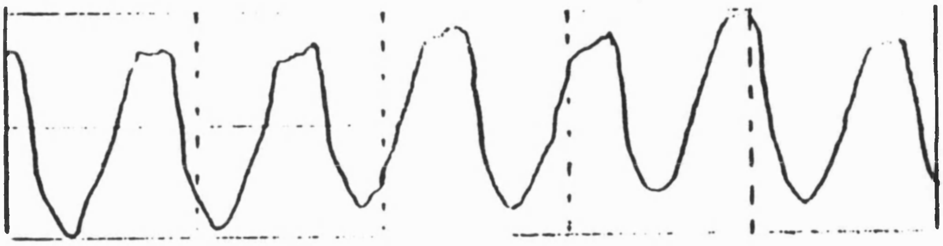
Figure 36

Table 4

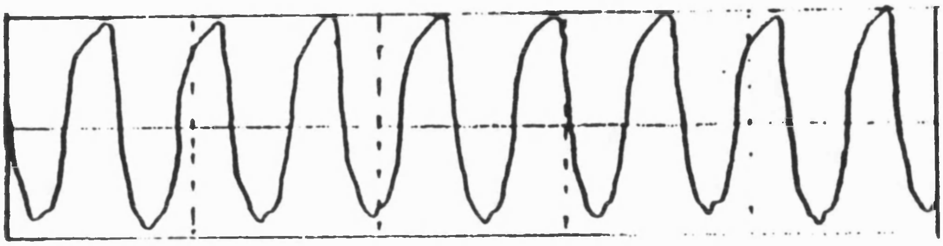
Diagnosis	EGG waveform features	% Subjects displaying feature
No dysphonia (N=50)	Curved peak and more or less uniform inclination of the ascending and descending wave. (Fig.36 a)	
Functional hypokinetic (N=66)	Particularly sharp peak and reduced amplitude (Fig.36 b)	93 % (7 % no change in EGG)
Functional hyperkinetic (N=85)	Plateau like waveform (Fig.36 c)	95 % (5 % no change in EGG)
Vocal nodules (N=92)	A single notch in the closing portion of the waveform. (Fig.36 d)	72 % (28 % nearly normal EGG)
Vocal Polyps (N=86)	Single notch (Fig. 36 d) Double notch (Fig.36 e) in closing phase	25 % 68 % (7 % almost normal EGG)
Reinke's oedema (N=53)	Single notch (Fig. 36 d) Double notch (Fig.36 e) Irregular wave (Fig. 36 f)	24 % 72 % 4 %

(From Motta et al, 1990)

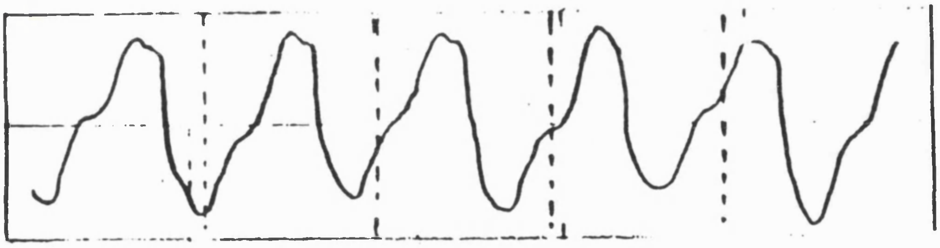
Childers and Krishnamurthy (1985) conclude their paper with illustrations of a variety of waveforms taken from normal and pathological speakers. They agree with Fourcin (1979) that EGG shows changes in the closing and closed phase in some pathological speakers, but they find more abnormalities in the opening phase for some. Looking at their examples of abnormal EGG due to cancerous lesions of the vocal folds (Fig. 37 a-d), and in any case of laryngeal pathology, one must remember the possible contribution of individual compensatory glottal and respiratory adjustments in response to changes in glottal mass, often unilateral, and vibratory pattern, as in the case of unilateral vocal fold palsy (Fig. 38 c).



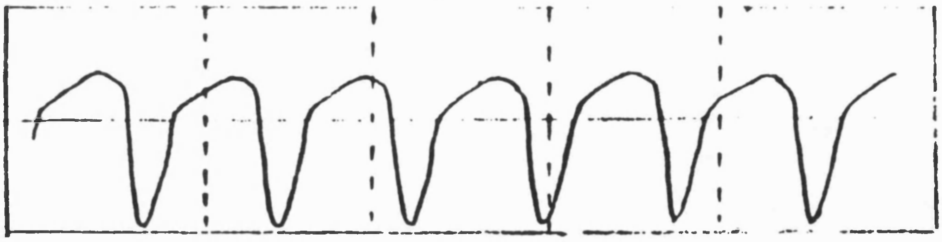
a) Subject 'CJ' Male, aged 66, Diagnosis: 'Cancer of the Throat', 'Hoarse'.



b) Subject 'CH', Male, aged 51, Diagnosis: 'Cancer of the left vocal fold, possibly fixed', 'Moderate hoarseness, complete closure'.



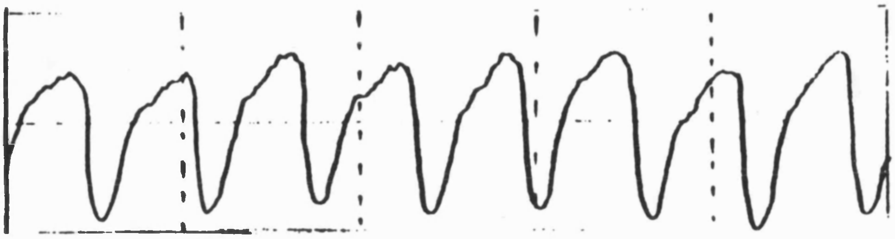
c) Subject 'HG' Male, aged 54, Diagnosis: 'Extensive Cancer, one partially fixed, vocal cord', 'Moderate hoarseness.'



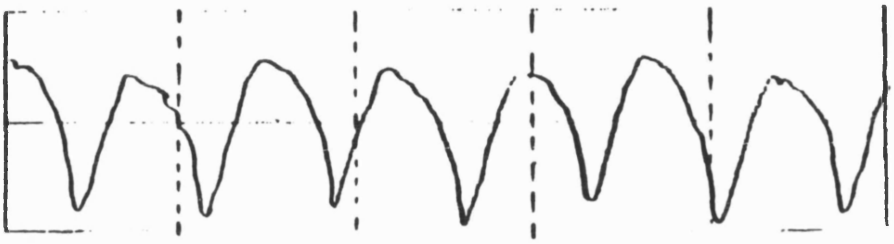
d) Subject 'WC' Male, aged 52, Diagnosis: 'Cancer' Voice weak, moderate hoarseness

EGG waveforms produced by speakers with malignant vocal fold lesions.
(From; Childers and Krishnamurthy, 1985)

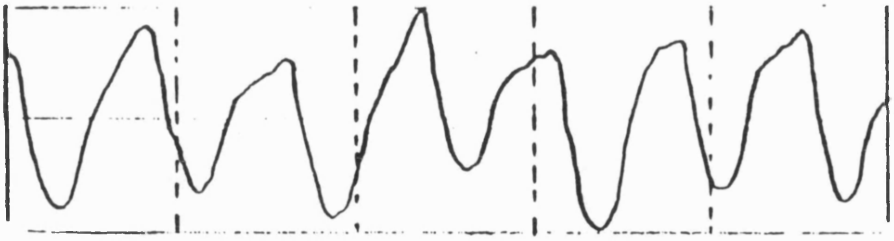
Figure 37



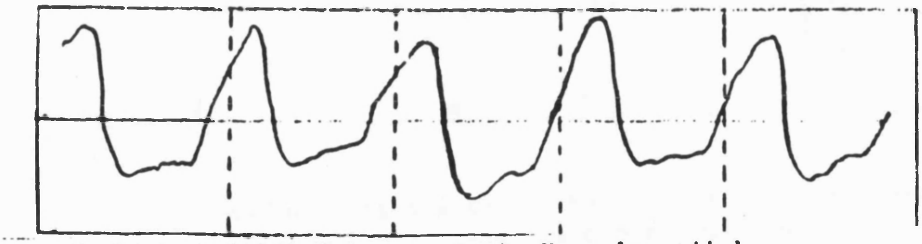
a) Subject 'RCB', Male, aged 63, Voice tremulous, high pitched with narrow range, no hoarseness, weak.



b) Subject 'JB' Male, aged 45, normal voice quality, pitch a little high, functional, non-organic, weak.



c) Subject 'RHM' Male, aged 62, Unilateral paralysis of the left vocal cord, cord in paramedian position.



d) Subject 'JCF' Male, aged 56, Very low pitch occasional aphonia.

EGG waveforms produced by speakers with benign vocal fold lesions. (From: Childers and Krishnamurthy, 1985)

Figure 38

Changes in the EGG may not only reflect interference in closure or opening due to interference by a mass or just a single vocal fold vibrating (Fig. 37 a,b,c,d Fig. 38 c), but may reflect habitual, abnormal patterns of vocal fold contact with or without interference of organic lesions. As Motta et al (1990) also found, however, there are a number of pathological voices that do not produce any significant abnormality in the EGG waveform (Table 4). The site and extent of the lesion must determine whether it affects the vocal fold contact pattern. It may only show up on production of certain pitches in sustained phonation or in running speech (Carlson, 1993 a).

The perceptual descriptions given by Childers and Krishnamurthy (1985) of the various voice qualities represented in Fig. 37 and 38 are very impressionistic and inadequate for the purpose of relating the waveforms to voice quality. All the cases of laryngeal cancer in Fig. 37 are described as 'hoarse'. In a and c there is evidence of amplitude perturbation and it would be interesting to know if their voice quality was worse (Wendahl, 1963) than the 'hoarseness' in b and d where there is no evidence of such gross perturbation.

The latter, d, would be likely to have a very 'breathy' quality reflected in the very short duration closure and long open phase and also low amplitude, possibly as a result of lack of complete closure at any point of the cycle. This may be reflected in the description of the voice as 'weak'. A similar situation is shown in the 'weak voice' in non-malignant examples in Fig. 38 a and b.

The voice in fig. 37 b, also described as 'moderately hoarse' but also with the observation of 'complete closure', shows a higher amplitude waveform as a result. There is a recurring slight irregularity in the beginning of opening, which may give rise to a suspicion of interference by the tumour along the inferior edge of the vocal folds.

Information is needed of whether the perceptual description of voice quality in their examples is based on sustained phonation producing these waveforms, or is based on perceived voice quality in continuous speech. Many dysphonic patients who are able to sustain relatively normal phonation on a vowel, produce very abnormal phonation and EGGpatterns during running speech.

Childers' and Krishnamurthy (1985) describe a number of their subjects producing 'double periodicity' in their waveforms, possibly due to the difference in the vibrational mass and contact between the vocal folds (Fig. 37 c) in cases of mass lesions. It is also evident in the non pathological cases of dysphonia in Fig. 38 b, and d, which are very similar to Motta et al's (1990) Hypo- and Hyperfunctional dysphonics' (Fig. 36 b and c). In the hyperfunctional dysphonia the voice is described as having very low pitch and occasionally being aphonic (Fig. 38 d).

Fig. 36 c and 38 d are reminiscent of Lx waveforms of subjects where there is an indication on indirect laryngoscopy, of ventricular band involvement e.g. one vocal fold approximating a bulky ventricular fold on the opposite side or where damping of vocal fold vibration is caused by ventricular bands. This type of phonation would also be of extremely low pitch. Some of our irradiated speakers produce such waveforms.

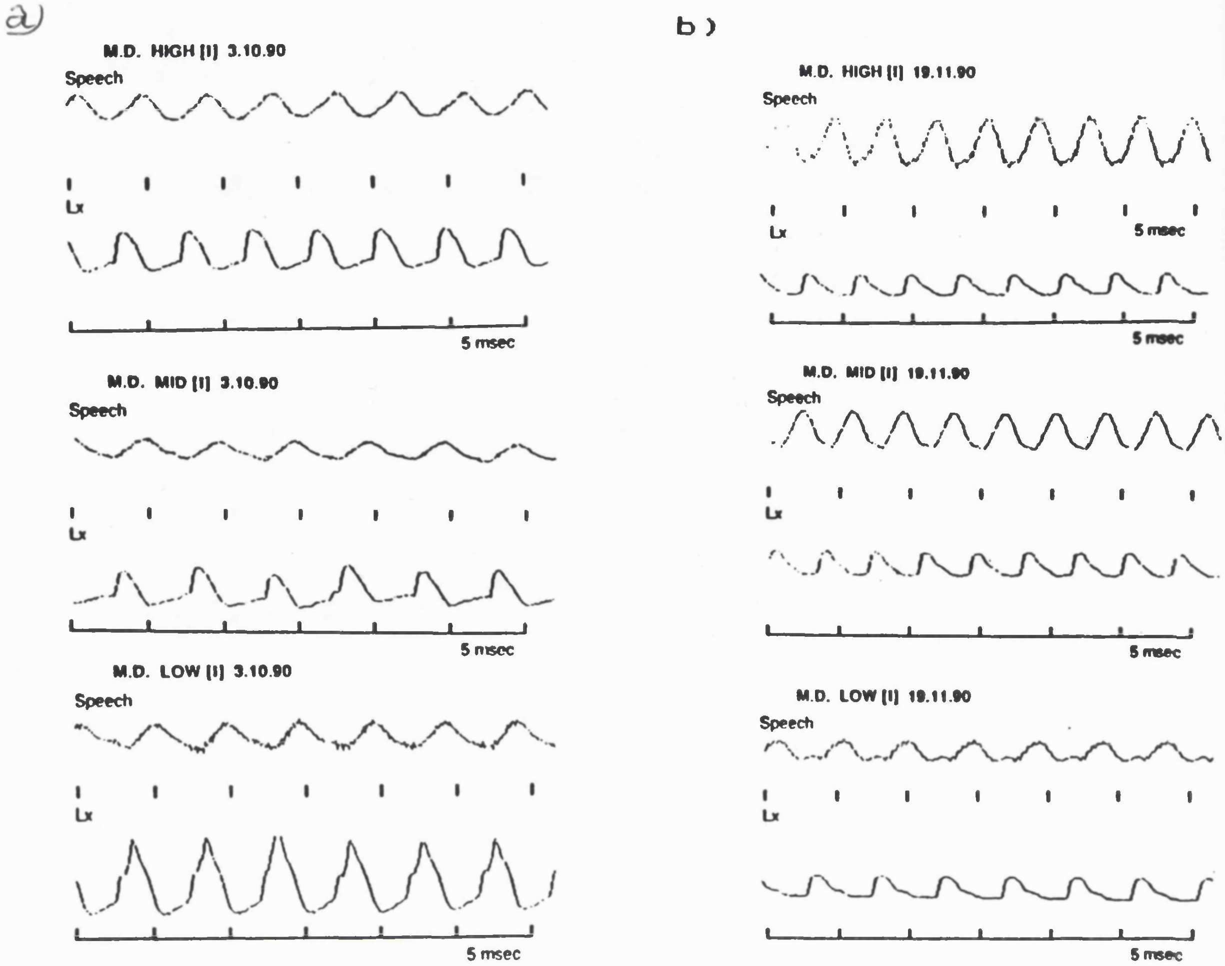
As shown in Lx waveforms produced by patients with puberphonia (Carlson, 1993 b) (Fig. 8 a, p. 72), or as shown by Motta et al (1990) in their two groups of Functional dysphonics (Table 4) (Fig. 36), EGG abnormalities may be the result of abnormal voice production as far as breath support and fundamental frequency are concerned, without evidence of organic pathology. Here the evidence of abnormality is in the open/contact phase relationship (Fig. 8 a and 37) and sometimes in the evidence of amplitude perturbation (Fig. 8 a).

Haji et al (1986) measured EGG amplitude and frequency perturbation and found, particularly amplitude perturbation (shimmer) in the EGG waveform, to be a sensitive index of phonatory stability. It differentiated between three groups of hoarse patients, slightly, moderate and severely hoarse. They also found that female speakers showed higher amplitude perturbation than male, which may be a reflection of their tendency to phonate with a posterior glottic chink. This may allow for more variation in the cycle to cycle degree of contact, which would be reflected in variability of the amplitude of the EGG waveform.

Colton and Conture (1990) agree with Painter (1988) that points in the 'idealized' EGG waveforms, particularly the opening and open phase, assumed to illustrate certain vocal fold behaviour, may or may not be present in the waveforms of both normal and abnormal speakers. They suggest that in people with vocal pathology, the vocal fold vibratory pattern may be very different from that expected for a normal voice. Interpreting an EGG signal from a pathological speaker by comparing it to an 'idealized' model waveform may be inappropriate.

They found that reliable measures of the 'closing time' or 'slope' can be made from the waveforms of pathological speakers, and that lesions on one or both vocal cords may affect the speed of vocal fold closure. Vocal nodules, polyps, oedema and cancer would tend to show longer closing times on EGG. This is evident in a case study by the author (Carlson, 1993 a, enclosed as subsidiary matter with this thesis) regarding a woman with a residual vocal fold polyp after surgery and voice therapy. It was, however, mainly evident in the waveform produced at habitual low pitch. At mid pitch there is some amplitude perturbation (Fig. 39 a). After the second operation to remove the polyp her waveforms were normalized (Fig. 39 b).

Despite their extensive report on pitfalls in the interpretation of EGG waveforms, Colton and Conture (1990) encourage more empirical studies correlating vocal fold vibratory patterns in pathology with corresponding EGG waveforms and proceed to describe their approach outlined above to quantifying aspects of the EGG waveform, measuring the duration from first to maximum contact (cf. Fig. 32, p. 149).



Lx waveforms of female patient with vocal fold polyp (From: Carlson, 1993 a)

a) before surgery b) after surgery.

Figure 39

CHAPTER VI - THE EFFECT OF RADIOTHERAPY ON LARYNGEAL FUNCTION.

1 Incidence of laryngeal carcinoma.

Tumours of the larynx constitute 2% of all cancers diagnosed, and are among the most common tumours of the head and neck. In 1980, 1725 laryngeal carcinomas were diagnosed in England and Wales, which gives an incidence of 4 per 100 000. (Sikora and Halnan, 1990). An article by Cann and Fried (1984) cites a statistic from the United States stating that 2.3% of all cancers diagnosed in males and 0.4% in females are cancers of the larynx (Young, Perce and Asire 1981, Flanders and Rothman, 1982). There is a great worldwide difference in the incidence rates of laryngeal cancer. The highest rates being reported from India (Waterhouse, 1976).

There is a strong tendency for laryngeal cancer to be more common in men (Young et al, 1981). An estimate for Britain is given by Shaw (1979) as a proportion 8:1. A tendency reported in the United States of the incidence among women aged 50 and over to have shown a marked increase, is assumed to reflect the increased smoking habits among women (Cann and Fried, 1984). A report by De Rienzo, Greenberg and Fraire (1991) noted a substantial increase in the incidence of laryngeal carcinoma both among men and women in a general hospital setting in Texas between 1947 and 1984. The incidence reported among men in 1947 was 5.6 per 100 000 and in 1984 9 per 100 000. The incidence among women was 0.5 per 100 000 which increased to 1.5 per 100 000. The male female ratio in two fifteen year periods 1959-1973 and 1974-1988 decreased from 5.6/1 to 4.5/1 (De Rienzo et al, 1991).

In a large retrospective study of laryngeal pathologies by age, sex and occupation in a treatment seeking sample Herrington-Hall, Lee, Stemple, Niema and McHone (1988) reported laryngeal carcinoma as the fourth most common diagnosis in 9.7 % of a total sample of 1262 cases. It was distributed equally across the age range 45-64 and over age 64, and developed approximately at the same age in men and women. It was most common in retired people. The male female ratio in their sample was reduced to 3/1. In a study by Maier, Dietz, Heller and

Jünemann (1990) of 95 male subject suffereing from laryngeal cancer the age distribution peaked at 51-55 years.

ii Risk factors.

There are certain factors which combine or enhance the risk of people developing cancer of the larynx. Rothman, Cann, Flanders et al (1980) cite sex, age, smoking, alcohol drinking and exposure to employment related risk factors such as inhalation of asbestos dust.

Laryngeal cancers are likely to be caused by chronic mucosal irritation produced by long term, heavy smoking and heavy drinking, particularly of spirits (Shaw, 1979), and possibly also by the chewing of tobacco and aromatic nuts, which may explain the high incidence reported in India (Waterhouse et al, 1976). De Stefani, Correa, Oreggia et al (1987) assessing risk factors for laryngeal cancer in Uruguay, report a significantly increased risk in those smoking cigarrettes made from dark (air cured) versus light (flue cured) tobacco. Starting smoking at an early age, before the age of 15, also increased the risk for developing laryngeal cancer in later life.

Evidence of the considerably increased risk of laryngeal carcinoma when heavy smoking and alcohol drinking are combined, is offered by the following studies: Herity, Moriarty, Bourke and Daly (1981), Flanders and Rothman, (1982), De Stefani et al, (1987), Guenel, Chastang, Luce et al, (1988) and Maier et al (1990).

Maier et al (1990) found that single working class men showed a higher incidence of laryngeal carcinoma, which was explained by their significantly higher tobacco and alcohol consumption. Another factor specific to this study was the finding that chronic exposure to cement dust increased the risk, which they hypothesise may be due to inhalation of chromium, which is a constituent part of cement and is a known carcinogen.

The combined effect of smoking and alcohol drinking is not simply additive but the risk is about 50% greater than the risk of cancer developing as a result of one of the agents alone (Flanders and

Rothman, 1982, Guenel et al, 1988). Guenel et al (1988) found that the relative risk of cancer as a result of the combined effect of alcohol and tobacco was greater for supraglottic than for glottic carcinoma. They also found that significant excess risk was produced by alcohol in only moderate smokers.

Explanations for the 'synergistic effect' of alcohol and tobacco in the causation of laryngeal carcinoma have been put forward. For instance Koopman (1981) suggested that the physical contact between tobacco and alcohol carcinogens facilitates entry of tobacco carcinogens into the epithelial cells, possibly with alcohol acting as a solvent. This explanation may be supported by the finding cited above by Guenel et al (1988). Mc Coy and Wynder (1979) however, suggest that it is more likely that alcohol vapour in expired air interacts with tobacco carcinogens as the alcohol would not come into contact with laryngeal epithelial cells. Further possibilities may be that alcohol changes the metabolism of the laryngeal epithelium or affects the metabolic conversion of carcinogens in the liver and indirectly interacts with the carcinogenic components of tobacco smoke (Mc Coy and Wynder, 1979).

iii Symptoms and site of lesion.

A tumour growing on or around one or both vocal folds results in hoarseness, usually noticed and investigated at an early stage in the history of the neoplasm. The cure rate and five-year survival rate after radiotherapy treatment or combined radiotherapy and surgical intervention, is high (Kaplan, Johns, Slaughter Fitz Hugh et al, 1983, Mendenhall, Parsons, Stringer et al, 1988). This is however also much dependent on the site and extent of the tumour (Shaw, 1979).

Malignancy arising in parts of the larynx other than the vocal cords, e.g. the anterior commissure, the supraglottis or the subglottic areas are less common than those of the glottis. They tend to be more advanced before they give rise to symptoms and particularly a subglottic tumour "gains early access to lymphatics, draining directly into the prelaryngeal, para-tracheal and lower deep cervical chains of nodes" (Shaw, 1979).

Mc Ilwain (1991) cites different modes of invasion of laryngeal tumours to the posterior glottis and a tendency of these to direct extension from this site into the subglottis, as the reason for a need to re-define the larynx from that commonly used in the staging of laryngeal tumours. He proposes that the posterior glottis be considered part of the subglottis. *"The posterior glottis as part of the subglottis requires a review of our present understanding of the structure and function of the larynx and in particular the spread of cancer posteriorly."* (Mc Ilwain, 1991)

Sikora and Halnan (1990) suggest that tumours involving the anterior commissure may have a worse prognosis than those arising on the middle third, the ligamental part, of the vocal fold. They also suggest that tumours extending to the posterior third of the vocal fold may easily invade the pyriform sinus. Gerritson and Snow (1991) suggest a preponderance of glottic tumours affecting the anterior portion of the vocal folds and the anterior commissure. In our study 61 % of T1 tumours and 50 % of T2 tumours are located in this area. However, 35 % of all the tumours are only localised to the right or the left vocal fold without further information (Appendix 2 A and B).

Shvili, Zohar and Rahima (1990) studied the control rates in 63 patients who had anterior commissure carcinomas. Of the 63 patients 47 were classified as having T1 lesions and 16 patients were classified as having T2 lesions. Twentynine had horseshoe shaped lesions across the anterior commissure and 34 had lesions of one cord extending to the anterior commissure.

Thirtyeight T1 lesions and nine T2 lesions were treated with primary radiotherapy. The initial control rate among this group was 72 % (34/47). Salvage surgery, seven partial and six total laryngectomies, were carried out successfully in 10/13 patients. All the three failures had T2 tumours, two had horseshoe lesions. The survival rate of patients with horseshoe lesions was 89 % (17/19) compared to those with vocal cord lesions extending to, but not across, the anterior commissure whose survival rate was 93% (26/28).

Shvili et al (1990) recommend surgery as primary treatment for lesions of the anteriror commissure as it achieved initial control of 94% compared to primary radiotherapy which only achieved a control rate of 72%.

Mendenhall et al (1988) however, in a retrospective study of control rates among 304 T1 and T2 lesions did not find evidence that anterior commissure involvement predicted radiation failure or necrosis, nor did Harwood, Bryce and Rider (1980). This is also confirmed by Benninger, Gillan, Thieme et al (1994) who did not find that anterior commissure involvement was associated with increased risk of recurrence in their retrospective study of 63 irradiated T1 and T2 tumours. Their mean follow up time was 6 years.

iv Systems for classification of laryngeal tumours.

A system of clinical classification of the severity and extent of tumours facilitates exchange of information and comparison of treatments and end stage results. Sisson and Pelzer (1985) point out the difficulties however, of developing a system that is simple enough to be readily employed but sophisticated enough to provide the neccessary information to plan the treatment most likely to result in the patient's survival.

The TNM system for classification of cancer was developed and has been regularly revised and updated to provide such a system (The American Joint Committee on Cancer 1978, 1980, 1983, 1987). The extent of each lesion is defined in terms of three parameters (Table 5).

The TNM classification system has also been adopted by The International Union against Cancer the UICC, and been revised in 1968, 1974, 1978, 1982 and 1987.

Table 5

TNM staging.

T1 - Tumour is limited to vocal cords with normal mobility.

T1a - Tumour is confined to one V.C. with normal mobility of cord.

T1b - Tumour involves both cords with normal mobility.

T2 - Tumour extends to subglottic and/or supraglottic region and/or with impairment of vocal fold mobility

T3 - Tumour as in T1 or T2 but with fixation of one or both cords.

T4 - Tumour invades through thyroid cartilage and/or extends to other tissues beyond the larynx.

N1 - Palpable lymph nodes homolateral and not fixed.

N2 - Palpable bilateral or contralateral lymph nodes, not fixed.

N3 - Fixed cervical lymph nodes.

M - Distant metastases.

(From Skolnik, King , Wheatley and Martin, 1975, and Spiessl, Beahrs, Hermanek et al, Eds. 1992).

The American Joint Committee on Cancer (1983) also added 'Tis - Cancer in situ, pre invasive carcinoma' to the classification.

Many studies have shown that impaired mobility but not fixation of the vocal cord in stage T2 lesions has a significantly worse prognosis after radiotherapy than superficial extension into the sub or supraglottis with normal mobility of the vocal folds (Harwood and De Boehr, 1980 a, Harwood, Hawkins, Keane et al, 1980, Dickens, Cassisi, Million and Bova, 1983, Kaplan et al 1983). Johns, Slaughter-Fitzhugh, Boyd et al (1983) therefore propose a subdivision of T2 lesions in T2 a, lesions with normal vocal fold mobility and T2 b, lesions with impaired cord mobility but not fixation. They found that T2 tumours with impaired cord mobility resulted in a reduced overall five year survival rate to 73% from 90% in T2 lesions with normal cord mobility.

Johns et al (1983) maintained that *"...fixed cord is usually the result of deep thyroarytenoid muscle invasion, and impaired cord mobility represents a lesser degree of invasion. Surface extension of the lesion, even involving the false cord should not be considered more extensive than one involving the deeper layers of the thyroarytenoid muscle."*

Kleinsasser (1992) agrees and suggests a 'metric system' of TNM classification where the main criterion for staging a tumour is the measured surface extension and depth invasion. He gives as a reason the finding that 10-20 % of small vocal fold carcinomas are overestimated but 30-50 % of larger T2-T4 tumours are underestimated using the present system of staging. He also proposes abandoning the separate group of subglottic tumours.

"..In all our clinical and histological investigations we have not seen a single tumour of the subglottic space (and the ventricle) that did not have a wide connection area with the squamous epithelium zone of the vocal cord" (Kleinsasser, 1992). Mc Ilwanney (1990) suggests: *"The spread of cancer to the posterior glottis from the different primary laryngeal and pyriform fossa sites, shows different modes of invasion but in particular direct extension and connection with the subglottis...the posterior glottis is so intimately related to the subglottis that it must be considered part of the subglottis."*

Limited mobility of the vocal cord is always an indication of deep invasion and Kleinsasser proposes that T2 tumours must not include those with impaired mobility, which should always be classified as T3 or T4 as they have the same poor prognosis after radiotherapy (Harwood and De Boehr,1980, Kaplan et al, 1983, Johns et al, 1983).

Kleinsasser (1992) suggests that there are embryological, functional and oncological reasons to divide the larynx into two main areas only, the supraglottis and the glottis or rather, he suggests, the vocal folds. He proposes that the division of T1 tumours into unilateral 'a' and bilateral 'b' should be extended to carcinoma in situ and to T2 tumours. Tis and T1 tumours should not show more than 15 mm surface extension, which corresponds to the membraneous part of the vocal fold.

T2 tumours should be those with a superficial diameter of 15-25 mm *"...which have not caused limitation of vocal fold mobility...and would thus include most bilateral vocal fold carcinomas extending to the vocal process or the arytenoid regions or to the subglottic region without affecting vocal fold mobility"* (Kleinsasser.1992).

T3 and T4 tumours would be those with reduced or absent vocal fold mobility irrespective of their surface extension. *" Limited mobility ...is always accompanied by histological evidence of deep invasion of the tumour, which in many cases has even penetrated the cricothyroid membrane reaching the T4 category".*

The degree of impairment of vocal fold mobility is dependent on the site of the tumour, the inflammatory reaction, the type of infiltration, whether this is into the vocal folds or with increased surface extension and whether there is infiltration into the cricoarytenoid joint. (Kleinsasser, 1992)

Finally, the depth of tumour invasion pertaining to the different stages Kleinsasser suggests, should be the basal membrane for carcinomas in situ, 3 mm for T1, 6 mm for T2 and more than 6 mm for T3 and T4 (Kleinsasser, 1992).

v Treatment for cancer of the larynx.

The initial diagnosis of a malignant laryngeal tumour is usually made through indirect laryngoscopy or endoscopic inspection of the larynx and the extent and stage of the tumour is determined. The diagnosis is subsequently confirmed through histological analysis of biopsied specimens from the affected area or areas. The most common laryngeal tumour is a generally well differentiated squamous cell carcinoma. Shaw (1979) reports a tendency for vocal fold tumours to be of this type. Gerritsen ans Snow (1991) report a predilection for glottic carcinomas to affect the anterior parts of the vocal cords and the anterior commissure.

The undisputably most important criterion when choosing a treatment for cancer of the larynx is that it should be a method which gives the patient the greatest chance of survival. The survival rate is generally high with early detection and intervention. Therefore, a secondary but also important aim is to ensure that patients enjoy the best possible quality of life after treatment.

The treatments available for laryngeal cancer on their own or in combination are:

1. Laser excision
2. Cordectomy
3. Partial Laryngectomy
4. Total Laryngectomy
5. Radiotherapy
6. Chemotherapy

For T1s tumours with no involvement of the anterior or posterior commissures laser excision may be used with good results (Strong, 1975). For early tumours of the larynx, stages T1 and T2, radiotherapy tends to be the primary form of treatment in Britain and Canada and surgical procedures are used for salvage in case of recurrence (Dickens et al, 1983).

Radiotherapy has obvious great advantages over surgery in that it leaves the patient with intact laryngeal structures and interferes minimally with communication in the long term. Dickens et al (1983) found that for early laryngeal cancer, control rates using radiotherapy with cordectomy or hemilaryngectomy in reserve, are similar to those reported using the surgical procedures as primary treatment (Neel, Devine and DeSanto, 1980). Total laryngectomy is also an option but leaves the patient with major limitations in communication compared to the other surgical procedures.

Shapsey and Hybels (1985) give a cure rate with radiotherapy for early glottic carcinomas of 80-90% three year survival without recurrence. Other researchers report control rates of early vocal fold tumours of 95 % (Korcok,1983) 80 % (Fletcher and Klein, 1964) and 87 % (Mantravadi, Liebner, Haas et al, 1983).

Mendenhall et al (1988) report a local control rate with radiotherapy of between 90 % and 100 % in a sample of 171 patients with T1 tumours. The lowest rate was for lesions with more than 15mm surface extension. The 100 % control rate was for tumours of less than 5 mm. The control rate with radiotherapy for 133 patients with T2 lesions was considerably lower, between 72 % and 85 %. The lowest control rate was for tumours of 'moderate size and reduced cord mobility'. As a result there were more partial and total laryngectomies carried out for salvage in the T2 group of patients.

These control rates are similar to those reported by Marks, Fitz - Hugh and Constable (1971). Sikora and Halnan (1990) comment that impaired mobility of the vocal fold in T2 tumours affects the recurrence rate. In the study reported by Kaplan et al (1983) the recurrence rate among T2a lesions (normal cord mobility) was 9% as opposed to 25 % in the T2b (impaired cord mobility) group. Castelijns and Snow (1991) suggest that the variation in reported cure rates is likely to be due to problems in staging the extent of the primary tumour.

Dickens et al (1983) reported 92% control with irradiation alone of T1 tumours. This increased to 98% with surgical salvage. Control rates for their patients with T2 tumours were 67 % with radiotherapy alone, which increased to 94% with surgical salvage. They emphasise

"...the chance of success is closely related to the anatomical distribution and the volume or bulk of the tumour" and *"Further analysis of T2 lesions shows that the degree of supraglottic and/or subglottic extension correlates with local failure."* (Dickens et al,1983).

vi Fractionation and Field size.

The treatment is given in fractions of radiation doses per day until a maximum of about 5 - 6000 cGy have been given. Botnick et al (1984) report studies of complications associated with varying fractions per dose. When the dose fraction size is increased to greater than 225 cGy per day a marked increase in complications such as necrosis of laryngeal cartilage, severe oedema of the vocal folds or subcutaneous fibrosis of the upper chest wall may occur (Wang, 1974). Reductions of the intensity and duration of the mucosal reaction may be achieved by altering the radiation dose by 5-10 % (Fletcher, 1980).

In a longitudinal, multicentre study (Wiernik, Bates, Bleheen et al, 1990) comparing the results of radiotherapy given in 3 fractions per week versus 5 daily fractions to patients suffering from cancer of the larynx and hypopharynx, the final report after ten years follow up, did not show any significant difference in acute or late normal tissue radiation damage, laryngectomy free rates or survival, between the two fractionation regimes. For the three fractions per week the total tumour dose was reduced by 11 % and 13 % for treatments lasting 6 weeks and 3 weeks respectively.

Studies have shown that the radiation field size, not the tumour dose, is the most important factor in determining local control in T1 and T2 tumours with normal mobility. Smaller field sizes, 5 by 5 cm or less, were associated with increased local recurrence (Harwood and Tiere, 1979, Harwood and De Boehr, 1980 a, Harwood, Hawkins, Keane et al 1980 c).

However, Inoue, Inoue, Chatani and Teshima (1992) studied the effect of field size on the control rates of T1 vocal fold cancers and found that a larger treatment field size resulted in higher proportion of minor complication rates such as persistent laryngeal oedema without improving the local control rate. A field size 5 by 5 cm resulted in a 93 % control rate with 4 % persistent oedema. A field size 6 by 6 cm resulted in 95 % control but 21 % persistent oedema.

This was also found by Mendenhall et al (1988) although they compared field sizes of 4 by 4 cms and 5 by 5 cms. They describe the extension of the radiation portal for T1 lesions as usually extending from the thyroid notch superiorly to the inferior border of the cricoid anteriorly. The posterior border depends on the posterior extension of the tumour. The portals for T2 tumours are larger because of the larger extent of the tumour.

Acute and late complications of radiation e.g. dysphagia and persistent laryngeal or soft tissue oedema are associated with large radiation field sizes (Fletcher and Klein, 1964, Stell and Morrison, 1973, Fu et al, 1982). Optimum size seems to be an area of 6 by 8 cm for supraglottic tumours and as little as 5 by 5 cm for glottic tumours on account of their sparse lymphatic drainage (Botnick et al, 1984, Wang, 1974).

Dickens et al (1983) reported mild arytenoid oedema in patients whose tumour volume required the use of a wedge to increase the dose posteriorly.

Acute effects of radiotherapy are troublesome to the patient in the short term but are reversible. Late effects are irreversible.

The sites and stages of tumours, histology, tumour dose, field sizes and fractionation used in the treatment of subjects in this study can be found in Appendix 2 A-C.

vii <u>The effect of radiotherapy on laryngeal tissues.</u>

Botnick et al (1984) state: *"The goal of cancer therapy is to cure the patient with the least amount of morbidity".*

Unfortunately ionising radiation cannot differentiate between normal and malignant tissue. The ultimate effect is brought about by differences in cell repair and repopulation in response to irradaiation.

Radiotherapy causes destruction of malignant disease both at the tumour site and at nearby microscopic extensions. It also sterilizes tumour cells within the local lymphatic system (Lo, Salzman and Schwartz, 1985). Laryngeal oedema is the main side effect as the radiotherapy causes an inflammatory response in normal tissues. It is thought to be due to an increase in vascular permeability in the acute phase followed by obliterative vascular damage over many months or years. This accounts for the hoarse or aphonic voice during the late stages of radiotherapy in some patients.

The duration of oedema after the end of radiotherapy treatment is affected by a) dose of radiation b) volume of tissue irradiated c) continued use of alcohol and tobacco and d) size and extent of the original lesion (Mendenhall et al 1988).

Mendenhall et al (1988) found a serious complication rate of 0.5 % in 184 patients with T1 lesions. This rate increased to 3.4 % in their group of 119 T2 lesions. Serious complications such as soft tissue necrosis leading to chondritis and cartilage necrosis were related to use of a single portal, T stage and a high total dose.

A study by Fu et al, (1982) found that in a group of 247 patients irradiated for cancer of the larynx, oedema was common after radiotherapy but usually subsided in 4-6 weeks. Thirtyseven patients had oedema beyond this time and 17 of these were found to have recurrent disease. Mantravadi et al (1983) report an overall incidence of laryngeal oedema of 18 %. Of these 54 % were free of cancer recurrence. This indicates that persistent oedema is an important warning sign of possible recurrence.

Factors that influence the persistence of oedema have been found to be: continued smoking (Karim, Snow, Diek and Hanjo, 1983, Mantravadi et al 1983, Rugg, Saunders and Dische, 1990, Whittet, Lund, Brockbank and Feyerabend, 1991) vocal abuse (Kagan, Calcaterra, Ward and Chan, 1974, Karim et al 1983, Fu et al, 1982, Mantravadi et al 1983) continued alcohol use (Kagan et al, 1974, Mantravadi et al, 1983, Whittet et al, 1991)

Harwood and Rawlinson (1983) and Benninger et al (1994) found a tendency to higher rate of recurrence in those patients who continued to smoke compared to those who stopped before treatment. The lack of significance of this tendency in the earlier study, may be due to the comparatively short follow-up time, only 9-15 months, whereas in the latter study, with a mean follow up time of 74 months, there was a highly significant, sixfold increased risk of recurrence.

Rugg et al (1990) demonstrated the detrimental effect of continued smoking on mucosal recovery after radiotherapy. They found that the duration of oedema was 13 weeks in those patients who stopped smoking and did not restart, as opposed to 21 weeks in those who smoked during and after treatment or started again 4 weeks after treatment. They found that in patients whose mucosal reactions healed quickly, recovery of the mucosa to normal appearance often occurred. Prolonged reactions were associated with permanently thinned and atrophic appearances.

The apparently long duration of mucosal reactions in their study may be due to the particular radiotherapy regime employed. CHART, continuous, hyperfractionated, accelerated radiotherapy, gives 36 fractions of 140 cGy, over a 12 day period with three treatments per day at six hourly intervals up to a total dose of 5040 cGy. Mucosal reactions appeared on days 13-15 in the majority of cases and persisted for 8-24 weeks (Rugg et al, 1990). They emphasise that the elimination of cigarrette smoking can halve the duration of mucositis and that under conventional fractionation regimes the effect is likely to be present but less marked.

Whittet et al (1991) used serum cotinine, a major metabolite of nicotine with a longer half life, as an objective measure of continued smoking in a sample of male patients undergoing radiotherapy for cancer of the larynx or the supraglottis. They confirmed Rugg's et al (1990) finding of the detrimental effect on the mucosa of continued smoking during and after radiotherapy. They also found a small group of non-smokers, who were however heavy drinkers, who developed significant mucositis during radiotherapy. Considering the findings reported earlier on the powerful contribution of alcohol drinking to increasing the risk of developing laryngeal cancer, this finding may not be surprising (Rothman, Cann and Flenders, 1980, Flanders and Rothman, 1982, Guenel et al, 1988, Maier et al 1990).

viii Different fractionation regimes, the BIR study.

In choosing the radiotherapy regimen that is going to maximise the patient's chances of survival, with the minimum of morbidity and maximum quality of life after treatment, the considerable cost in time and money to both the patient and the health service is a factor.

A major multicentre, prospective study was set up by the British Institute of Radiology (BIR) in 1963 (Wiernik et al, 1982, 1990) for the purpose of finding out whether there was any difference between treatments given 3 times per week versus 5 times per week in patients' survival, tumour free rates and laryngeactomy free rates. Tumour free rates were defined as "... *the probability that a given patient has not suffered from either a persistent tumour or a recurrent tumour at the primary site at a specified time*".

The trial involved random allocation to each arm of the study of 734 patients with laryngeal and laryngopharyngeal carcinoma diagnosed in either of 17 centres round Britain between 1963 and 1975. Patients were followed up for ten years, which is commonly the point at which these patients stop coming for regular review, if there has been no sign of recurrence. The mean age of patients in each arm of the trial was 62.7 years.

Tumours were staged according to the UICC directions for TNM classification available during this period (UICC 1978, 1987), but comparisons were ulitmately made between three groups, according to site - larynx, pharynx and 'dual sites', and according to nodal - status as suggested by Johns, Neal and Cantrell (1984). The reason given for using this classification is that prognosis is heavily dependent on nodal status. In the final report the total data of 713 patients entered into the trial is analysed (Wiernik et al 1990).

Survival of patients through the ten year period was, not surprisingly, heavily dependent on age, stage and site of the tumour. Comparing the two treatment regimes, there was a slight but statistically not significant trend in favour of the 5 F/week arm of the trial. After 'age correction' of the raw data the trend decreased.

Comparing 'tumour free rates', the study recorded a total of 320 recurrences in the 10 year period. The number of recurrences were very similar and differences not significant in the 3F/week and the 5F/week group. Site and stage of the tumour were important prognostic variables for recurrence.

The third criterion used for comparison of the two treatment regimes was 'laryngectomy free rates'. It turns out not to be a very useful one as not all patients who recur are suitable for laryngectomy. It was only used for 151 out of 320 patients who recurred. However, there was found to be little and non significant difference between the two treatment regimes when the nodal involvement was taken into account.

Of interest for the current study, which is going to look at voice quality in patients after radiotherapy for T1 and T2 vocal fold carcinoma, are the findings in the BIR trial of acute and late normal tissue effects.

The acute reactions are divided into 'marked mucous membrane reactions', such as ulceration or fibrinous reactions, marked mucosal oedema, marked skin reactions, dysphagia, and perichondritis. The incidence and duration, of these acute reactions was high, but there was no significant difference between the different fractionation regimes. Marked acute reactions were noted in 94% and 96% of patients in the 3F/week and the 5 F/week groups respectively. The duration of this reaction lasted longer than four weeks in 75 % and 84 % of patients respectively.

The late onset normal tissue reactions, 3-12 months after the end of radiotherapy, were divided into Skin reactions such as telangiectasia, fibrosis and atrophy, oedema and dyspigmentation; Mucous membrane reactions - telangiectasia, atrophy and oedema; Cartilage reactions - perichondritis; Pain, in or outside the treated area and inside the treated area, 3 cases of 'Probable myelitis' and Dysphagia. There was no significarnt difference between 'All late reactions' in the 3F versus 5F per week treatments with 59 % and 62 % recorded reactions respectively (Table 6) (Viernik et al 1990).

Table 6 shows the number of patients in each arm of the study who showed late mucous membrane reactions:

Table 6

Number of patients with late mucous membrane reactions

Reaction	3F/week	5F/week
Telangiectasia	58 (16%)	69 (19%)
Atrophy	16 (5%)	25 (7%)
Oedema	125 (36%)	134 (37%)
3-6 months	92 (26%)	107 (30%)
6-12 months	20 (6%)	15 (4%)
After 12 months	13 (4%)	12 (3%)
All late reactions	208 (59%)	225 (62%)

(From Viernik et al 1990)

Viernik et al (1990) point out the difficulty of diagnosing laryngeal telangiectasia, and the fact that it was performed only using indirect laryngoscopy, which may render this parameter less reliable as a criterion. The fact remains however, that the incidence was similar in the two arms of the study. The incidence of any type of late reaction was similar and not significant.

ix The Effect of Radiotherapy on Vocal Fold Function

Mendonca (1975) reported common findings on indirect laryngoscopy being hyperaemia or visible vessels on the surface of the cords, sometimes also involving the false vocal folds. The hyperaemia may be due to mucosal atrophy or thinning as a result of radiation, which would render bloodvessels more visible. Chronic inflammation of the mucosa is also common.

Among 31 hoarse patients in Mendonca's study, the following contributory causes were found:

1. Thickening of the affected cord/s but no atrophy (10/31)
2. Post-radiation hyperkeratosis revealed as pink or white leukoplakic patches on the upper surface of the vocal cords. Histology revealed mild to moderate cellular atypia and chronic inflammation. (7/31)
3. Irregular punched out defects on the affected cords, possibly the sites of deep biopsies used in diagnosis (3/31)
4. Ventricular band phonation with hypertrophied (overdeveloped, thickened) false cords meeting in the midline (2/31)
5. Bowing of cords due to bilateral weakness of the internal tensors. (2/31)
6. Anterior glottic fibrous web across the anterior commissure joining the anterior third of the cords. This was in the position of the original lesion (1/31)
7. Prolapse of the laryngeal ventricle (1/31)
8. Benign vocal cord polyp (1/31)

In an earlier study by Riska and Lauerma (1966) stroboscopy was used to observe the dynamics of vocal fold vibration. Of the twentyfour patients who had undergone primary radiotherapy for early glottic cancer, only three turned out not to have any observable abnormality in vocal fold function. All the others exhibited degrees of redness or thickening of vocal folds. atrophy or paresis of the internus muscle, asymmetries and even vocal folds positioned at different levels in the larynx.

The effect of this was irregular or asymmetrical vocal fold vibrations. Particularly the vocal fold originally affected by cancer appeared red, thickened or atrophic resulting in deviant vibration patterns. Despite this the majority (15) of this group were judged to have voice qualities that fell into the top two out of five categories. Interesting to note, however, is that one third of these patients were still not satisfied with their voices, particularly those who had gone back to work.

A comprehensive more recent study by Lehman, Bless and Brandenburg (1988) compared a group of 20 male subjects, aged 55-80, 1-7 years after radiotherapy for T1 tumours of the vocal folds, to a group of normal age and sex matched subjects. They used videostroboscopy to observe the irradiated larynges, and found that 12 of 20 subjects showed irregular closure of the glottis. The vibratory edge of the vocal fold margins were irregular in 17 subjects, often on the side contralateral to the tumour. The amplitudes of vibration were reduced and abnormal mucosal wave and phase abnormalities were observed.

Some degree of extraneous supraglottic activity was also observed in 13 subjects

All the subjects had at least had one direct laryngoscopy and biopsy prior to irradiation and Lehman et al (1988) suggest the explanation for the observed abnormalities may be radiation fibrosis and breakdown of elasticity. Depth of biopsy or stiffness resulting from it, is also a possibility, but Lehman et al (1988) found that patients who had stripping of the affected vocal fold did not show significantly worse acoustic perturbation measures than patients who had more localized biopsies.

x The effect of radiotherapy on acoustic measurements of voice function.

Werner-Kukuk, von Leden and Yanagihara (1968) used ultra-high-speed photography for observation of the larynx at different times during and after radiotherapy in one subject. The mass lesion and inflammatory reaction of the laryngeal mucous membranes resulted in the restriction of the amplitude of vibration of the vocal folds and prevented their close approximation.

Some of their objective measurements consisted of a registration of airflow during phonation of vowels measured by a pneumotachograph. As a result of the impaired vibration and closure of the vocal folds they registered increased airflow values, both rate, volume and fluctuation and also a reduction in the patient's maximum phonation time. They stress the value of aerodynamic observations, as they found those signalled improvement in laryngeal function before recovery was observable through other means. This is confirmed by Murry, Bone and von Essen (1974) who also measured mean airflow-rate during sustained vowel phonation in one subject, before, during and after radiotherapy. This was the only one of their measures. which showed a constant reduction over four occasions.

Lehman et al (1988) measuring airflow in patients one year or more after radiotherapy, found wide variations in their subjects both at normal and loud intensity, but the airflow means fell witin normal limits.

Werner-Kukuk et al (1968), Isshiki, Okamura, Tanabe and Morimoto (1969), Iwata and von Leden (1970), Hecker and Kreul (1971), Kakita et al (1977) and Mook Yoon, Kakita and Hirano (1984) have all used versions of the soundspectrograph to measure the effect of malignant laryngeal lesions on fundamental frequency, changes in harmonic structure and energy distribution of noise components.

Yanagihara (1967) in his description of how soundspectrographic data can be used in classifying degrees of hoarseness, suggests that the increase of noise components in second and third formant ranges, with increasing hoarseness, may originate from *"turbulent airflow due to incomplete closure of the glottis during vibratory cycles, or irregular vibratory attitudes of the glottis"*. The above studies seem to substantiate this claim.

Mook Yoon et al (1984) compared sonagram recordings of three groups of subjects, a) normals, b) patients with T1 lesions of one vocal fold with normal mobility, and c) patients with T3 lesions confined to the larynx with fixation of one or both vocal folds.

They found five parameters which seemed to correlate with the stage of glottic carcinoma. The median of all of the following measures increased with advancement of the lesion:

1. the extent of Fo fluctuation
2. the speed of Fo fluctuation
3. the extent of amplitude fluctuation
4. the relative level of higher harmonic components
5. the relative noise level

Lehmann et al (1988) in an objective assessment of voice production after radiotherapy for T1 cancer of the larynx, found that both jitter and shimmer measurements of the pathological group of speakers were significantly increased compared to a sex and age matched group of normal speakers. The vowel /α/ had the greatest degree of perturbation compared to /i/ and /u/. The latter showed the least. This agrees with Horii's findings (1980).

Hoyt, Lettinga, Leopold and Fisher (1992) analysed the voices of two groups of patients before radiotherapy and 6 months after completion. One group (N=10) was irradiated for glottic carcinoma the other (N=25) for other head and neck lesions but the radiation field included the larynx. They found that the patients with laryngeal lesions showed an improvement in perturbation measures post radiotherapy, the non-laryngeal group showed a deterioration in perturbation measures. The explanation lies in the beneficial effect of radiotherapy on the voice of a patient with a glottic tumour but the detrimental effect on the presumably normal voices of patients with tumours in the naso- or oropharynx of radiation treatment.

Looking at Hoyt's et al data it seems the glottic group before treatment show perturbation measures of 4.359 % the non-glottic group 1.547 %. Post radiotherapy they show very similar measures of 3.611 % and 3.343 % respectively. The task consisted of sustained vowel /α/ phonation.

Koike (1973) demonstrated that for sustained phonation, subjects with laryngeal carcinoma showed significantly greater mean magnitude of perturbation than subjects with vocal fold paralysis. Both showed significantly greater perturbation than normal speakers. Murry and Doherty (1981) similarly found both directional and magnitudinal perturbation factors for subjects sustaining the vowel /α/ and reading the third sentence of the Rainbow Passage, significantly greater in subjects with laryngeal cancer.

Murry (1978) extracted the third sentence from a reading of the first paragraph of the "Rainbow Passage" and found that the mean speaking fundamental frequency, the standard deviation and the semitone range of the voices of patients with vocal fold palsy were significantly reduced compared to a sample of normal speakers. However, Mean speaking fundamental frequency failed to separate normal speakers from the other two groups of pathological speakers a group with benign mass lesions and a group with cancer of the larynx. Hecker and Kreul (1971) found that patients with laryngeal cancer had more restricted pitch range than normal speakers reading the second sentence of the "Rainbow Passage".

There are few studies which compare frequency measures of normal and pathological voices over longer stretches of speech or reading. Until recently speech technology did not allow analysis of more than a few seconds of phonation, which accounts for the great number of studies that report measurements taken from spoken or repeated sentences.

Stoicheff (1975) measured mean speaking fundamental frequency (SFF) in a group of normals and a group who had undergone radiotherapy for glottic cancer. She used a Fundamental Frequency Indicator and the subjects were recorded reading a standard passage. Her results indicate consistently lower mean SFF for radiotherapy subjects.

There was also a tendency for smokers to have lower SFF. This was the case both within the normal and the irradiated group and Stoicheff considers this finding a possible confirmation that smoking causes some vocal fold oedema.

Lehman et al (1988) recorded subjects reading the 'Rainbow Passage' and in a two minute conversation. Compared to the normal group, both jitter and shimmer, and the variability of these measures, were significantly higher in the post radiotherapy group. Signal to noise ratio was significantly lower. There was no signifcant difference in fundamental frequency. Nor were there any differences in these measures when different techniques of radiation, type or number of biopsies, location of tumour or time since the end of radiotherapy were compared.

In the study which will follow, irradiated speakers are recorded in a conversation and in reading aloud, 2-3 paragraphs from the Rainbow Passage (Fairbanks, 1960), using Electrolaryngography (ELG) (Fourcin and Abberton, 1971, 1976, Wechsler, 1977, Abberton and Fourcin, 1972, 1984, Abberton, Howard and Fourcin, 1989). This is a very reliable method for deriving fundamental frequency information straight from the voice source via two superficial electrodes placed either side of the thyroid cartilage. A detailed description of the method will be given in a later section.

xi The effect of radiotherapy on perceptual voice quality parameters.

Isshiki et al (1969) used an early version of the GRBAS scale to attempt differential diagnosis of hoarseness resulting from vocal cord polyps and nodules as opposed to cancerous lesions. The voice quality resulting from a cancerous lesion, of the same size and on the same site on the vocal cord as a benign lesion, was rated significantly more 'Breathy' than the voice quality produced by the benign lesions. The voice of a patient with laryngeal cancer would have degrees of both 'Rough' and 'Breathy' quality as opposed to the benign laryngeal lesions that produced 'Rough' quality. The acoustic correlate of 'Breathy' quality, the 'B' factor, was characterised by a marked noise component in the spectrum, with reduced or negligible harmonic components. The acoustic correlate of factor 'R' (rough) was found to be fundamental frequency perturbation. The latter was hypothesised to be due to asymmetry of the vibrating cords.

Isshiki et al (1969) suggest the explanation for the difference in voice quality may be the generally less elastic and rougher surface of a cancerous lesion compared to the benign lesion. This would interfere more with the Bernouilly effect during vocal cord vibration and limit the amplitude of vibration as well as interfere with closure of the glottis. This seems to have been confirmed in studies using stroboscopy evaluation of vocal fold function.

The study by Stoicheff, Ciampi, Passi and Fredrickson, (1983) showed significant differences in the degree of perceived 'dysphonia' between normal speakers and laryngeal cancer patients both before and after radiotherapy for T1 tumours. The treatment changed the voice quality from predominantly 'strained', 'hoarse' and 'breathy' to 'hoarse' and 'rough'.

The longitudinal retrospective study by Benninger et al (1994) reported 67 % of irradiated patients' voices rated as 'Normal' or 'Near normal', 25 % as 'Raspy' or 'Weak' and 12 % as having 'Poor' voice quality. 40 of the 63 patients were rated by the researchers in the clinic and were also asked to rate their own voice quality using the same terms as above. There was found to be a 90 % agreement between the ratings.

Many of the above studies of voice quality in irradiated patients, report a return to perceived normal or near normal voice quality in a majority (Mendonca, 1975, Stoicheff, 1975, Karim et al, 1983), Lehman et al (1988) conclude, however, on the basis of their extensive measurements and comparisons of irradiated subjects with normal speakers that... '*..radiation therapy of stage I glottic carcinoma results in an abnormal voice. It is produced with greater than normal effort. This appears to be the result of a diffuse process that affects more of the larynx than the area involved with tumour.*'

In this study, parts of the Vocal Profile Analysis scheme (VPA) (Laver, Wirz, MacKenzie and Hiller, 1981), described earlier, will be used for perceptual evaluation by the researcher of irradiated patients' voice quality. Electrolaryngograph based objective measurements of speaking and reading fundamental frequency and regularity of vocal fold vibration will also be used. Subjects will also be rating their own experience of vocal limitations and hoarseness at different times after radiotherapy.

xii Voice and Quality of life after treatment for laryngeal carcinoma.

Successful cancer treatment is usually measured in terms of five year survival or tumour free rates. The survival rate after treatment for early laryngeal tumours is generally high with early detection and intervention. Therefore, a secondary but important aim should be to ensure that patients enjoy the best possible quality of life after treatment. Some attempts at evaluating this with respect to laryngeal cancer patients are reported below.

Slevin, Plant, Lynch et al (1988) reported poor agreement between patients and their clinicans' judgements of 'Quality of life'. They maintain: "*If measurement of a patient's quality of life is required, it should be done by the patients themselves and not by their doctors and nurses*". This echoes a statement by Harwood and Rawlinson (1983) "*...studies based on quality of life should be heavily weighted in favour of data obtained from the patients themselves.*"

Information about patients' subjective perception of voice quality and limitations in voice use and function, has been gained by interview (Mendonca, 1975, Karim et al, 1983, Harwood and Rawlinson, 1983, Lehman et al, 1988) or by asking patients to fill in a questionnaire (Stoicheff 1975, Llewellyn-Thomas, Sutherland, Hogg et al, 1984) at varying times after treatment.

Many studies report on the treatment's effect on voice quality in terms of 'normality' (Mendonca, 1975, Stoicheff, 1975, Karim et al, 1983). An attempt to summarize some of these findings is found in

Table 7.

According to these studies, a majority of patients report a return of the voice to 'normal' within a certain time after radiotherapy. The discrepancy between the studies is likely to be due to different criteria being used for judging 'normality'. Karim et al (1983) asked both the patient and a close relative to rate the voice, as well as the interviewer. Categories and scores used were as follows:

1. Excellent and voice better than predisease states rated 100%
2. Normal or near normal rated 80-99%.
3. Reasonable 60-79%.
4. Unsatisfactory or worse 60%.

"A patient observed speaking with a reasonably good voice was scored 70% even if he and/or his wife might have emphatically stated that the voice was excellent, normal or near normal." (Karim et al, 1983)

TABLE 7

Table showing patients' self rated voice recovery rates in different studies.

	N	Time post Rx	Normal voice	Reasonable voice	Unsatisfactory voice	Time for voice to recover
Mendonca, 1975	68	2-12 yrs	69%	25%	6%	3-6 mths
Stoicheff 1975	227	3-11 yrs?	44.5%	38.8%	11.5%+ 3.5%*	4 mths
Karim et al 1983	150	2 yrs	76%	10%	14%	2-6 mths

*includes the two lowest rating categories

Mendonca (1975) divided patients into three categories: Group 1 - 'Normal speaking voice' was a group that included 10 patients who reported 'normal' voice but who were judged as having a residual 'mild hoarseness'. Group 2 - 'Improved' - included 17 patients with 'mild hoarseness' and 11 with 'marked hoarseness'. Group 3 - showed

'No voice improvement'.

Lehman et al (1988) report poor agreement between patients' self perception of voice quality and objective measurements, which show that the post radiation voice is not normal. *"It may be that the patients' standards of what is normal are lowered when they are faced with a neo-plasm that could completely destroy their voices".* This may explain both Mendonca's (1975) and Karim's et al (1983) observations regarding patients' more favourable opinions of their voice qualities than their clinicians'.

Stoicheff (1975) carried out a comprehensive study of the effects of radiation treatment on patients' 'Quality of life'. The patients made self-ratings of voice using the following categories:

1. Normal
2. Improved but not quite normal
3. Improved but still a problem
4. A definite problem

Patients were also asked to give information regarding the time the voice took to recover (Table 7). Stoicheff (1975) found that they tended to rate their voices more favourably with increasing time after radiotherapy. She attributes this to the possibility of continued microscopic changes in tissues subjected to radiotherapy, compensation in vocal cord function, decreased abuse in the form of smoking and reduction in abusive vocal behaviour such as shouting and too much talking. Continued employment with or without changes in duties due to voice problems was enjoyed by 160 out of a total of 172 respondents who were working before radiotherapy.

Despite the high proportion of patients in her study reporting 'Normal' or 'Near Normal' voice after radiotherapy (83%), a majority (80%) also report persisting problem with fatiguing of voice, difficulty in singing/speaking loudly or shouting and hoarseness of the voice. This may be explained by Lehman's et al findings (1988), on the basis of their comprehensive objective measurements, among them measures of airflow and subglottic pressure, that voice production requires considerably more effort after radiotherapy. This was confirmed by observation of extraneous supraglottic activity on videostroboscopy, and patients' subjective report of an increased effort required to speak (Lehman et al, 1988).

Stoicheff did not find any correlation between voice rating and smoking during or after treatment or the time the voice took to recover nor did Lehman et al (1988) find any significant differences due to smoking, in their objective measures of voice function in subjects assessed between 1 and 7 years after radiotherapy. Karim et al (1983) did however find significantly higher smoking habit scores among those patients who had unsatisfactory voices.

An extensive 'Quality of life' study of 129 patients, 9-15 months after radiotherapy for early (T1 and T2) and advanced tumours (T3 and T4) of the larynx and after surgery, partial or total laryngectomy s reported by Harwood and Rawlinson (1983). They compared the quality of life and aspects of voice quality using interviews and patients' self ratings.

Patients were divided into three groups, those irradiated for T1 and T2 tumours, those with T3 and T4 tumours and the surgery group. Not surprisingly, irradiated patients enjoyed a considerably better 'quality of life' than the surgery group and the authors strongly recommended their policy of using radiation with surgery in reserve for the treatment of laryngeal cancer, including more advanced T3 and T4 tumours.

xiii <u>Reliability of rating scales.</u>

In an effort to find reliable and valid measures for patients' self ratings of voice quality, that would be sensitive to clincal changes in the larynx, for instance during a course of radiotherapy, Llewellyn - Thomas et al (1984) used 'Linear analogue self assessment' (LASA) which is a set of scales equivalent to Visual analogue scales (VAS) described in Chapter III about perceptual analysis of voice quality (Sederholm, McAllister, Sundberg and Dalkvist, 1992).

An example of some items included in the LASA scales used by Llewellyn - Thomas et al (1984) is seen in Fig. 40 . The anchor points at each end of the 100 mm line, describe the opposite extremes of a voice symptom or function. Patients were asked to mark the point on the line that best described their own state at the time. Scores were calculated in millimetres measured from the 'severe end' of the scale. Low scores consequently indicated more severe problems than high scores.

Similar symptoms and disabilities were selected to those that had been identified by Stoicheff (1975) and Harwood and Rawlinson (1983), as frequently causing problems after radiotherapy.

Example of a Linear Analogue Scale for self assessment

Over the past week I have been....

Able to speak without any effort	______________________________	Able to speak only with great effort
Able to be heard in a noisy environment	______________________________	Able to speak only in a whisper.

(From : Llewellyn-Thomas et al, 1984)

Figure 40

Llewellyn - Thomas et al (1984) interviewed two groups of patients. One group (N=30) was just starting radiotherapy for laryngeal carcinoma, the other (N=29) were 18 months post radiotherapy. The group beginning treatment (On-treatment) was interviewed twice at the beginning of treatment, one or two days apart then again twice, one or two days apart at the end of radiotherapy. They were asked to rate their voices on the LASA scales on each occasion. The 18 months post treatment group was interviewed and rated their voices twice, about six weeks apart.

The On-treatment group's LASA scores were compared for their sensitivity to change in symptoms from the beginning of radiotherapy to the end of treatment. For each patient the average of their two scores from the beginning of treatment were compared to the average of those obtained at the end. Mean scores were systematically higher at the start of therapy than at the end for all items rated, indicating worse voice quality as an immediate result of radiotherapy. Statistical significance was achieved for 3 of 8 symptoms: mouth/throat dryness, effort to speak, and voice fatigue (Table 8) and 5 of 8 functions: 'ability to whistle', 'use of voice in usual leisure activities', 'use of voice in usual work activities', 'ability to converse with family and friends' and 'ability to use the telephone' (Table 9). All voice symptoms and functions did show change, however, from the beginning to the end of treatment.

TABLE 8

Test - retest reliability coefficients of vocal symptom scales applying to two groups of patients with laryngeal carcinoma. One group tested before radiotherapy (RT) and at the end of RT. The other tested twice at 18 months post RT.

LASA SCALES	On treatment group RT begins+ (N=34)	RT ends++ (N=30)	Overall intraclass corr. coeff.+++	18 months Post treatment group (N=29)+++
SYMPTOM SCALES				
Selfconciousness re: voice	0.79	0.93	0.84	0.84
Effort required to speak	0.68	0.86*	0.76	0.57
Decreased range of expression	0.72	0.82	0.75	0.76
Mouth/throat dryness	0.69	0.80*	0.74	0.47
Change in voice sound	0.73	0.67	0.71	0.64
Hoarseness	0.67	0.73	0.70	0.63
Voice fatigue	0.65	0.73*	0.68	0.65
Loss of voice	0.62	0.63	0.63	0.50

(From Llewellyn - Thomas et al, 1984)

* Significant difference between pre- and post ratings
\+ Correlations between two pre Rx ratings
++ Correlations between two post Rx ratings
+++ Relationship between test and retest scores

The 'Overall intraclass correlation coefficient' column in the table above and in Table 9 pertains to the relationship between the test and retest scores for the On treatment group.

TABLE 9

Test - retest reliability coefficients of voice function scales applying to two groups of patients with laryngeal carcinoma. One group tested before radiotherapy (RT) and at the end of RT. The other thested twice at 18 months post RT.

LASA SCALES	On treatment group		Overall intraclass corr. coeff.+++	18 months Post treatment group+++ (N=29)
	RT begins+ (N=34)	RT ends++ (N=30)		
FUNCTION SCALES				
abilitY to...				
Use the telephone	0.81	0.90*	0.83	0.55
Shout	0.75	0.91	0.82	0.83
Be heard in a noisy environm.	0.76	0.87	0.81	0.85
Use voice in usual work	0.90	0.76*	0.81	0.09
Sing	0.75	0.71	0.69 (N=15)	0.79 (N=20)
Converse with family and friends	0.57	0.71*	0.64	0.68
Whistle	0.54	0.65*	0.61 (N=23)	0.88 (N=20)
Use voice in usual leisure activ.	0.67	0.56*	0.57	0.09

(From Llewellyn - Thomas et al, 1984)

* Significant difference between pre- and post ratings
\+ Correlations between two pre Rx ratings
++ Correlations between two post Rx ratings
+++ Relationship between test and retest scores

There were lower reliability coefficients in the group of subjects rating themselves at 18 months post radiotherapy with a six week interval, the 'Post-treatment group' (Table 8 and 9). Llewellyn - Thomas et al (1984) suggest that the reason for this may be twofold a) the longer time period between test and retest and b) a lower interest level in respondents a long time after treatment, when they may not experience much difficulty. There were particularly low reliability coefficients in this group for 'ability to use the voice in usual work' and 'ability to use the voice in usual leisure activities' and this is put down to the extremely low variance in scores obtained from this group 18 months post radiotherapy. Presumably because their voices did not change significantly in function between assessment occasions at this late stage.

In the following study a questionnaire will be used for irradiated subjects to rate their subjective experiences of symptoms of hoarseness and vocal fatique and limitations in voice function at different times before, during and after radiotherapy. A seven point equal appearing interval scale is used, which is somewhat easier to score than the LASA scales described above.

xiv <u>The role of Voice therapy in the rehabilitation of patients after radiotherapy.</u>

Entering the field of voice evaluation after radiotherapy from the Speech Pathologist's perspective, Stoicheff (1975) remarks: *"Only one attempt to direct the vocal efforts of such patients during treatments has been reported in the literature (Fex and Henriksson, 1969). It would be unusual indeed if these patients with an original serious organic involvement of the vocal cords were to make the easiest best use of their speaking mechanism following treatments."*

Fex and Henriksson (1969) report an attempt at reducing the risk of secondary vocal abuse during and after radiotherapy by offering voice therapy in parallel with irradiation.

Fifteen patients were examined and recorded before and after treatment. They were advised of the difficulties to come and instructed in 'voice hygiene' which is taken to mean an explanation of how not to abuse or force the vocal mechanism while it is undergoing radiation treatment. They remark: *"...while therapy is given the patient may have appreciable laryngeal trouble; dysphonia passing into aphonia for one or more weeks, pain on swallowing and occasionally a painful cough."*

Voice therapy consisted of instruction in and practice of relaxation and breathing techniques and was aimed at teaching the patient to *"use the vocal cords as circumstances permit and to adapt to changed conditions in the larynx"*. Most patients were aphonic immediately after radiotherapy. Three weeks' later most of them reported having a useful voice and at four weeks there was no evidence of 'infiltrative process' in any of the patients.

All patients except one *"had voice qualities well within normal limits"*. The majority of the patients reported finding the voice treatment comfortable and of good use, also when the voice had returned. Fex and Henriksson conclude that their results support the assumption that the effect of radiation damage can be reduced by offering 'phoniatric' (in this country - voice therapy) treatment.

Stoicheff (1975), on the basis of her study of the change in speaking fundamental frequency (SFF) following radiotherapy, suggests that *"...the study of irradiated patients' voices before, during and following treatments by means of audio, aerodynamic and acoustic instrumentation would yield additional information of diagnostic value on recovery of laryngeal function."* She agrees with Fex and Henriksson's (1969) recommendation of the contribution of the voice therapist in the assessment and guidance of patients during and following treatment.

Increased mass and stiffness of vocal fold tissues resulting from oedema or the presence of a mass on one or both cords, leads to increased laryngeal effort in the attempt at maintaining appropriate pitch level, range and volume. This may not least be the case in the period before laryngeal cancer is diagnosed, when increasing hoarseness is the presenting symptom. Increased laryngeal effort is observed after radiotherapy (Lehman et al, 1988).

Radiotherapy to the larynx gives rise to vocal fold stiffness, oedema and mucosal dryness, which develops during, and persists after radiotherapy. This results in discomfort during speaking and deterioration in voice quality. Radiotherapy patients are likely to react in similar ways to such symptoms as patients with benign diagnoses such as chronic laryngitis, Reinke's oedema, polyps or nodules.

Common vocally abusive behaviours consist of e.g. habitual or professionally required, talking over noise, continued talking despite vocal fatigue, persistent throat clearing and/or coughing, excessive laryngeal tension and poor breath control. Most people have little awareness of the role of breathing in voice and speech production. They concentrate efforts at producing and maintaining phonation in the larynx itself, thereby preventing optimum function of the vocal folds under the temporarily extremely unfavourable circumstances, which exist before, during and after radiotherapy.

Based on their findings of great variability in patients' acoustic and aerodynamic measurements, which could not be explained by differences in radiotherapy treatment or demographic variables, Lehman et al (1988) suggest:

"...there may be variable ability among patients to use a poor vibratory structure to maximum advantage through good vocal habits or compensatory vocal manoeuvres. The differences in vowel formants indicates that some compensation for glottic abnormalities is being attempted through vocal tract positioning".

Some vowels in the Lehman et al study were produced with lower jitter and shimmer and higher signal to noise ratio. This suggested to the researchers that voice therapy might help to improve these patients' voice production.

This study hopes to demonstrate that voice therapy can help patients improve and maintain good voice quality after radiotherapy for early glottic tumours.

CHAPTER VII -THE LARYNGOGRAPH PROCESSOR IN ASSESSMENT AND MEASUREMENT OF VOCAL FOLD FUNCTION.

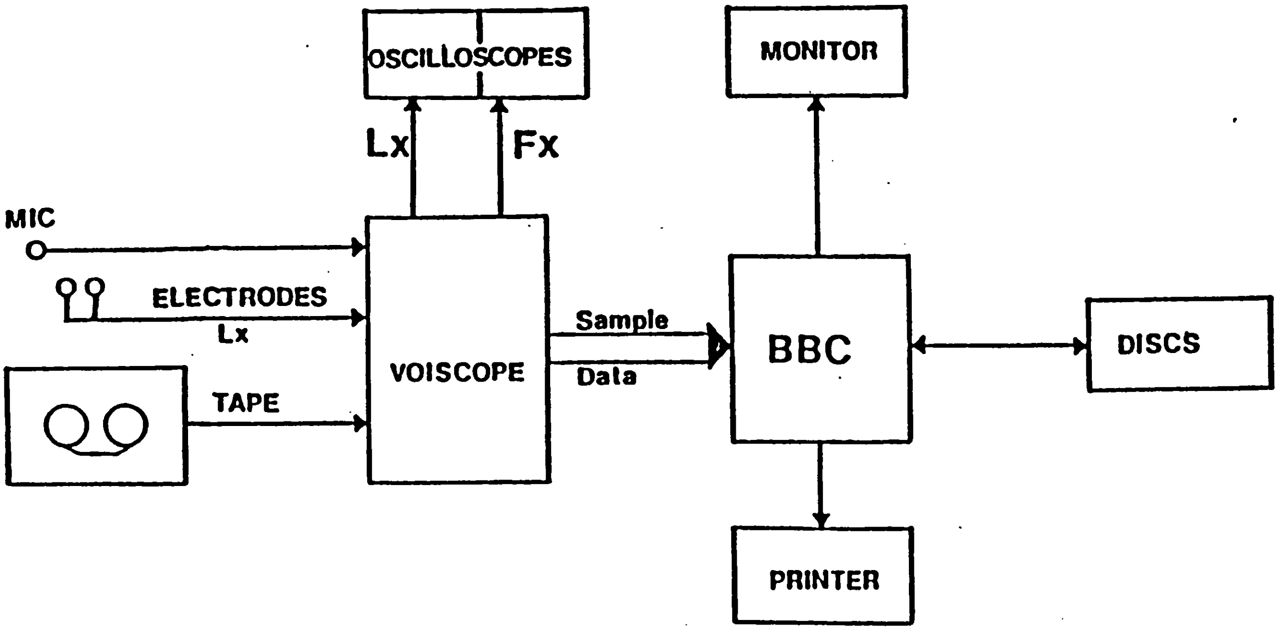
i The derivation of Larynx excitation (Lx) parameters using the Laryngograph Processor.

Very accurate registration of the timing of vocal fold contact by means of superficial laryngeal electrodes placed either side of the thyroid cartilage, makes EGG and ELG instruments reliable and relatively simple tools for registration and measurement of vocal fundamental frequency (Fourcin and Abberton,1971, Abberton and Fourcin, 1972, Abberton, 1976, Wechsler 1977, Neil, Wechsler and Robinson, 1977, Kitzing, 1979, Askenfelt, Gauffin, Sundberg and Kitzing, 1980, Abberton and Fourcin, 1984).

Askenfelt et al (1980) showed that the fundamental frequencies derived from strong EGG signals were very similar to fundamental frequencies derived from contact microphone input except for very breathy voices. They therefore recommended EGG for detailed studies of fundamental frequency as a function of time, at least for speakers who do not have breathy phonation. This has been confirmed more recently by Ohlsson (1988 b)

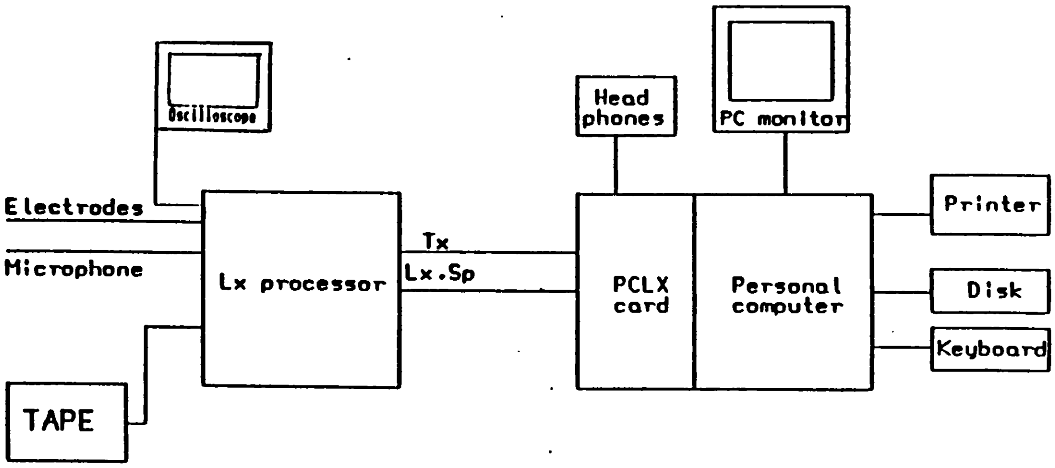
The Voiscope incorporating the Laryngograph (Abberton and Fourcin, 1984) displayed the Lx waveform and the Fx fundamental frequency contour, derived from Lx, on two oscilloscope screens (Fig. 41). The development of the Laryngograph (Lx) Processor allows the alternative display of either of these features on a personal computer monitor and the printing of hard copy. The instrument line-up used in this study most recently retains one oscilloscope for simultaneous Lx display (Fig. 42).

The Lx processor carries out an analogue to digital conversion of the time span (Tx) between successive vocal fold closures (Lx) (Fig. 43). This is then used as the basis for the calculation of fundamental frequency or Fx, data (1/Tx=Fx) (Fourcin, 1981, Abberton, 1976, Leff and Abberton, 1981, Ison, 1985, Abberton, Howard and Fourcin, 1989,



THE VOISCOPE

Figure 41



PCLX system

Figure 42

Ball, Faulkner and Fourcin, 1990, Barry, Goldsmith, Faulkner and Fuller, 1990 a,b, 1991, Carlson, 1986, 1988 b, c, 1993 a and b).

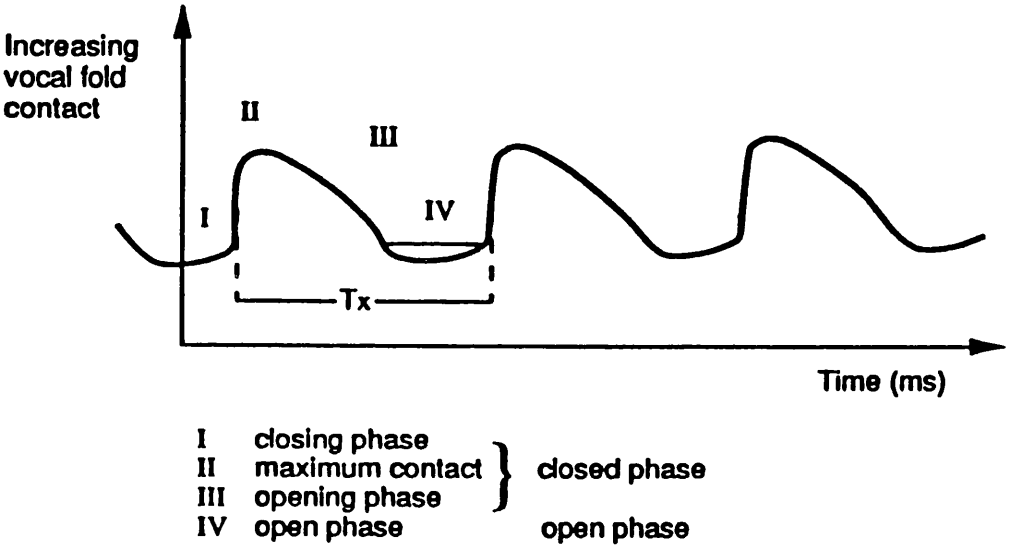
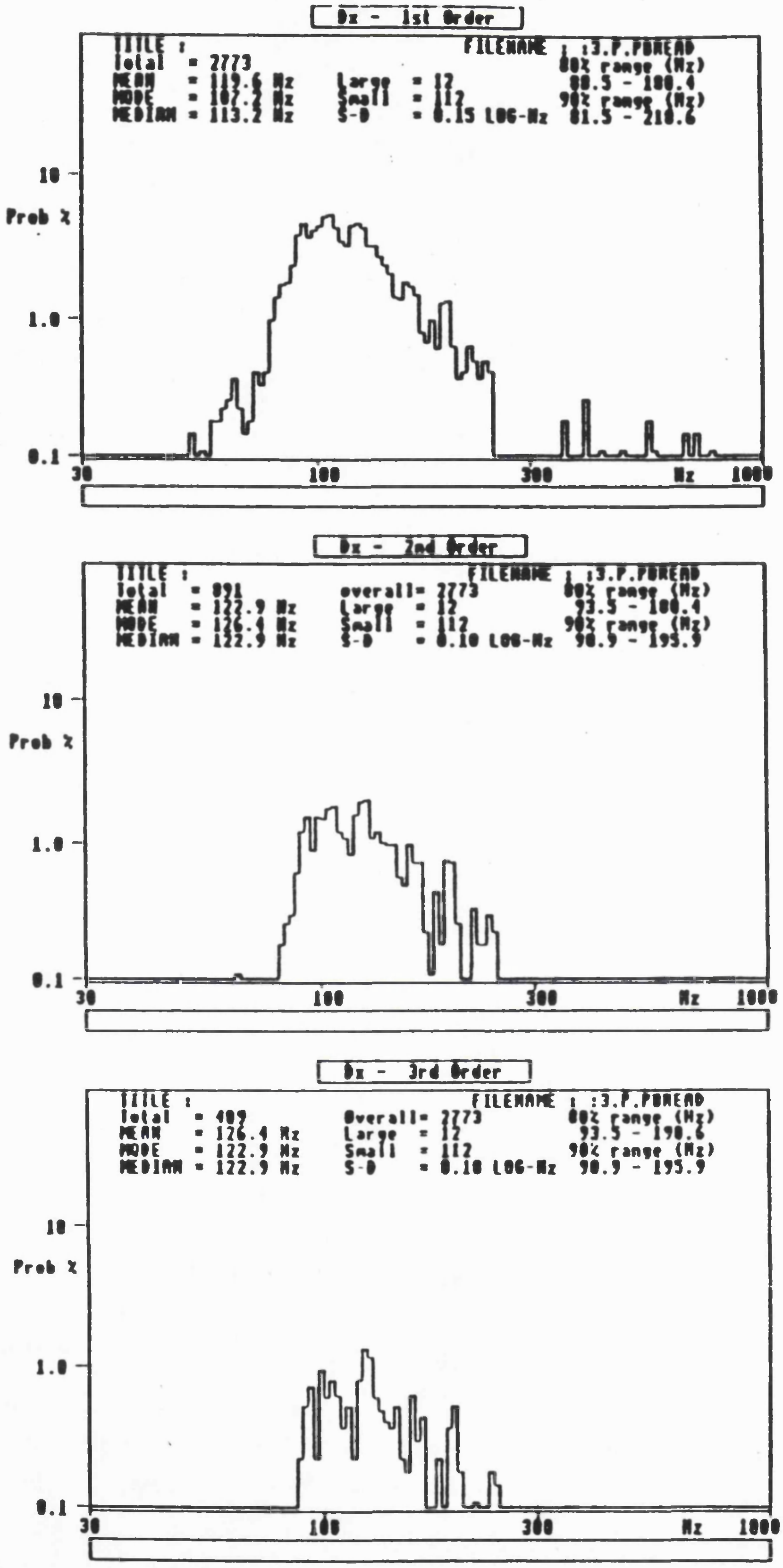


Figure 43

ii Distribution of excitation - Dx plots.

Analysis of the recorded, or live, fundamental frequency (Fx) data of long stretches of speech or reading allows a number of Fx parameters to be calculated, tabulated and displayed. The 'Distribution of excitation' (Dx) program plots the data in histogram form. (Fig. 44).

Dx plots show the probability of occurrence of a range of Fx values registered during a recorded voice sample. Dx is plotted on a logarithmic scale to correspond to perceived pitch and the frequency axis is divided into 128 logarithmically equally spaced 'bins', a few Hz in width, in the range 30.52 Hz to 1000 Hz (Ison, 1985, Abberton et al, 1989, Ball et al, 1990, Barry et al, 1990 a,b, 1991)



1st 2nd and 3rd order Dx plots
TPS software.

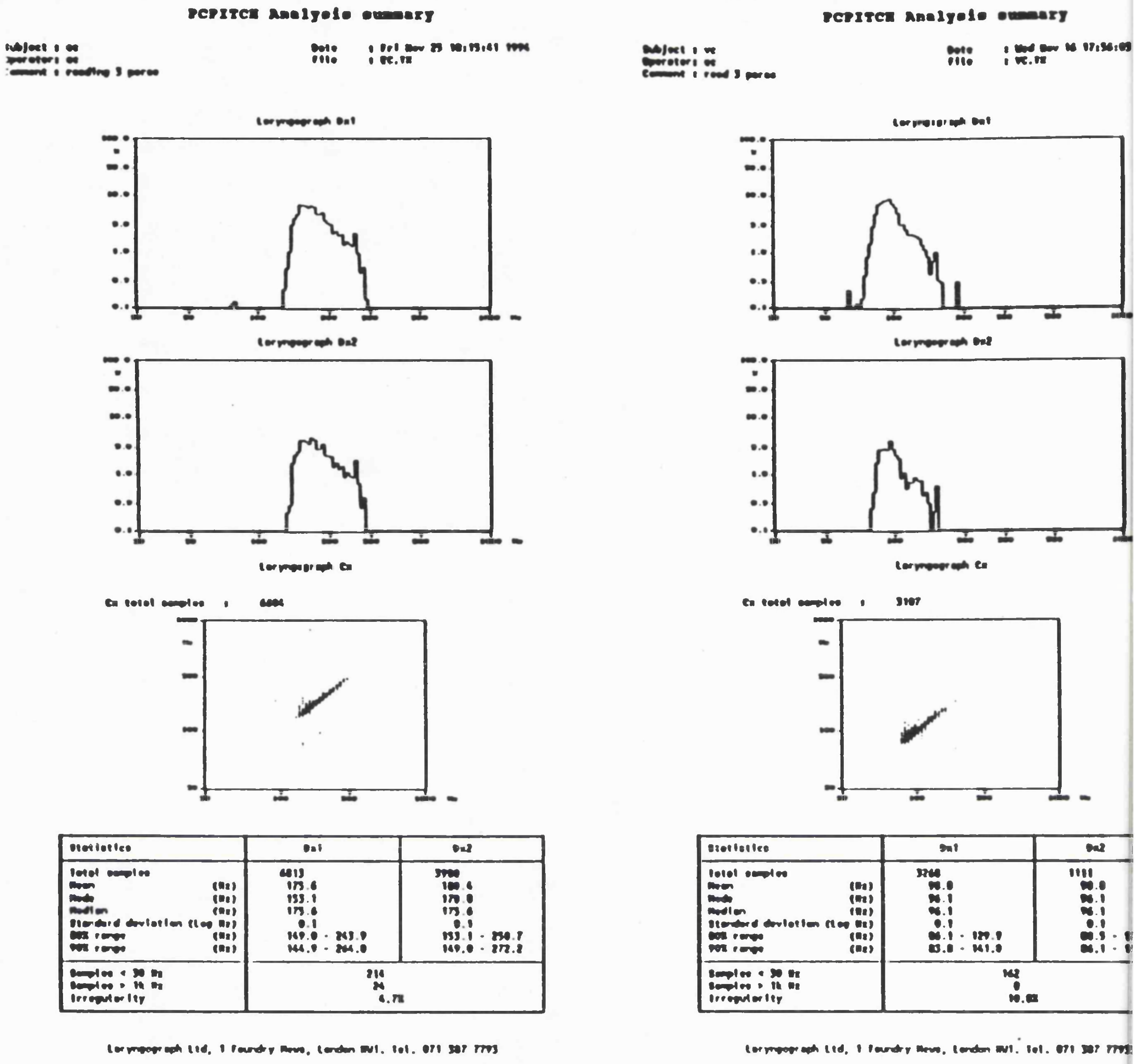
Figure 44

The range of frequencies within each 'bin' varies somewhat due to the use of a logarithmic scale. The range of frequencies within bins at the lower end of the scale is therefore smaller than at the top. Each 'bin' at the lower end of the distribution is more 'selective' about 'admitting' Fx values than at the top end. As the bins are incremented on the basis of the ocurrence of certain frequency intervals, the bins at the top end are incremented faster than at the lower end because of the more frequent occurrence of vibration at higher frequencies. This tends to bias the Dx plots towards the top (Ison, 1985).

Dx distributions are displayed in the form of first, second, and in the original analysis program, TPS, also third 'order' (Fig. 44). Second and third order distributions illustrate the amount of regularity of vocal fold vibration present in a sample. Second order includes only the instances, in the original total sample of Tx values, where two consecutive period values fell into the same 'frequency bin'. Third order plots include only adjacent triplets which fell into the same 'bin'.

In a more recent analysis program, PCLx, Tx values are admitted into 2nd order not on the basis of 'bin' range but on whether they differ by more than 10% from their preceding neighbours. This also eliminates irregularities from the distribution, but allows a maximum number of samples to be carried into second order and is less biased in favour of high pitched voices as the admission of samples is not related to bin width (Fig. 45 a and b).

Fx data related to the second order distribution rather than first order are often chosen to describe Fx characteristics of a recorded voice. Gross irregularities of the input data have been eliminated that may affect the calculations of Fx parameters. The irregularities may be part of the total recorded Tx sample, but not necessarily the result of vocal fold activity but of extralaryngeal movement, e.g. swallowing, or excessive gross laryngeal movement during recording. (Leff and Abberton, 1981, Comins, 1988, Ball et al, 1990, Barry et al, 1990 a,b, 1991, Ogle and Maidment, 1993).



a) Normal female voice

b) Normal male voice

Figure 45 - 1st and 2nd order Dx plots, PCLx software.

The shape of Dx distributions have been analysed by several researchers. Kurtosis scores from 2nd order Dx plots were calculated. These define the degree of 'peakedness' of the histogram. High kurtosis scores are related to auditory perception of a monotonous voice (Fig. 46) (Abberton, 1976, Leff and Abberton, 1981, Barry et al, 1990 a, b).

Leff and Abberton (1981) found a clear distinction between kurtosis scores for blunted (BS) and non-blunted schizophrenics (NBS) (Fig. 46) indicating that '*laryngograph recordings can be used as an objective clinical measure of blunting of affect*'. Besides, they found that a group of patients described as 'retarded depressives' (RD) had significantly higher kurtosis scores than either of three other groups of patients including the 'non-retarded depressives' (NRD).

Barry et al (1990 a,b) found systematic differences in kurtosis scores between text types. Subjects reading the 'Environmental Passage', had significantly lower kurtosis scores (i.e. flatter, broader Dx distribution) than in their reading of the 'Numbers Passage', a text including a listing of a lot of numbers, and both texts showed significantly lower scores than the free monologue, which was performed by subjects describing pictured material.

Another feature of the Dx histogram studied by Abberton (1976) and by Barry et al (1990 a,b,) is the degree of 'Skew' (Fig. 47) calculated as Mean-Mode/Standard Deviation. It may be positive, symmetrical or negative depending on where the speaker's modal value falls, in the lower, middle or upper part of the distribution. They found a tendency for Skew to be positive in reading because of more varied intonation patterns used, but there was no consistent tendency for skew related to speaker or task and they conclude that skew remains an inherently unstable measure.

Barry et al (1990 b) also found that mode, kurtosis and frequency range were more consistent than mean frequency in reflecting differences in frequency distribution resulting from differences in

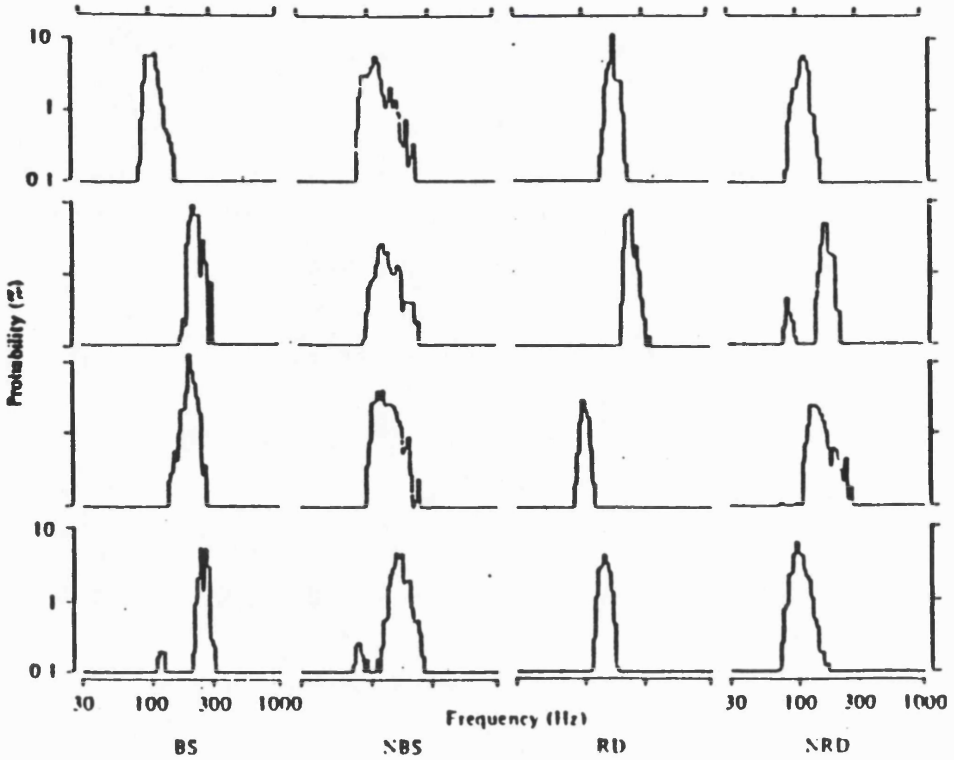


Illustration of Dx kurtosis in subjects with schizophrenia and depression. (From: Leff and Abberton, 1981)

Figure 46

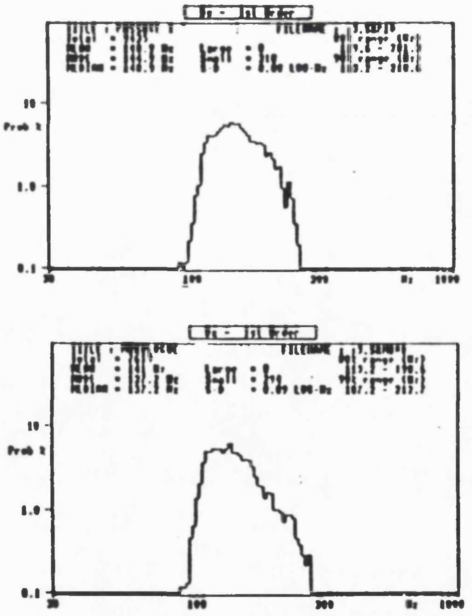


Illustration of degrees of Dx 'skew'. (From: Barry et al, 1990)

Figure 47

the speech material. For a given task they found speaker mode and kurtosis reliable individual measures that remained relatively stable over time.

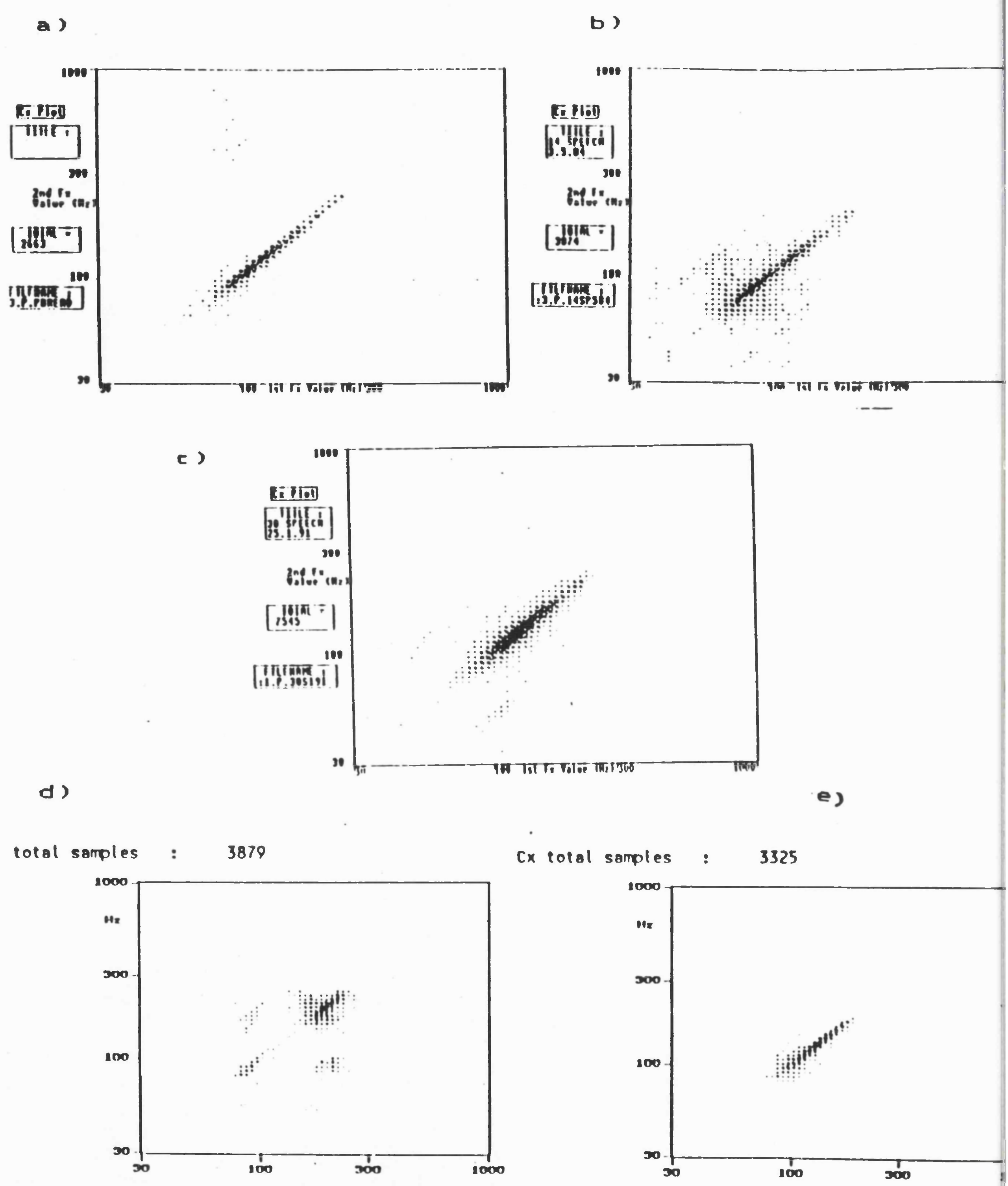
A study by Ogle and Maidment (1993) showed that measures of central tendency i.e. Mean and Modal fundamental frequency, and Fx Range of ELG recorded speech of mothers talking to children, were higher than when talking to another adult. The Mean was more affected than the Mode as it was the upper range of the voice that was raised in child directed speech.

iii Cross-plots - Cx.

Another way of illustrating regularity, or irregularity, in a recorded voice sample, apart from using 2nd or 3rd order Dx plots, is by means of a Cx plot or 'scatter plot' (Cross plot of excitation Fx) (Fig. 48). This plots adjacent pairs of Fx values against each other in 64 x 64 number of 'boxes' on the logarithmic fundamental frequency axes. Each 'box' illustrates the number of times a particular pair of Fx values occurred in a given order by the degree of darkness or density of the trace. The more frequent the occurrence of particular pairs of values, the higher the probability of transition between those two values and the greater the intensity (blackness) displayed on the plot (Ison, 1985).

Cx plots give a striking visual illustration of the amount of regularity or otherwise in a recorded speech or reading sample. The length of the diagonal reflects the range of frequencies used by a speaker. The plot is based on the raw data of the total sample and is comparable to the first order Dx plot.

A regular voice with a wide range will be shown as a well defined, long, dense diagonal (Fig. 48 a). Creaky voice common in males in particular is shown as a widening of the plot at the low frequency end (Abberton and Fourcin, 1984) (Fig. 48 b). A 'harsh' voice with irregularity throughout the frequency range, will be illustrated by a thick 'cigar shaped' Cx plot showing a diffuse, sometimes wide scattering of points along the diagonal (Fig. 48 c). Any frequency



Scatter plots (Cx)

Figure 48

a) Regular phonation
b) Creaky voice
c) Harsh voice
d) Fuberphonic voice before treatment
e) Same speaker as in d) after voice th

area showing reduced or absent vocal fold vibration, as in the puberphonic voice illustrated in fig. 48 d, will also be reflected in the Cx plot. The same voice after therapy is illustrated in Fig. 48 e.

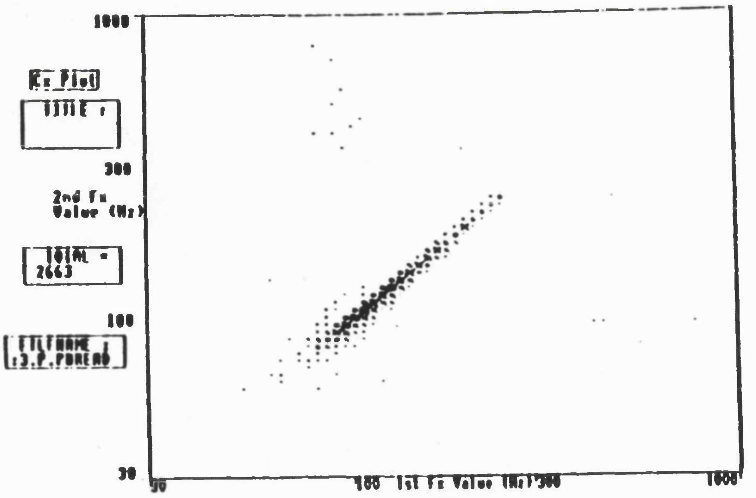
Barry et al (1990 a,b) studied Cx plots in detail and concluded that an offset of the diagonal towards the top left hand corner indicates a preponderance of sharply rising pitch movements i.e. rising intonation. An offset of the diagonal towards the bottom right corner indicates a tendency towards falling intonation patterns. They found significantly more falling intonation in men than in women (Barry et al, 1990 b).

iv %TS as a measure of degree of voice 'Regularity'.

Second and third order Dx plots (Fig. 44 and 45) also give an indication of the regularity of vocal fold vibration in a recorded voice sample. In the TPS - software program used for most voice analyses in this study, 2nd and 3rd order distributions only admit instances when two and three adjacent Fx samples fall into the same frequency 'bin'. A very 'rough' voice with a lot of irregularity of vocal fold vibration will show a considerable decrease in the 'Total Sample' carried into 2nd and 3rd order distributions (Fig. 49). There will, however, always be instances where adjacent Fx samples only vary by one or two Hz but happen to be allocated to adjacent 'bins' on the frequency axis and will therefore be excluded from second order. Some fundamental frequency variation, that will also reduce the carry-over of Fx samples into 2nd order distributions, is related to fast Fx variation associated with changing intonation contours of speech and reading.

An estimate of vocal fold Regularity of vibration (Voice regularity) can be expressed as the proportion of the 'Total sample' carried into 2nd order or %TS. (%TS = Sample total 2nd order/Sample total 1st order * 100). It has been found to be a useful clinical measure of voice quality in patients before and after surgical and vocal therapeutic intervention (Carlson, 1986, 1988, 1993 a and b) and in this study of voice quality after radiotherapy (Fig. 49 b).

a)



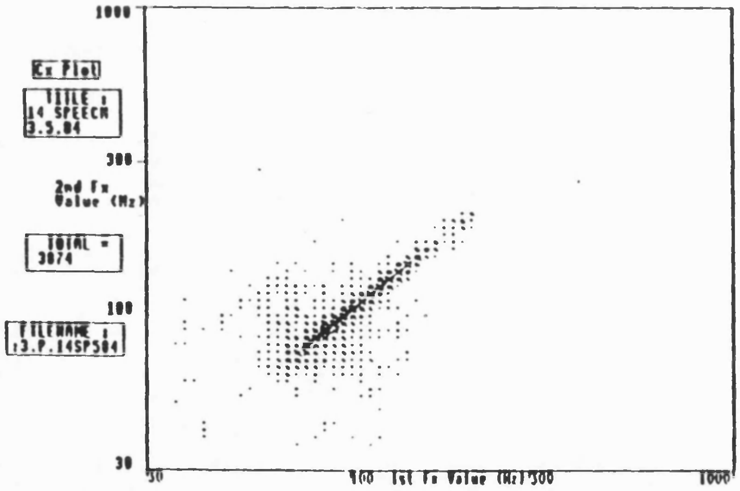
Normal voice

STATISTICS TABLE

TITLE :   FILENAME ::3.P.FOREAD

DISTRIBUTION TYPE	Dx 1st Order	Dx 2nd Order	Dx 3rd Order
SAMPLE TOTAL	2773	891 %TS=32.1	409 %TS=14.7
MEAN	119.6 Hz	122.9 Hz	126.4 Hz
MODE	107.2 Hz	126.4 Hz	122.9 Hz
MEDIAN	113.2 Hz	122.9 Hz	122.9 Hz
STANDARD DEVIATION	0.15 LOG-Hz	0.10 LOG Hz	0.10 LOG Hz
80% RANGE	88.5 180.3 Hz	93.5 180.3 Hz	93.5 190.6 Hz
90% RANGE	81.5 210.6 Hz	90.9 195.0 Hz	90.9 195.0 Hz

b)



Creaky, harsh voice

STATISTICS TABLE

TITLE : 14 SPEECH 3.5.84   FILENAME ::3.P.14SP584

DISTRIBUTION TYPE	Dx 1st Order	Dx 2nd Order	Dx 3rd Order
SAMPLE TOTAL	4147	863 %TS=20.8	332 %TS=8
MEAN	90.9 Hz	98.7 Hz	98.7 Hz
MODE	86.1 Hz	86.1 Hz	86.1 Hz
MEDIAN	90.9 Hz	96.1 Hz	98.7 Hz
STANDARD DEVIATION	0.12 LOG-Hz	0.10 LOG-Hz	0.09 LOG-Hz
80% RANGE	67.3 133.5 Hz	75.1 137.2 Hz	79.3 133.5 Hz
90% RANGE	62.0 148.0 Hz	71.1 148.0 Hz	75.1 144.0 Hz

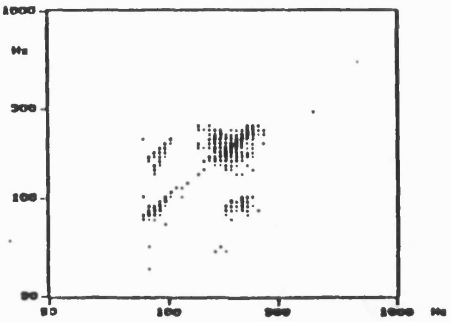
Fx statistics tables showing %TS into 2nd and 3rd order and corresponding Cx plots using the TPS software.

Figure 49

a) Puberphonic before voice therapy

Laryngograph Cx

Cx total samples : 3879

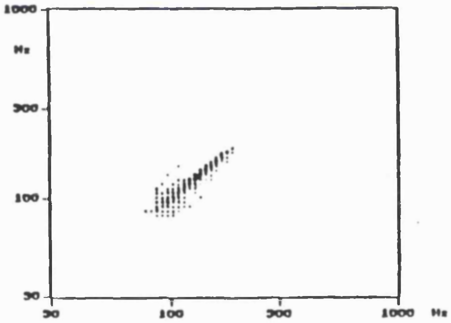


Statistics		Dx1	Dx2
Total samples		4183	1032
Mean	(Hz)	180.4	185.5
Mode	(Hz)	195.9	190.6
Median	(Hz)	190.6	195.9
Standard deviation (Log Hz)		0.2	0.1
80% range	(Hz)	119.6 - 224.7	175.6 - 218.6
90% range	(Hz)	91.0 - 237.3	116.4 - 230.9
Samples < 30 Hz		261	
Samples > 1k Hz		101	
Irregularity		35.0%	

b) Puberphonic after voice therapy

Laryngograph Cx

Cx total samples : 3325.



Statistics		Dx1	Dx2
Total samples		3522	1435
Mean	(Hz)	123.0	126.4
Mode	(Hz)	126.4	126.4
Median	(Hz)	126.4	129.9
Standard deviation (Log Hz)		0.1	0.1
80% range	(Hz)	104.3 - 153.1	113.3 - 161.7
90% range	(Hz)	98.8 - 166.2	104.3 - 170.8
Samples < 30 Hz		206	
Samples > 1k Hz		0	
Irregularity		7.8%	

Fx statistics tables showing
% Irregularity and corresponding Cx
plots using the PCLx software.

Figure 50

% TS is not a pure measure of jitter, for reasons explained above, but it will contain such Fx perturbation as a component. Normal male speakers have been found to carry over between 30-50% of their Total Sample into 2nd order (Fig. 49 a) (Kramer, 1989). Although fairly gross, %TS will be shown to differentiate a group of Normal speakers from a group of irradiated speakers in this study.

In the most recent software produced with the PCLx system a measure of **'% Irregularity'** is introduced (Fig. 50 a and b). This is the converse of %TS in that it reflects the proportion of Tx samples that differ from their adjacent neighbours by more than 10 % and are excluded from second order Dx distributions.

It was the acquisition of the Voiscope incorporating the Laryngograph, allowing non-invasive, objective clinical assessment of voice quality before and after intervention, which inspired the following study.

CHAPTER VIII - THE RADIOTHERAPY STUDY

A THE PILOT STUDY

i Introduction

Patients who had undergone radiotherapy for vocal fold carcinoma were chosen for study as they were not routinely referred for voice assessment or therapy but were expected to have a certain amount of difficulty with voice production after irradiation.

The answers to four questions were sought in the Pilot study:

a) Will ELG clearly demonstrate normalized vocal fundamental frequency with increasing time after radiotherapy?

b) Will ELG analyses reflect common characteristics or patterns of recovery in vocal function after treatment?

c) Is Lx sensitive to interference in vocal fold contact patterns resulting from recurring tumour.

d) Will Lx demonstrate improved vocal function after voice therapy?

Apart from objective measurement of vocal function, a questionnaire was devised to try to monitor patients' smoking habits before and after radiotherapy, the amount of talking they did during and after treatment and their 'quality of life' in terms of experienced vocal limitations and degrees of 'hoarseness' (Appendix 1 A). The therapist used the Vocal Profile Analysis (VPA) (Fig. 10, p. 83) by Laver, Wirz and MacKenzie (1986) for perceptual evaluation of voice quality other than fundamental frequency parameters. The questionnaire and perceptual evaluation were aimed at enabling comparisons with the studies by Stoicheff (1975), Mendonca (1975), Karim, Snow, Diek and Hanjo (1983), Riska and Lauerma (1966) and Harwood and Rawlinson (1983).

Permission to carry out the study was sought from the senior consultants in the departments of Oncology and Otolaryngology (ENT). Ethical approval was not considered neccessary, as the assessments carried out were not invasive, and the purpose at this stage was not randomisation for different treatment regimes. The ELG assessments and questionnaire responses were seen as an extension of the routine evaluation and review procedures for this patient group.

ii Subjects.

The original protocol was more ambitious in its aims than turned out to be possible to realize. It proposed a three part study; firstly investigating subjects any time post radiotherapy; secondly assessing subjects before, during and after radiotherapy to determine at which point the voice was considered 'normal' as measured by ELG and as rated by therapist and patient and to what extent this was related to smoking habits and amount of talking during radiotherapy; thirdly, one group of subjects was to be randomly allocated to a period of voice therapy and advice on voice conservation (Fex and Henriksson, 1966) or to a group which would act as a control and be assessed at various times post radiotherapy.

As the hospital was not a major centre for treatment of head and neck cancer, the number of patients referred was not large and criteria for referral were therefore not stringent to allow the maximum number of subjects to be assessed. Any patient irradiated for laryngeal carcinoma was seen irrespective of age, sex, tumour stage or time after completion of radiotherapy.

Subjects were allocated a number in the order of referral. The subject numbers will be referred to throughout the text. Subjects 1-14 were referred between June 1983 and June 1984. The original aims were by this time considered to be unrealistic and too ambitious for a pilot study, but as some interesting findings were emerging, it was decided to carry on data collection for another few months. Subjects 15-21 were referred to the Pilot study between 17.8.84 and 28.2.85 at which point around 50 recordings had been carried out. Tumour stages and numbers of subjects in each diagnostic group is shown in Table 10below.

Table 10

Stage	**T1**	**T2**	**T3**	**T4**	**Supraglottic**	**Total**
N	10	7	1	1	2	21

Table showing the number of subjects and tumour stages represented in the Pilot study.

Throughout this text and in all figures the time of individual subjects' assessments will be expressed as 'months post radiotherapy' abbreviated to **'MPRx'**.

For details of individual subjects' age, tumour stage, site of lesion, histology and fractionation see Appendix 2. Most subjects in the pilot study had received treatment in 3 fractions per week over a period of approximately one month or less. Subject 1 with a poorly differentiated supraglottic tumour had been treated in 6 fractions in hyperbaric oxygen. Subject 12 with a T2 invasive tumour of the left vocal cord with right subglottic extension had received a split course of radiotherapy.

Patients were asked if they were willing to take part in the study, told about the purpose and what it would entail and that every effort would be made to record them on the occasion of their regular follow up appointments in the Joint clinic. No patient refused an initial assessment, but some were not enthusiastic about further requests for recordings and questionnaire responses. The possible reasons for this reluctance will be discussed.

All subjects were offered an explanation of laryngeal structure and function using diagrams of the vocal tract and a model of the larynx. They were also given advice regarding vocal hygiene and voice conservation measures according to their perceived and self expressed needs.

Two subjects were eventually excluded from the pilot study: Subject number 5's objective measures, as the original recordings were erased by mistake, and subject 9 as this was the only woman referred to the pilot study and the intention was to look at the group as a whole in terms of fundamental frequency data.

iii <u>Subjects referred before radiotherapy.</u>

Three subjects, 13, 16 and 20 were referred and assessed before, during and after radiotherapy (Appendix 3 A and B).

Subject 13, with a T2 tumour of his right vocal cord, experienced severe reaction to the treatment. He resumed smoking 20 cigarettes per day one week after the end of treatment. Two weeks after his 3 MPRx assessment in June 1984 (Appendix 3 A and B), he developed perichondritis for which he received treatment with steroids. Subject 13 was also on continuous asthma medication using inhalers. He attended ENT appointments erratically througout the year, blaming this on pressure of work.

Subject 16 had a two month history of hoarseness before the diagnosis of a T1 tumour of the right vocal fold. His recordings after the end of radiotherapy were unfortunately erased by mistake, but he reappears in the subsequent 'Main study'.

Subject 20 had originally been referred for voice therapy in September 1984 after a six year history of voice problems. He had vocal nodules removed four years previously resulting in a normalisation of his voice, which had however started to deteriorate again two years later. He had received a 15 month course of voice therapy with limited improvement. He had two biopsies, one in July 1984, which did not show any malignant changes but 'reactive and hyperplastic' epithelium.

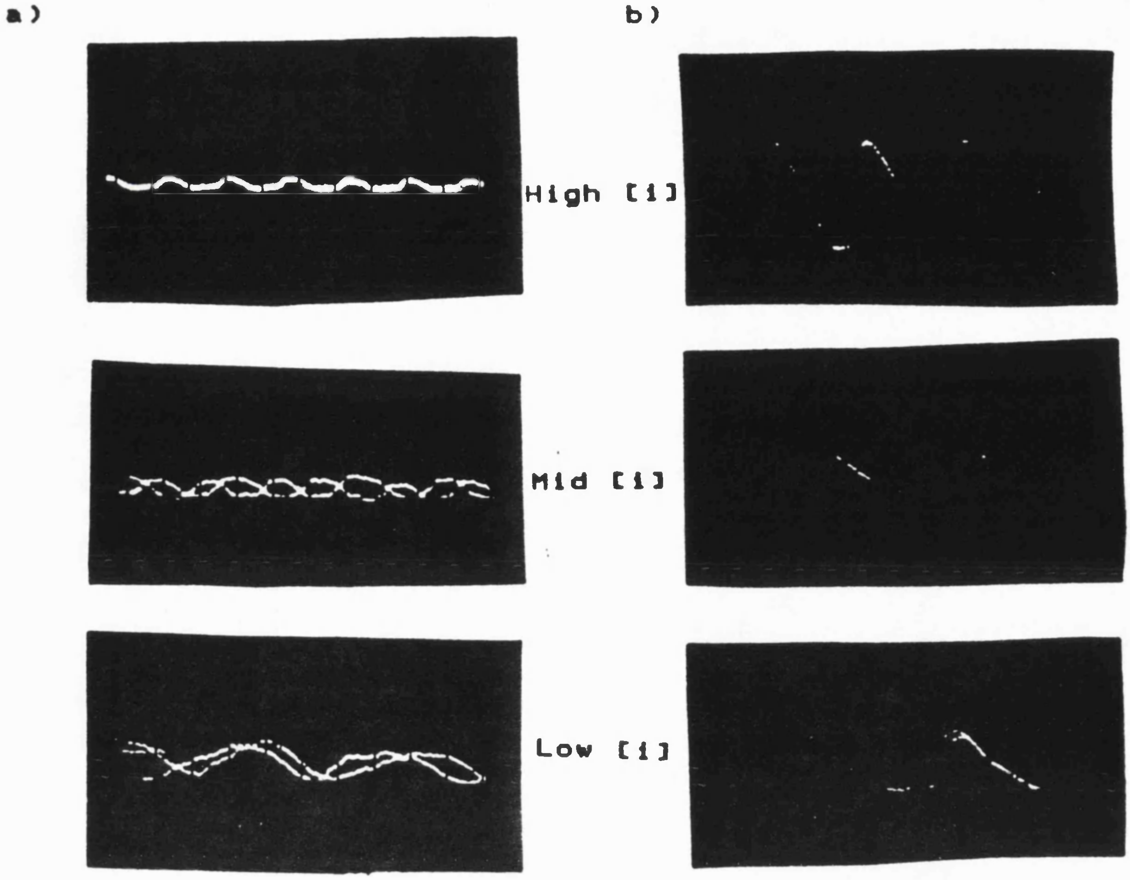


Figure 51 - Subject 20 - Lx waveforms
a) before diagnosis
b) after biopsy

After a course of voice therapy in September - October 1984, which resulted in some limited improvement in vocal function, Lx waveforms still indicated severe limitation and irregularity in vocal fold contact patterns (Fig. 51 a) and he was referred back for review in the ENT department. A large area of leukoplakia was observed on the anterior 2/3 of his right vocal fold. This was biopsied and found to show malignant changes corresponding to a stage T1 tumour.

Subject 20 was encouraged to carry on using his voice as he had been taught in voice therapy with good breath support and much reduced laryngeal tension. He was reassessed after the biopsy just before the beginning of radiotherapy, when his voice was found to be much improved. This is confirmed by his Lx waveforms (Fig. 51 b) and his fundamental frequency and %TS measurements. (Appendix 3 A/B 'Pre'). He continued to be followed up for the Pilot study at intervals up to 14 MPRx and also reappears in the 'Main study'.

iv Assessment occasions.

The intention was to assess patients, when they attended their regular review appointments in the joint ENT and Oncology clinic held on the second and fourth Friday morning of every month. This would ensure assessments and reassessments at approximately the same time of day, between 10.30 a.m and 12.30 p.m., at monthly, three, six and twelve monthly intervals after radiotherapy (Months post radiotherapy = MPRx). When there had been no sign of recurrence of the tumour ten years post radiotherapy, patients were discharged from the joint clinic. Some patients with recurring voice problems have continued to attend the clinic for longer which will be illustrated in the 'Main study'.

This was a clinical study where no dedicated time was available for data collection, as a result, regular reassessments turned out to be more difficult to achieve than anticipated. Patients failed, cancelled or changed appointments, or were in a hurry to get away. The therapist could not always be present in the joint clinic and patients were therefore 'lost' to follow up or not referred until the therapist was present on a later occasion. An attempt to get patients to attend separate additional appointments for the purpose of recording met with poor attendance rates. The reasons for this will be discussed.

v Self assessments of voice quality and vocal function.
The Questionnaire.

Based on the findings in the studies by Stoicheff (1975), Karim et al (1983) and by Harwood and Rawlinson (1984) a questionnaire was constructed for this study, including items that their patients reported difficulties with. The questionnnaire presented to the Pilot study subjects is shown in Appendix 1 a. 'Q' in the following text refers to the numbered questions in this.

Our study assessed the effects of irradiation for mainly early, highly curable, laryngeal tumours. The focus is therefore, on 'quality of life' in the 'narrow' sense of subjects' experiences of symptoms of hoarseness and limitations in voice function. i.e. the effect on quality of life in terms of their ability to communicate.

The rating scale used is a kind of linear analogue scale similar to that described by Llewellyn - Thomas et al (1984) (Fig. 40, p. 194), but here providing seven scale points at equal appearing intervals to facilitate scoring. A similar scale has been used by Schipper, Clinch, Mc Murray and Levitt (1984) in their Functional Living Index: Cancer (FLIC), which is an instrument for rating quality of life of cancer patients. In our scales a low rating indicates none or little difficulty, hoarseness or experienced limitations in voice use. A rating of '7' indicates severe problems.

Different occupations demand different amounts of talking. The oedema and laryngeal discomfort caused by radiotherapy sometimes leads to complete loss of voice during the latter stages of the treatment and may or may not prevent subjects from going to work. Patients who may need to use the telephone, address meetings and whose 'tool of their trade' is their voice, are dramatically affected by the temporary, maybe even permanent change and limitation in voice function during and after treatment. It may even lead to a need to stop working or to change jobs. Some early studies reported less favourable self ratings of voice quality in patients who returned to work compared to those who did not (Riska and Lauerma, 1966). Therefore we requested information from the patients on their work history (Q. 1 a,b).

Less good voice quality was reported by Karim et al (1983) in subjects who did not stop smoking. The questionnaire asked the subjects to report on their smoking habits (Q.2 a and b) before and after radiotherapy.

We attempted to tap the tendency for subjects to perceive their voice quality to be better than objective measures (Lehman et al,1988) or clinicians ratings suggest (Mendonca, 1975, Karim et al, 1983) by asking subjects to rate their voices in terms of 'hoarseness' (Q. 3) and whether the voice posed a 'Problem'or not (Q. 5).

We hoped to get an indication of the amount of 'daily voice use' before and after radiotherapy by questions 4 and 15.

Patients' ratings of five voice 'functions'; 'voice tiring after a lot of talking', 'ability to sing', 'ability to shout', 'ability to talk over noise' and 'ability to use the telephone' (Q. 6-10), were combined and averaged into what we call a 'Mean Problem-score' or 'Mean P-score'. Not everybody felt all the items applied to them, particularly not 'ability to sing or 'hum' a tune. (cf. Table 9 below, p. 196, Llewellyn -Thomas et al, 1984).

The questionnaire also requested subjects to rate how much talking they did while having radiotherapy (Q. 14), and how long the voice took to recover after treatment (Q. 15). Stoicheff (1975) found a tendency for the voice to recover more quickly when subjects had reduced the amount of talking they did during radiotherapy. The last question (Q. 16) asks the patient to rate the time the voice took to recover or whether it has not.

A point regarding the construction of 'quality of life' questionnaires is raised by Fallowfield (1990), when she suggests that the time frame within which the patient is asked to recall a symptom or function should be as short as possible, as this will otherwise influence the reliability of a rating. In our questionnaire (Appendix 1 a) questions 13-16 go against this rule. They require subjects to recall, in some cases what had happened several years ago. In 'The Main Study' subjects were therefore not required to respond to these questions but only to the ones that ask them to report the current state of voice quality, use and function.

If a method for self assessment is going to be of use in evaluation of 'quality of life' in terms of vocal symptoms and function, it must be simple, easy to understand and not take excessive amounts of the subjects ' time to respond to (Llewellyn-Thomas et al, 1984). The vast majority of the pilot study subjects seemed to have little or no difficulty in filling in the questionnaire, although there was occasional reluctance, when it was presented at several different times post radiotherapy. Most subjects did however not seem to mind once they understood the purpose of the exercise.

The self assessment must also be reliable in the sense that "...*it should produce results with little random error when applied repeatedly to subjects whose clinical state is not expected to change between assessments*". It should be valid in the sense that it measures attributes, here of aspects of voice quality and function, that "...*correspond to other methods of assessing the voice, that change in an appropriate way with known clinical changes and shows a gradient in association with known differences in disease status.*" (Llewellyn-Thomas et al, 1984).

In this study the reliability or validity of the rating scales have not been formally tested, but their construction is based on previously tried scales, as reported. As subjects were seen and asked to rate their voices at a great variety of times after radiotherapy, a few also before and immediately after the end of treatment, we will illustrate the differences in experiences of voice quality and function in groups of subjects at different times using scatter plots of their test and retest ratings and of individual ratings as a function of MPRx (Appendix 6 A, B and C).

vi Voice samples.

The speech tasks chosen for recording and ELG analysis on each occasion were the following:

a) Conversation, approximately 2 minutes, on a 'neutral' topic e.g. spare time activities, holidays.

b) Reading aloud a standard passage 'The North wind and the sun' (113 words) (Appendix 4).

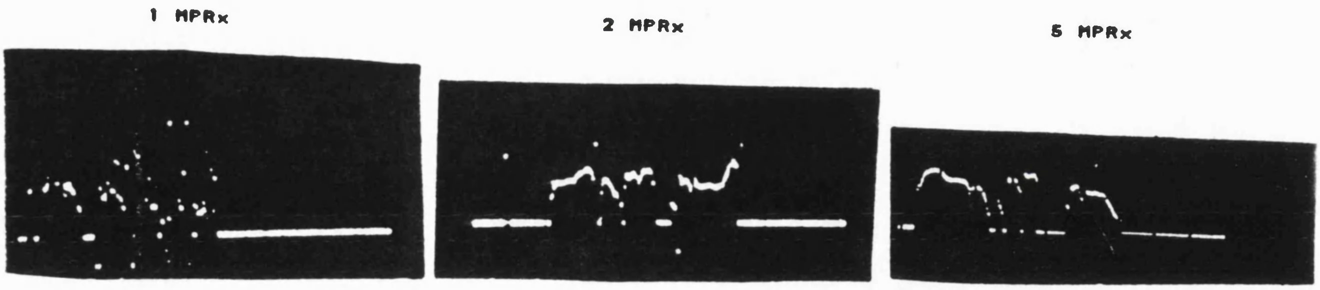
c) Sustained phonation on [i] at comfortable volume at mid, high and low pitch. The Lx waveforms produced at this stage had to be photographed 'live' off the oscilloscope screen (Fig. 51 a, b, p. 218).

The vowel [i] was chosen as it had been found to be most affected by degrees of "hoarseness" (Cooper, 1967, Bradley, 1985, Netsell, 1989).

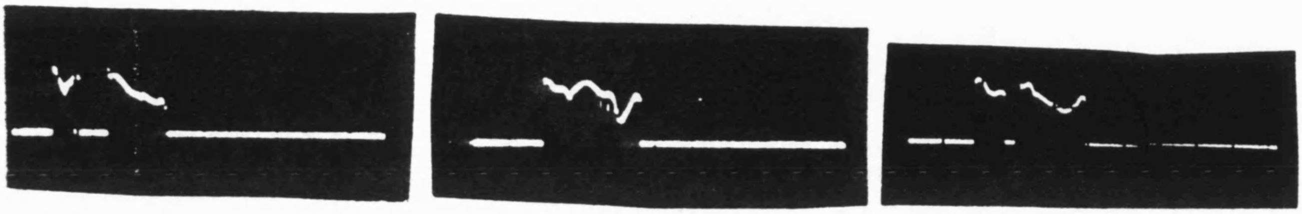
d) Repetition of three sentences after the therapist

"My name is..."
"Hello, how are you?"
"I'm fine, thankyou"

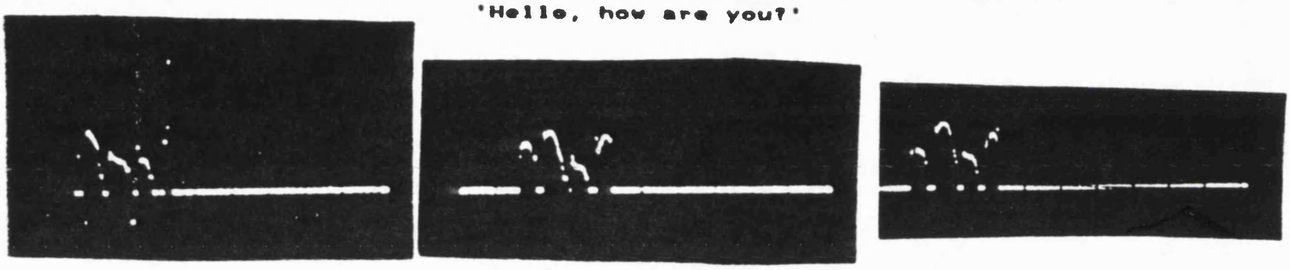
The aim of the latter was to enable qualitative evaluation of fundamental frequency contours of the same utterances at different times after radiotherapy for their 'smoothness' and regularity (Fig. 52). The Fx contours could be replayed at a later date, 'frozen' on the oscilloscope screen and photographed



'My name is...'



'Hello, how are you?'



'I'm fine, thankyou'

Figure 52 - Subject 8 - Fx contours at 1, 2 and 5 months post radiotherapy

Each recording took approximately 15 - 20 minutes. The subjects were asked to fill in the questionnaire (Appendix 1 a) on each occasion to reflect subjective experiences of hoarseness and limitations in voice function. Occasionally subjects asked to take the questionnaire away and post it back filled in. They hardly ever did, resulting in occasions where there are no self ratings corresponding to the objective and perceptual assessments.

vii Result of the ELG analysis.

Some results of the Pilot study were reported in an informal paper circulated to the referring consultants and later presented in a revised form at a national conference (Carlson, 1988 b).

The group of subjects was small (N=21), subjects were not selected at random and represented all tumour stages including two supraglottic tumours (Subject 1 and 19, Table 10, p. 216). The intervals after radiotherapy that subjects were assessed (MPRx) were also extremely wide and varied (Appendix 3 A and B). Because of these methodological problems, statistical analysis of the data was not attempted. A descriptive summary of some of the findings will now follow.

The graphs in Appendix 3 A and B illustrate the variation in a) regularity of vocal fold vibration expressed as %TS and b) 2nd order Mean fundamental frequency, as a function ot time post radiotherapy, (MPRx).

Crosses in the graph indicate subjects with T1 tumours, circles subjects with T2 tumours and squares subjects with more advanced tumours. Individual subjects are identified by their numbers on each assessment occasion.

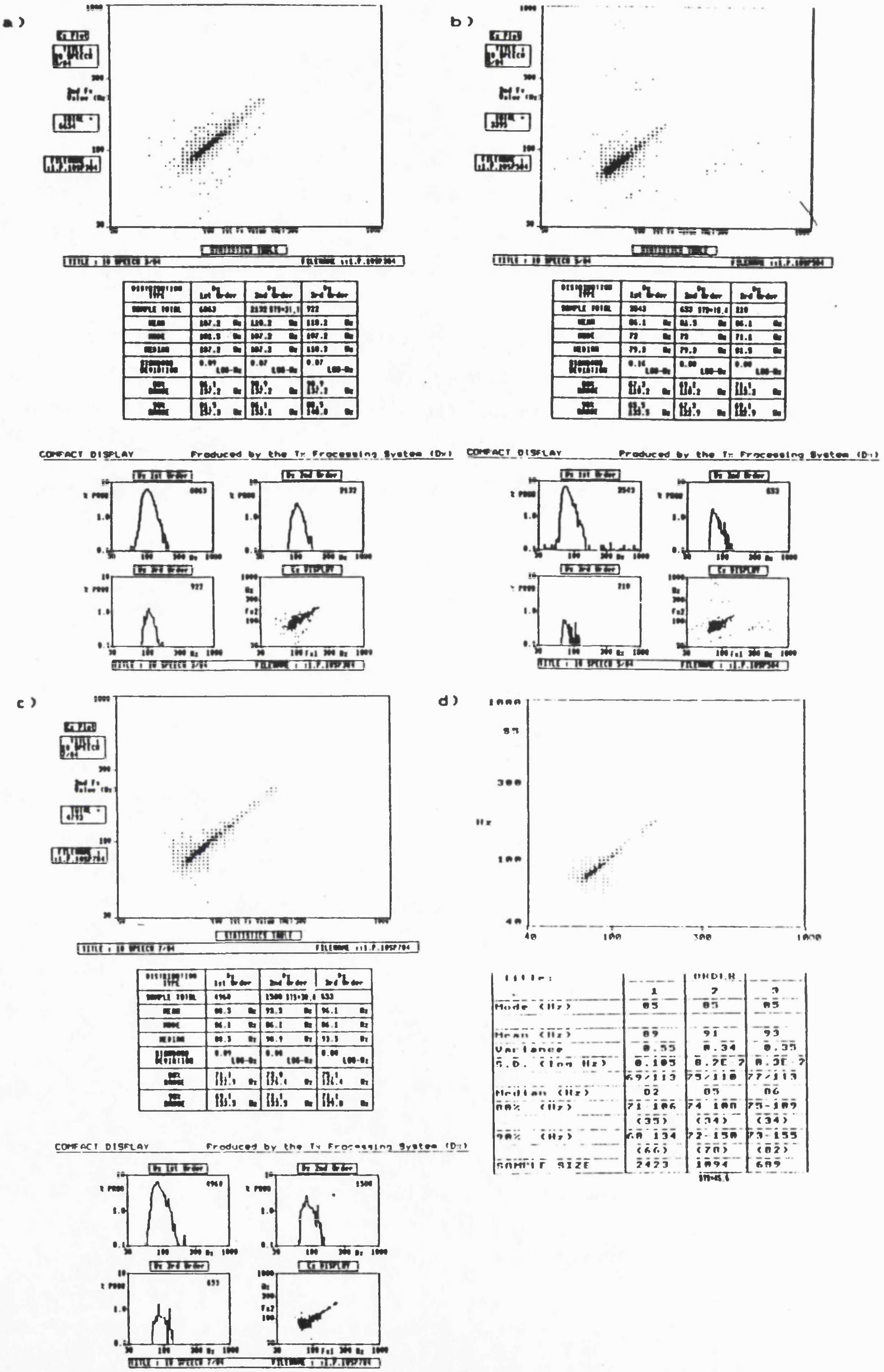
As the age range is very wide in our sample (39 - 77 years) the norms for Mean fundamental frequency for males within this age range reported in Table 1 below (p. 52) vary between 100 - 119 Hz for speech (increasing after the age of 60) (Mysak, 1959), and between 108 - 127 Hz for reading (Mysak, 1959, Hollien and Shipp, 1972, Pegoraro-Krook, 1988, Brown, Morris et al, 1991).

In a 'Normal speaker' sample of ten males aged 50-75, drawn at random from Kramer's subjects (1989), the average 2nd order Mean fundamental frequency for Speech was found to be 121.9 Hz with anaverage of 35 % TS into 2nd order. In Reading aloud the average fundamental frequency was 119.8 Hz and regularity 33 % TS. The levels of these 'normative' values are indicated by the solid lines drawn across the graphs, marked 'Normal average', in Appendix 3 A and B. It will be demonstrated that the differences in Mean fundamental frequency between the Normal speakers and the T1 subjects do not reach significance. The regularity measure, however, is significantly different for both the speech and reading tasks.

Inspection of the plots reveals a tendency towards a general lowering of the fundamental frequency for speech and reading during radiotherapy and in the first few months afterwards compared to Pre-radiotherapy recordings (Subjects 13, 16 and 20) and compared to the 'Normal Average'. Major changes in Fx and %TS seem to occur mainly in the first 2-6 months, after which the Mean Fx seems to stabilize but stay below 100 Hz for Speech and below 110 Hz for Reading with increasing time post radiotherapy. ~~The previously observed higher Fx in reading applies to irradiated speakers as well as to 'normal' speakers (Appendix 3 A/B).~~

There does, however, seem to be some fluctuation in fundamental frequency after 2 MPRx in some subjects. Subject 10, whose diagnosis had been a T1 tumour of the anterior two thirds of his right vocal cord, shows a marked reduction in Mean Fx and a reduction in %TS between 2 and 5 MPRx, more marked in conversation that in reading (Appendix 3 A/B) with a subsequent increase in %TS and levelling off in Fx at 7 and 10 MPRx (Fig. 53 a-d). He had been advised to avoid clearing his throat, which he experienced as very dry, and to increase hydration. At 5 MPRx he complaned his voice fluctuated and 'tired' after a lot of talking.

He was seen for four voice therapy sessions before the next assessment at 7 MPRx. On this occasion he complained of feeling depressed and of having to make a conscious effort to keep the pitch of his voice raised. However, he is quite successful at doing this , with beneficial effect on his voice regularity measure, %TS (Fig. 53 c). He was found to be suffering from myxoedema, treated with thyroxine and felt better on the occasion at 10 MPRx. This may be reflected in his increased %TS for Speech and stabilising of his Fx in both speech and reading which follows the same pattern over time in both tasks (Appendix 3 A/B, Fig. 53 d).



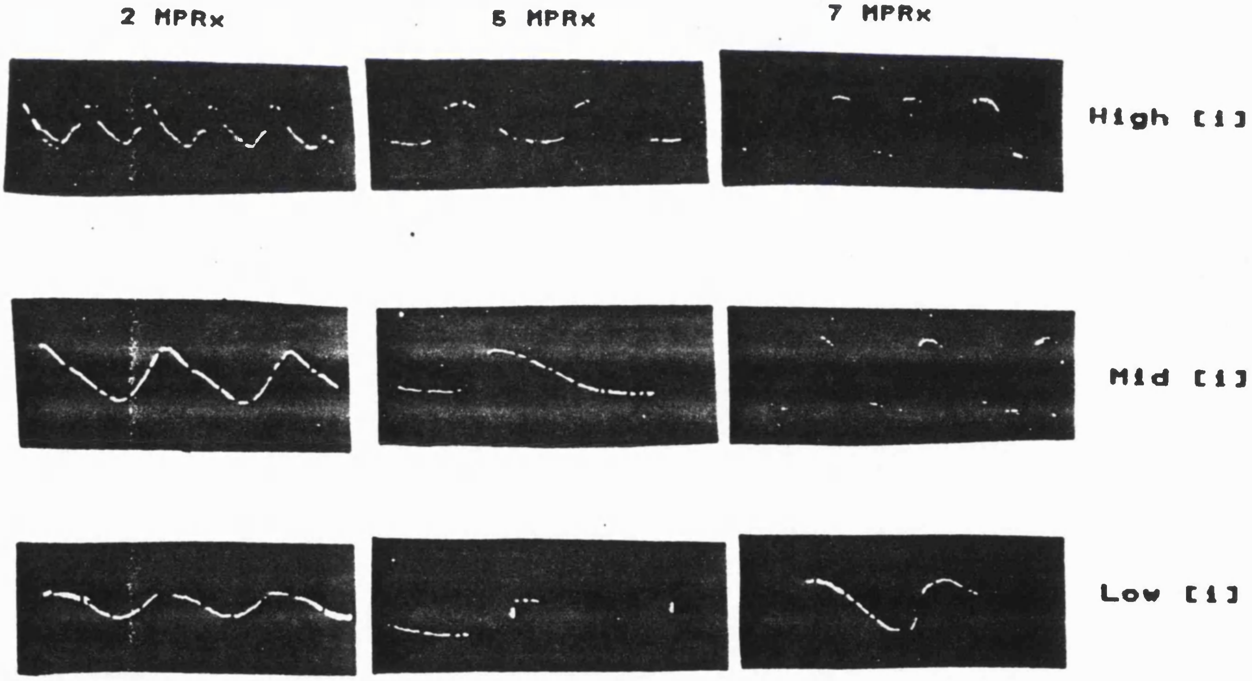
Title:	ORDER 1	2	3
Mode (Hz)	85	85	85
Mean (Hz)	89	91	93
Variance	0.55	0.34	0.35
S.D. (log Hz)	0.105 69/113	0.2E-2 75/110	0.3E-2 77/113
Median (Hz)	82	85	86
80% (Hz)	71-106 (35)	74-108 (34)	75-109 (34)
90% (Hz)	68-134 (66)	72-150 (78)	73-155 (82)
SAMPLE SIZE	2423	1094	689

Subject 10 - Speech, Dx and Cx plots and Fx parameters a) 2 MPRx b) 5 MPRx c) 7 MPRx and d) 10 MPRx

Figure 53

The same reading passage was used on all occasions and %TS does not show quite such a dramatic increase on the last occasion, but is at its lowest at 5 MPRx when both Speech and Reading Fx have dropped dramatically. This corresponds to a high degree of 'creaky' voice rating by the therapist on the VPA on this occasion (Appendix 5, Table 1), which may be a confirmation of Wendahl's (1963) finding that aperiodicity in a low pitched voice is perceived as more abnormal than in a higher pitched voice.

Lx waveforms (Fig. 54) at 2, 5 and 7 MPRx show a gradual improvement in speed of vocal fold closure (steepness of the closing slope) and in increased duration of the closed phase over time. This subject was reassessed some years later for the Main study.



Subject 10 - Lx waveforms at 2, 5 and 7 MPRx

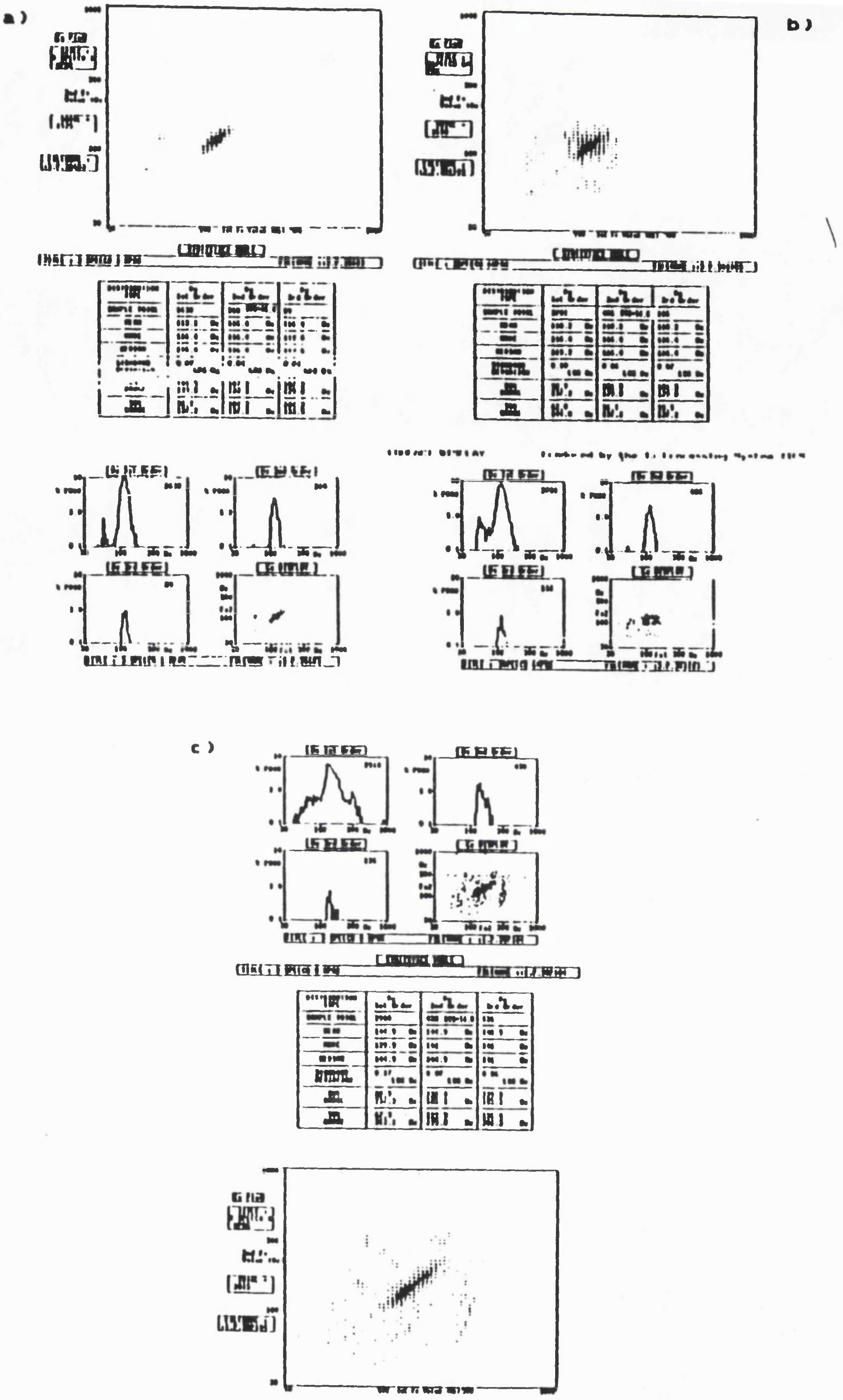
Figure 54

Another subject, who shows major fluctuation in Fx after 2 MPRx from his second assessment at 6 MPRx to his assessment at 8 MPRx, is Subject 3, who had a T2 tumour of his right vocal cord with subglottic extension. There is an abnormal increase in fundamental frequency between assessment occasions as shown in (Appendix 3 A/B, Fig. 55). He was found to have a recurrence of his tumour and had a hemilaryngectomy a few months later. The likely reason for his increased Fx measurements with recurrence of the tumour is an increased stiffness of the affected vocal fold (cf Hirano et al 1991).

Lx waveforms illustrate a change in vocal fold closure patterns between 1 and 6 MPRx (Fig. 56). Initially (1MPRx) there is poor control of pitch and the frequency is higher at attempted mid [i] than at either high or low [i]. This is put down to the subject's low intensity of phonation at this early time post radiotherapy. There is a short duration closure and poorly defined start of the closing phase. The latter is better defined on the second occasion (6 MPRx) and there is improved pitch control but an abnormal variation in amplitude (shimmer) between successive cycles at mid and high [i], indicating variability in the amount of vocal fold contact from one cycle to the next. On the last assessment occasion at 8 MPRx, high pitch results in a very distorted waveform (Fig. 56).

Subject 3 complained, on the occasion at 6 MPRx, of a deterioration in voice quality over the past few weeks. This was reflected in his self ratings on the questionnaire. His 'Hoarseness rating has gone up from 4 to 5/7, 'Problem' from 1 to 4/7 and his Mean P-score from 2.8 to 4.8/7. All these ratings have increased further by 8 MPRx when recurrence had been diagnosed (Appendix 6 C).

Subject 15 with a T4 tumour shows a slight decrease in Speech Fx from 4 to 7 MPRx (Appendix 3 A/B); this is however accompanied by a dramatic decrease in voice regularity on the second occasion as indicated by %TS. On this occasion indirect laryngoscopy revealed a fixed right vocal fold and the subject died a few months later from recurring tumour.



Subject 3 - Speech, Dx and Cx plots and Fx parameters a) 1 MPRx b) 6 MPRx c) 8 MPRx

Figure 55

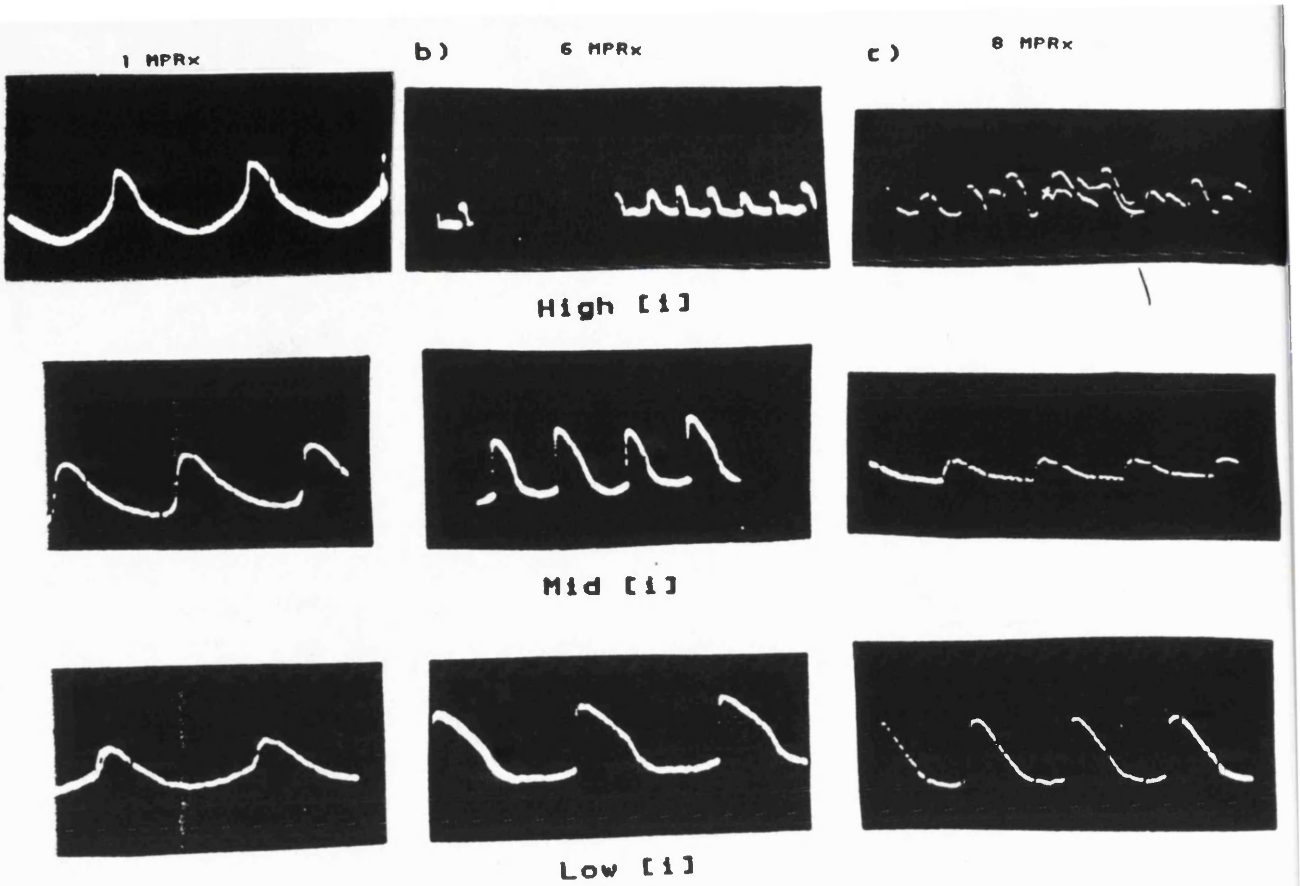


Figure 56 - Subject 3 - Lx waveforms

a) 1 MPRx b) 6 MPRx c) 8 MPRx

Subject 19 had a supraglottic tumour and was only assessed on one occasion 6 MPRx (Appendix 3 A/B). Three months later his tumour recurred and he had a total laryngectomy.

Three subjects referred to our study had completed their radiotherapy more than four years previously.

Subject 1 had radiotherapy for a poorly differentiated supraglottic tumour, not involving the vocal folds. He complained on the first occasion of a feeling of dryness ahd 'strangulation' when speaking. As he was a sales representative his job involved a lot of talking, often on the telephone and sometimes in a smoky atmosphere, which he felt aggravated his problem.

He was advised to reduce volume when speaking, which would help reduce laryngeal effort. He was also advised to ensure the mouthpiece of the telephone was close to his lips to avoid having to speak loudly. His Mean P(roblem)-score on the self assessment questionnaire was reduced from 5/7 to 4.2/7 on the second occasion on the grounds of him finding telephoning much less of a problem following the advice given (Rating on this item is down from 5/7 to 2/7).

He shows remarkably consistent Speech fundamental frequency values over a period of a year when he was assessed on three occasions, at 50, 55 and 62 MPRx, and highly regular vocal fold vibration as evidenced by high %TS. In reading aloud there is a consistent increase in Fx between occasions accompanied on the last occasion by a considerable increase in %TS (Appendix 3 A/B).

Subject 2 was seen for his first ELG assessment 53 months post radiotherapy for a poorly differentiated T1 tumour involving the anterior left vocal fold and the anterior commissure. He had severe residual problems with his voice. A year after the end of radiotherapy an anterior glottic web had been observed on indirect laryngoscopy. He had withdrawn from social situations and given up work as a result of his voice problem.

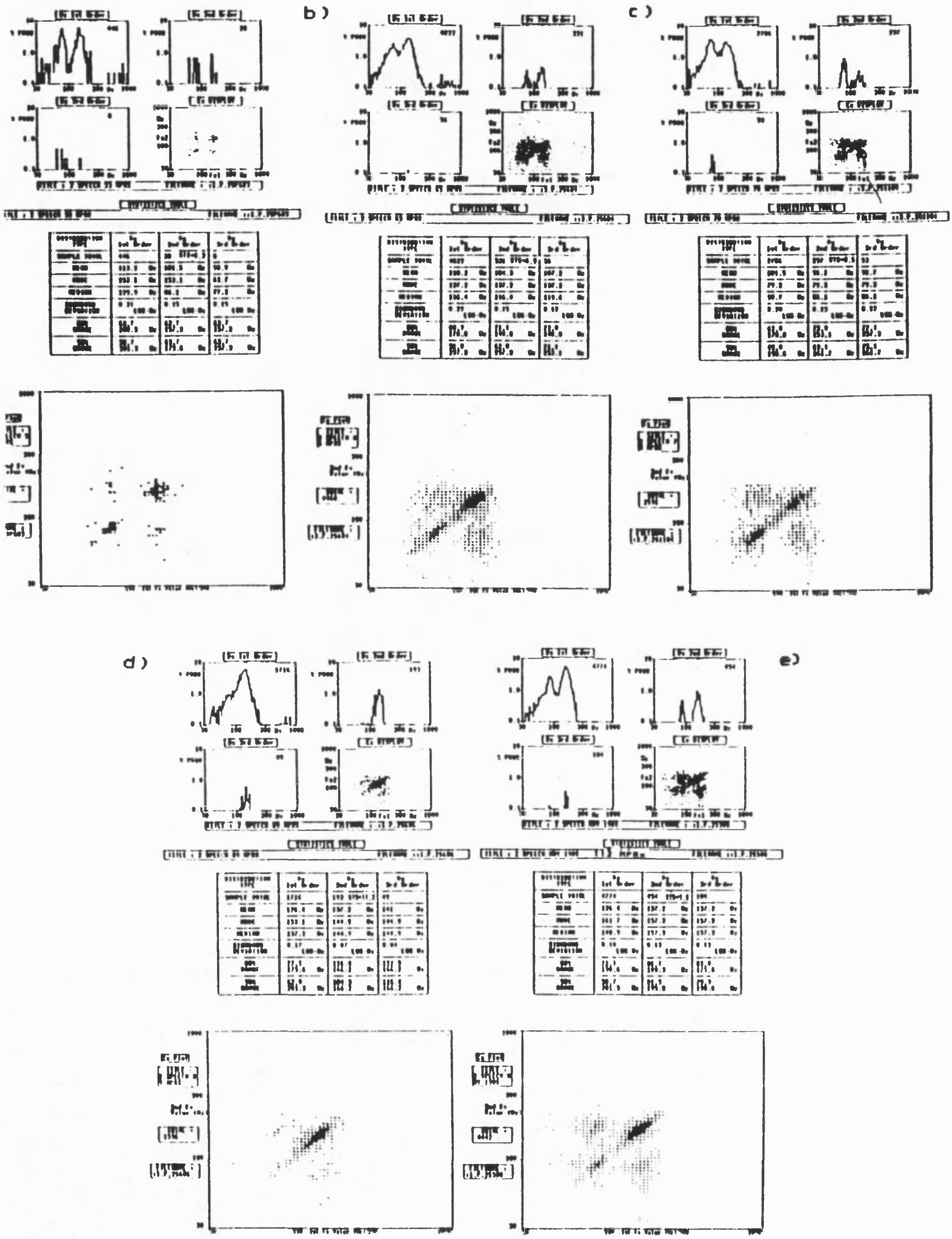


Figure 57 - Subject 2 - Speech, Dx and Cx plots and Fx parameters a) 53 MPRx b) 65 MPRx c) 70 MPRx d) 89 MPRx e) 112 MPRx

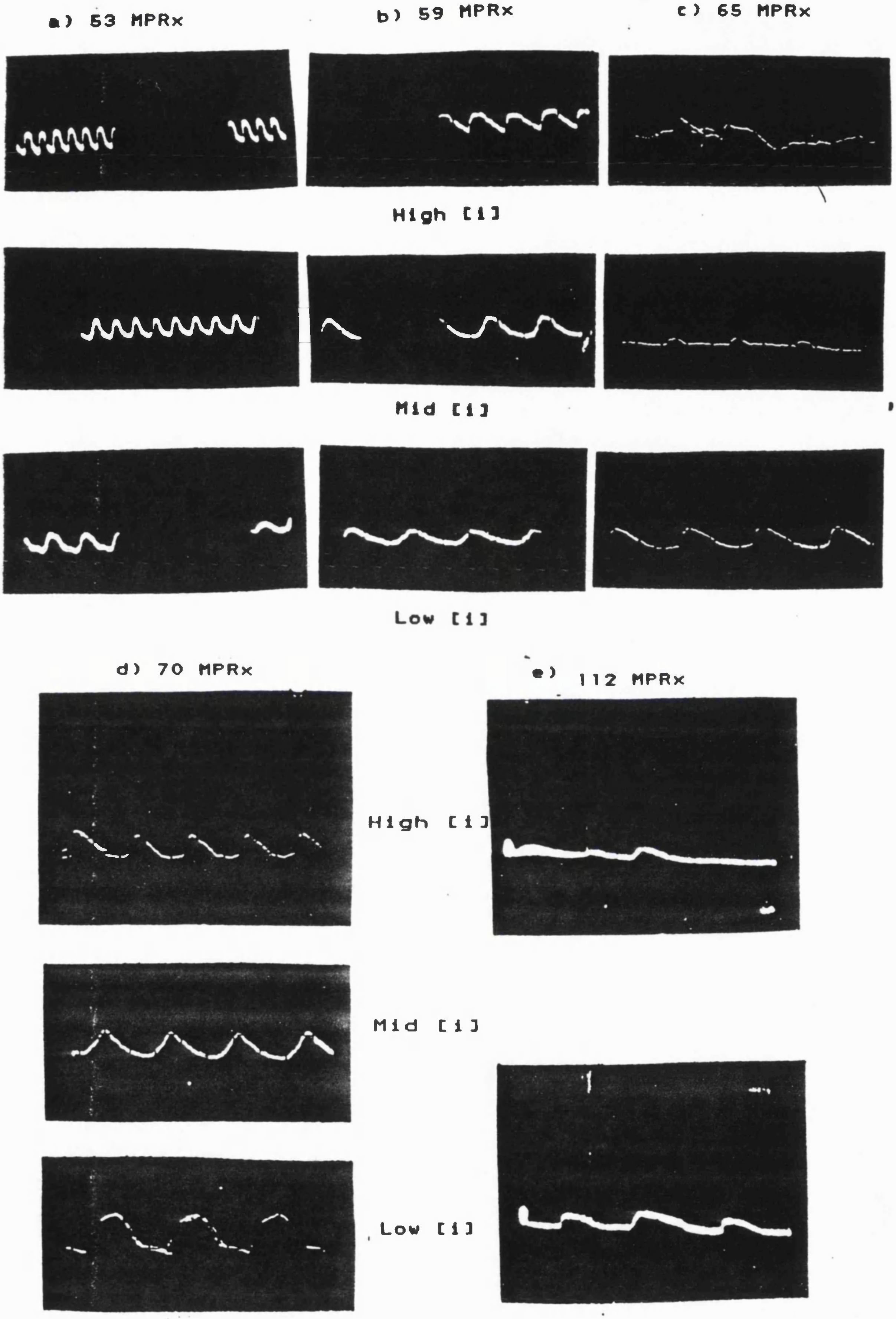


Figure 58 - Subject 2 - Lx waveforms

Appendix 3 A/B illustrates his very abnormal voice with extremely high Fx values, which in this case are not representing 2nd order Fx Means but Modal values. The severe limitation in vocal fold vibratory function resulted in very abnormal, sometimes bimodal Dx distributions (Fig. 57 a-e). In such distributions central measures are not really representative but we feel the 2nd order Mode could be used in this case as representing the main fundamental frequency used by subject 2. %TS and Cx plots reflect very limited vocal fold regularity of vibration. His Lx waveforms (Fig. 58 a and b) illustrate limited amount and short duration of vocal fold contact, mostly poor but varying control of pitch throughout his five year follow up period.

Self assessment (Appendix 6A)and the therapist's perceptual evaluation (Appendix 5, Table 1) all confirm subject 2's severe residual difficulties with voice production. We assume these are the result of the anterior glottic web tethering the vocal folds anteriorly. This has the effect of reducing vocal fold contact as evidenced by limited Lx amplitude (Fig. 58) thereby also reducing his ability to vary and control pitch.

Subject 14 had a long history of attendances at the hospital with chronic bronchitis and high blood pressure before the diagnosis of glottic cancer. A year previous to diagnosis he had developed hoarseness which was investigated. A biopsy was taken from his left vocal cord but found non-malignant. Malignant changes were diagnosed in February 1979 and he received radiotherapy in April the same year.

By July his vocal cord was reported smooth and mobile, but in October hoarseness had developed again. Oedematous false vocal folds and aryepiglottic folds were noted on every review occasion from then on until in November 1983, when he had an infected mucous cyst removed from his aryepiglottic fold. An area of leukoplakia was biopsied but not found malignant. He reported getting more hoarse when he felt tense.

In January 1984 he was diagnosed as suffereing from angina and had four months off work. His first ELG assessment for this study was carried out in May 1984, five years after radiotherapy.

Subject 14 was extremely tense with very poor vocal habits, shallow breathing and tended to speak on residual air. He complained of frequent choking attacks and breathing difficulties. He worked as a civil servant and used his voice a lot both at work and socially. He was about to return to work and it was considered imperative that he received a course of voice therapy to try to stop further vocal abuse to reduce his laryngeal discomfort.

He was seen for seven voice therapy sessions between the assessment at 61 MPRx and the following one at 63 MPRx (Appendix 3 A/B). It seems the treatment resulted in increased fundamental frequency and regularity of vocal fold vibration in both speech and reading. The improvement is also reflected in his self ratings (Appendix 6 A) and in the therapist's perceptual evaluation below (Appendix 5, Table 1).

viii Result of Perceptual and self-rated evaluation of voice quality.

Because of the earlier described methodological difficulties, no statistical analysis of results of either objective measures, subjects' self-ratings on the Questionnaire nor of the therapist's perceptual evaluation of voice quality on the VPA, were attempted.

Other studies had indicated 2-6 months before the voice returns to normal after radiotherapy (Mendonca, 1975, Stoicheff, 1975, Karim, Snow et al, 1983). The following descriptive summary is based on the first assessment of each Pilot subject at 2 MPRx or more after the end of radiotherapy (N=18). This was to avoid any influence on subjective measures of any information or advice given to the subjects regarding vocal hygiene or voice conservation. Subjects 13, 16 and 20, who had been assessed before and during treatment are therefore excluded from this summary.

The results of the therapist's perceptual evaluation on the VPA of the subjects' voice quality (on the first assessment occasion), where aspects of voice quality are rated on a six point scale, 1-3 indicating the presence of a feature within 'normal' limits, 4-6 indicating 'abnormal' degrees, revealed the following tendencies (average scalar degrees for the whole group are shown):

'Harshness'	'3.0'
'Whisper'	'4.0'
'Creak'	'4.5'
Laryngeal tension	'4.0'

The voices also tended to be judged as produced with a 'Low mean pitch' ('3'), which impression seems to agree with the objective Fx measurements (Appendix 3 A/B).

Stoicheff et al (1983) found that the degree of dysphonia in their sample of male patients rated one year after the end of radiotherapy (N=46), was significantly higher than in a control group of 'normal' speakers. They also found that 35 % of their irradiated patients had dysphonia ratings that fell within normal limits. They point out however, that this is not a 'majority' of patients as claimed by earlier studies.

The questionnaire (Appendix 1 a) required subjects to rate on a seven point equal appearing interval scale, aspects of voice usage, voice quality (hoarseness), whether the voice was experienced as a 'problem' and particular aspects of voice function. **The higher the rating the greater the difficulty.** Subjects were also asked to indicate the approximate time the voice took to recover after radiotherapy.

A 'Problem - score', hereafter referred to as the 'Mean P-score', was created by averaging the sums of each subject's ratings on 5 questions 12 a-e, which specify particular instances of voice function (Appendix 1 a).

A summary of subjects' (2 MPRx or more) self ratings **(average group scale ratings 1-7 in brackets)** indicated that on average they felt the voice had returned to 'normal' within approximately 2 months which is similar to reports in the literature. They reported having talked little during radiotherapy (2.4) and using their voices almost as much as usual (3.9).

Subjects indicated quite low degrees of residual hoarseness (2.5) and felt their voices did not pose much of a problem (1.9). The Mean P-score, however, indicated slightly more difficulty in terms of experienced functional limitations (3.2). The subjects did not find telephoning much of a problem, but using the voice at extremes e.g. singing, shouting or speaking over noise revealed considerable difficulties in many subjects. As these are activities that most people can avoid, this may explain why they still did not consider the voice as much of a 'Problem'. Vocal fatigue after a lot of talking was also common.

Due to the high average age among the pilot subjects (Mean = Median age 63, Range 39-82) only five had returned to work after radiotherapy, Subjects 1, 7, 8, 14 and 21. Riska and Lauerma (1966) had found a tendency for 'workers' to be less satified with their voices than people who retired or were unemployed.

Our 'Workers' (N=5) seemed to indicate only slightly higher voice usage after radiotherapy than their 'Retired or Unemployed' (N=13) colleagues, '4' vs '3.8', but they reported greater experiencce of 'Voice being a Problem', '3' vs '1.5' and slightly higher Mean P-score '3.7' vs '3.0', which seems to correspond more closely to their experience of the voice as a 'Problem'. The rated degree of hoarseness was very similar, '2.4' vs '2.5', in both groups.

Only two subjects were smoking after radiotherapy on their first and only assessment occasion, subject 5 and 21. They reported smoking 5 and 2 daily cigarettes respectively and neither had any major vocal problems.

Stoicheff (1975) suggested that the amount of talking patients did during radiotherapy may influence their subsequent voice recovery. The pilot study subjects were asked to rate the amount of talking they estimated they did while coming for treatment (Q 5, Appendix 1 a). This question, however, suffers from the problem pointed out by Fallowfield (1991) of requiring the subjects to recall over increasing lengths of time, which renders responses unreliable. The only subjects who reported having spoken 'As much as usual' in our study were subjects 1 (Supraglottic), 12 (T2), 13 (T2), 15 (T4), 17 (T2), 18 (T3). All others reported speaking less than usual during treatment.

The same problem of recall applies to question 6, "Roughly how long did your voice take to recover after the end of treatment?". Question 10 a asks the subject how much he uses his voice in his daily life and question 10 b, to compare this with how much he used his voice before radiotherapy. Again, this will be affected by the subject's ability to recall over increasing periods of time.

The above is a crude way of analysing the responses of our subjects and of perceptual voice quality ratings. It does not illustrate the very serious problems and limitations with voice production experienced by individual subjects e.g. subjects 1, 2, 3, 13, and 14 nor the almost normal voice quality and absence of residual vocal limitations in others e.g. subjects 4 and 7.

ix The effect of voice therapy on voice quality after radiotherapy..

In common with patients referred to the voice therapist with benign laryngeal disorders, all pilot study subjects responded positively to the explanations offered about the structure and function of the voice mechanism, how and why the voice was affected and to advice regarding vocal hygiene and voice conservation. The sometimes poor response to further reviews may have been due to the fact that the radiotherapy subjects had been cured of cancer and did not consider hoarseness a 'Problem', and also, compared to their voice qualities before radiotherapy, their post radiotherapy voices were no doubt better in the majority of cases, and they did not consider a residual dysphonia of much consequence.

Stoicheff et al (1983) did perceptual voice quality ratings of 46 male patients irradiated for vocal fold carcinoma before treatment but after biopsy, and one year after radiotherapy. Most patients deemed their voices significantly better after the biopsy compared to before. Eight experienced judges asked to rate voice quality parameters before and after radiotherapy showed, however, that although the degree of dysphonia was significantly reduced, it was still significantly higher than that of an age and sex matched control group. Our average VPA ratings for the group reported earlier (pp. 237) seem to agree, that perceptually the irradiated voices are not normal. They tended on average to be rated towards the 'Abnormal', 4-6, end of the six point scale.

As anticipated, some of the pilot subjects had more or less severe problems with their voices and were offered periods of voice therapy if they expressed concern about vocal limitations and discomfort e.g. subjects 2, 8, 10, 12, 14, 16, 20.

Table 11 below shows the times post radiotherapy that subjects received active therapy. The numbers before and after the oblique slash indicate self ratings given before and after voice therapy.

The reader is reminded that subject 20 had been referred before the diagnosis of laryngeal carcinoma and had received a course of voice therapy with some beneficial effect before diagnosis. Subject 16 attended voice therapy before and during radiotherapy and found this helpful in controlling his voice. He continued working throughout the treatment and did not lose his voice at any time.

All the other subjects received voice therapy some time after their initial assessments (Table 11).

The subjects with the poorest voice quality post radiotherapy were subjects 2, 12 and 14. Subject 12 did not return his selfratings

(Table 11) but subjects 2 and 14 seem to have found the therapy beneficial, subject 14 more so than subject 2, whose anterior glottic web severely limited his potential for improving his voice quality.

TABLE 11

Subject:	**2**	**8**	**10**	**12**	**14**	**16**	**20**
First assessment (MPRx)	53	0.5	2	0.5	61	pre	pre
Assessment pre-voice Rx (MPRx)	65	5	5	13	61	2	1
Assessment post voice Rx (MPRx)	71	6	7	15	63	3	2
Self-ratings:							
Mean P-score	6/5.8	1.3/1	3/3.2	-	5.6/4.2	1.8/1.4	6.4/5.8
Hoarseness	4/4	2.5/1	2/2	-	4/2	1/1	6/4
Voice a Problem?	5/4	1/1	2/2	-	4/1	2/1	7/3
Therapist's VPA ratings:							
Harshness	5/4	3/2	2/1	6/5	4/2	4/0	4/4
Whisper	5/5	4/4	2/0	5/3	3/3	5/3	5/4
Creak	6/6	2/3	5/4	6/6	5/5	4/4	5/5
Laryngeal tension	6/6	2/2	3/4	6/6	5/4	4/4	3/2

Table showing subjects' self-ratings of aspects of voice function and hoarseness and the therapist's perceptual (VPA) voice quality ratings before and after voice therapy.

Subjects 8, 10, 16 and 20 had no major problems with their voices. Subjects 2 and 10 had retired and subject 12 was unemployed.

Subjects 8, 10 and 14 smoked before radiotherapy, but gave up after. Subject 12 smoked before and had resumed smoking on reassessment at 15 MPRx. Subjects 2, 16 and 20 were non-smokers before and after radiotherapy.

All subjects reported talking as little as possible during radiotherapy.

On reassessment, subjects 10, 14, 16 and 20 report speaking as much or more than before radiotherapy. Subjects 2 and 8 speak a little less than before, but all report an increase in voice use after voice therapy.

Subjects 8, 10, 14 and 16 found voice use less tiring after voice therapy.

Table 11 above shows that the therapist seems to judge most of the subjects' voices less 'Harsh' after voice therapy and some also less 'Whispery'. The therapy does not seem to have had much effect on degrees of perceived 'Creak' or 'Laryngeal tension'. The severe difficulties of subject 2 are evident. The relatively high VPA ratings of subjects 16 and 20 are likely to be due to their being seen very early after the end of radiotherapy. Their self rated improvement may also be due to improvement in tissue status with increasing time after irradiation.

Most of the corresponding objective voice measurements are shown in Appendix 3 A/B. Not all the subjects were recorded both speaking and reading, and subject 16's and 20's recordings before and after voice therapy have been lost. Subject 20 did, however, attend for further assessments throughout the first year and the graphs in Appendix 3 show how his voice stabilised in fundamental frequency and regularity.

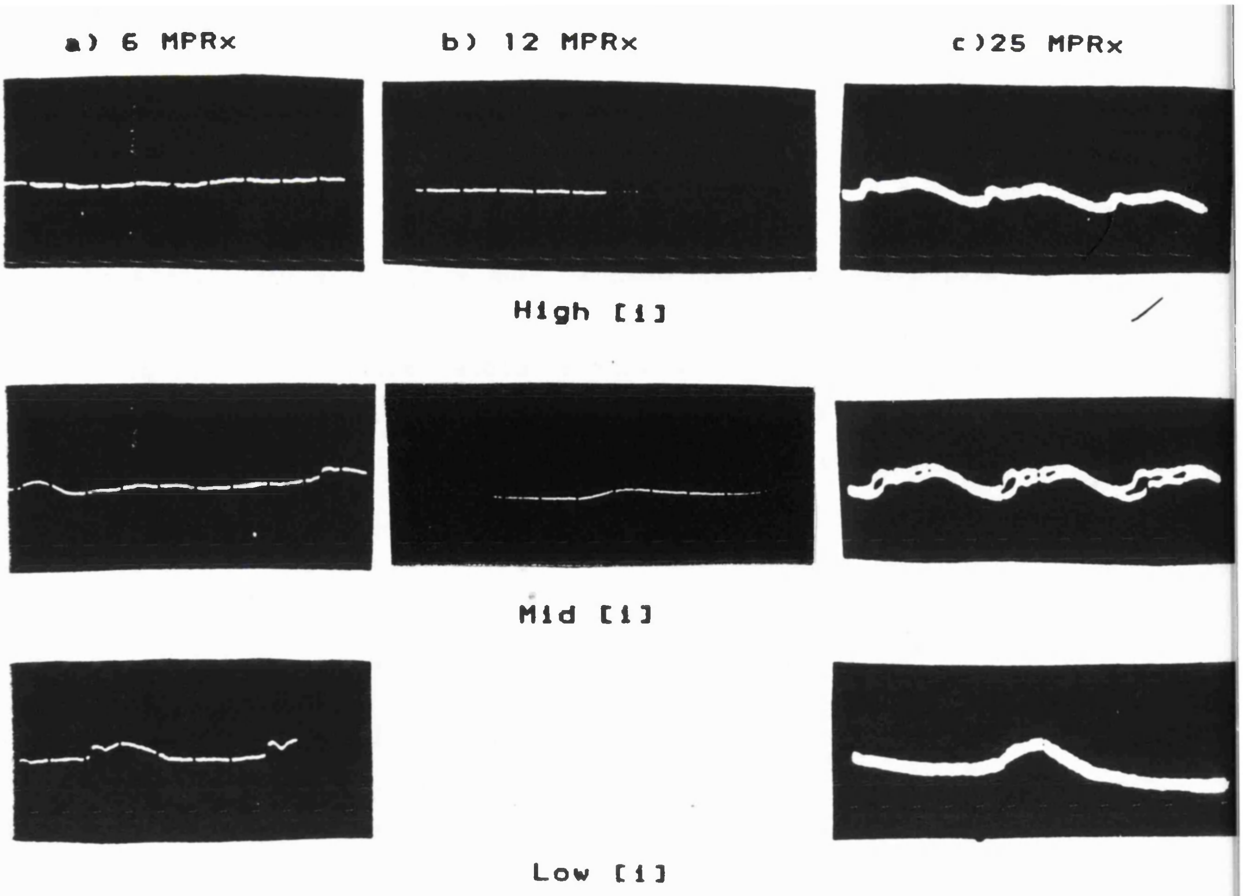


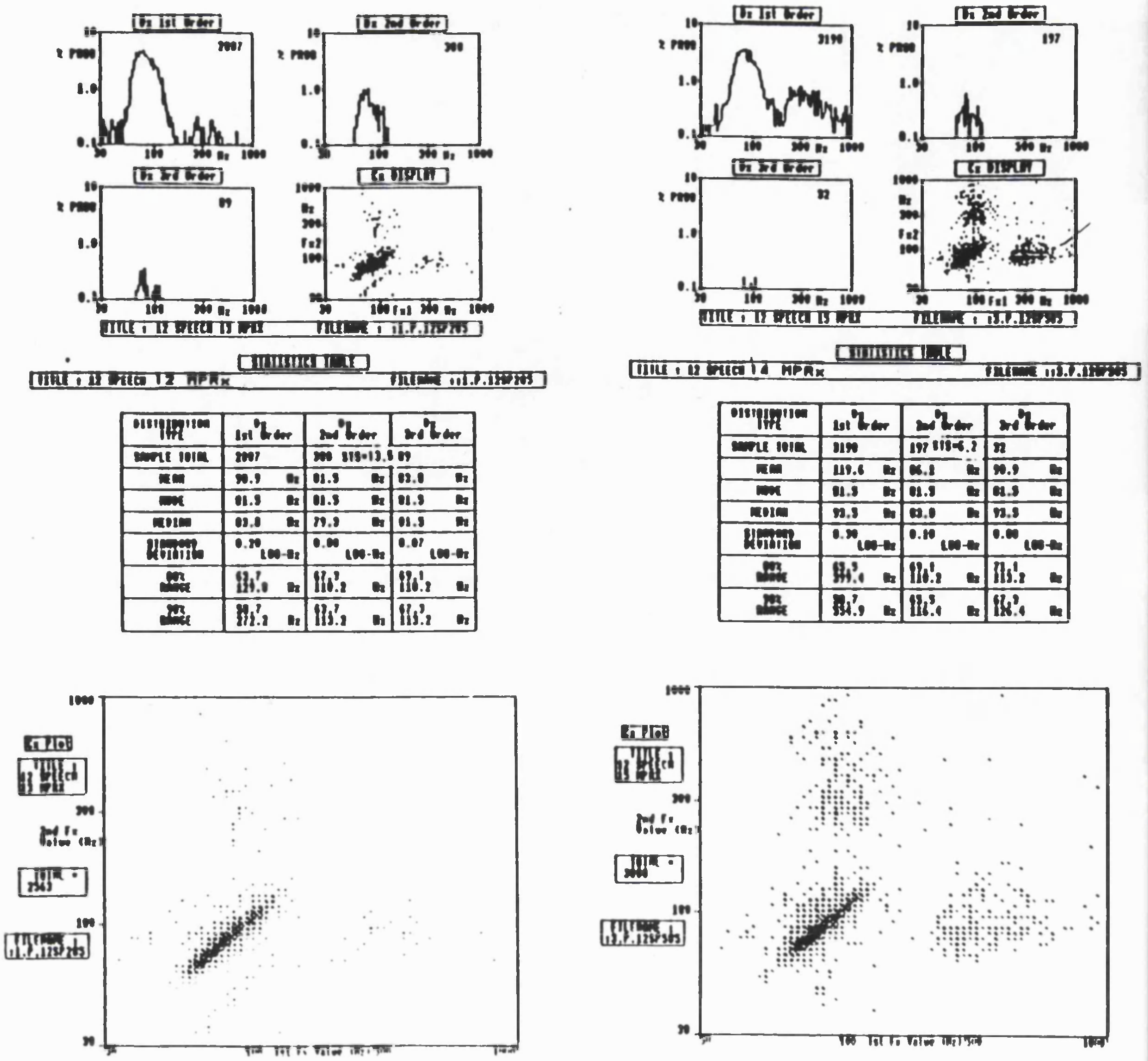
Figure 59 - Subject 12 - Lx waveforms

The Fx measurements for subjects 2 and 12 suffer from the fact that it was extremely difficult to obtain reliable and consistent Lx signals (cf. Fig. 58 a and b, p. 234, Fig. 59), their Fx data must therefore be interpreted with great caution. In subjects 10 and 14 there is evidence of increased fundamental frequency and regularity of vocal fold vibration after voice therapy at 7 and 63 MPRx respectively (Appendix 3 A/B).

Subject 12 had very abnormal voice quality. He refused to allow indirect examination of his larynx. His objective measures show extremely low fundamental frequency, the lowest of any speaker in the sample, and very low %TS regularity measures (Appendix 3 A/B) Perceptually he seemed to be using a well controlled ventricular band voice, which may account for the very high degrees of creak ('vocal fry') and laryngeal tension on the VPA (Table 11). Monsen and Engebretson (1977) found that 'creaky' voices consistently show fundamental frequencies below 100 Hz, which is certainly the case here. Keidar (1983) showed that perception of 'vocal fry' phonation is mainly fundamental frequency dependent and frequencies below 70 Hz are consistently judged as vocal fry. This is confirmed here, although the source of the creak in this case is assumed to be the false vocal folds.

Subject 12's Dx and Cx plots and the statistics pertaining to his speech sample on the two occasions before and after voice therapy (Fig. 60 a and b), show extremely low fundamental frequency distributions where the lower end of the 90 % range in 1st and 2nd order is below 70 Hz. The 1st order Dx plot in Figure 60 b shows massive high frequency 'noise' as a result of the gain setting being turned up on analysis due to the poor Lx signal on recording as a result of presumed ventricular band phonation.

It was impossible to capture Lx waveforms on the oscilloscope screen during sustained phonation on this occasion and subject 12 had little ability to vary his pitch up or down, further supporting the suspicion of ventricular band phonation.


a) Before voice therapy

b) After voice therapy

Figure 60 - Subject 12 - Speech, Dx and Cx plots and Fx parameters

His extremely low regularity measure %TS (Appendix 3 A/B, Fig. 60 a,b) is likely due to the poor Lx signal. Therapy was aimed at helping him to raise his pitch, which he did successfully in exercises, but was unable to sustain reliably in conversaaion.

Subject 12 had a very hoarse voice for four years before attending the ENT department. He admitted to having abused his voice by extremely heavy smoking, up to 60/day, singing, talking, continuously forcing and often losing his voice. His ventricular band phonation post radiotherapy may have been the result of reverting to what was established as 'normal' phonation for him before radiotherapy. It is often considered to be a component of hyperfunctional dysphonia, which may develop in response to poor vocal fold movement (Colton and Casper, 1990).

The last subject included in the Pilot study, subject 21, was referred in February 1985 but follow up assessments of some subjects continued throughout 1985 and early 1986 e.g. subjects 12, 16 and 20.

x SUMMARY

The subjects' low self-ratings of degrees of hoarseness and the therapist's relatively high VFA ratings may reflect a degree of discrepancy between her and the subjects' perception of the 'normality' of their voices, the subjects being more positive than the clinician (cf. Karim, Snow et al, 1983).

The objective voice measurements illustrate on the whole lower than 'normal' speech and reading fundamental frequency and vocal fold regularity of vibration as measured by %TS, modifying the claim of 'normality' of the voice two months post radiotherapy (Appendix 3 A/B).

It seemed our pilot study had to some extent answered four questions posed in the introduction:

1) Fundamental frequency parameters, Dx and Cx plots did demonstrate, if not normalisation, at least stabilisation of measures and Lx waveforms with time after radiotherapy.

2) Regarding a pattern of recovery, there seems to be a tendency towards a lowering of fundamental frequency during and for the first 6 months after the end of radiotherapy followed by a stabilisaiton over time (Appendix 3 A/B).

3) Fundamental frequency and %TS measures (Appendix 3 A/B) demonstrated in some subjects the effect of extensive tumour on vocal fold vibratory characteristics e.g. subjects 15, 18, 19 (T3, T4 and Supraglottic); Lx waveforms reflected the effect of recurring tumour e.g. subject 3 (Fig. 56); undiagnosed tumour subject 20 (Fig. 51); possibly of continued smoking and habitual vocal abuse e.g. subjects 12 (Fig. 59) and the effect of an anterior glottic web in subject 2 (Fig. 58).

4) Regarding the effect of voice therapy on ELG measurements there is some evidence of a beneficial effect on Subject 10's Fx and % TS measures between 5 and 7 MPRx (Fig. 53) and on subject 14's fundamental frequency and regularity measures between 61 and 63 MPRx (Appendix 3 A/B).

The purpose of the pilot study had been fulfilled and the experimenter had gained valuable experience in analysing Lx waveforms and fundamental frequency plots, which had become part of the routine assessment of dysphonic patients in the Speech and Language Therapy department. It had been found particularly helpful for intra-individual comparisons at various times during and after voice therapy and/or surgical intervention for organic laryngeal pathology (Carlson, 1986). Lx waveforms had also been found extremely helpful for visual feedback of laryngeal aerodynamics in voice rehabilitation.

The original protocol had proposed a three part study, the first two of which may be considered to have been carried out although not on the scale originally envisaged due to the small numbers of subjects referred. Subjects had been assessed at a great variety of times post radiotherapy and a few had also been assessed before, during/and after radiotherapy.

The third part of the study proposing a random allocation of subjects to voice therapy treatment or to a control group was not possible to realize due to the small numbers referred. Nor would it have been appropriate for all subjects. A number of Pilot subjects had, however, been offered periods of voice therapy with beneficial effects as indicated above.

xi Missing data.

As mentioned in the earlier description of methodology, the intention of reassessing subjects using the same techniques at regular intervals on several occasions after the end of radiotherapy was difficult to implement. Where subjects had missed or cancelled appointments, a request by the therapist for them to attend on an extra occasion, met with a mixed response.

Some subjects when approached gave excuses not to be recorded on a second or third occasion, or agreed to attend but failed the appointments. The reasons for this may have been that, despite some subjects having quite considerable vocal limitations as reported in their self ratings on the questionnaire, they still did not consider their vocal limitations as a 'Problem'. Having been cured of a potentially fatal disease, they seem to consider the residual vocal difficulties as insignificant and may therefore not have been motivated to attend for further voice assessment and advice.

Subject 13, having been told of the strong causative relationship between smoking and laryngeal carcinoma, was smoking again one week after the end of radiotherapy. This may have been the reason why he was reluctant to attend regular appointments.

B THE MAIN STUDY

i Introduction and aims.

Two years after the last subject had been assessed for the Pilot study, a patient was referred for voice therapy after complaining that his voice had not returned after radiotherapy for a T1 tumour of the vocal fold five months previously. He was a 49 year old lifelong non-smoker, managing director of a company and used his voice a lot. A routine voice assessment was carried out using the Electrolaryngograph and he filled in the questionnaire, which had been devised for the Pilot study.

This post radiotherapy patient successfully treated with voice therapy as evidenced by objective, perceptual and subjective measures rekindled the idea of continuing data collection of irradiated laryngeal cancer patients to study the long term effect of radiotherapy and, where appropriate, voice therapy, on voice quality.

There was no reason to believe there would be any greater numbers of subjects referred than had been referred to the Pilot study. Random allocation for voice therapy and to a control group was considered unrealistic and inappropriate as we had found that not every irradiated patient needs therapy.

In 1988, 'Hypofractionation' regimes were introduced at St Thomas' Hospital for treatment of laryngeal tumours, i.e. five fractions per week (5F/wk) instead of three (3F/wk) which had been the treatment received by the subjects in the Pilot study. 3F/wk had been the regime used in the treatment of the new referral, subject 22. The earlier described BIR study (Wiernik, Bates et al, 1990, Chapter VI, viii) had tried to establish whether there were any significant differences in control rates, acute or late normal tissue reactions between the two regimes but failed to prove this was the case. Some oncologists maintain, however, that 3F/wk regimes delivering higher dosage per fraction, although the total tumour dose is reduced compared to 5F/wk, will cause more severe reaction.

Due to the wide range of variables influencing individual vocal fundamental frequency parameters, added to which were the effects of varying sites and stages of malignant vocal fold lesions and subsequent radiotherapy treatment, interpretation of the results of voice assessments for the group of pilot subjects was tentative.

The plots did, however, as shown, reflect changes and developments in individual subjects' measurements with time post radiotherapy e.g. subjects 8, 10, 12, 13 and 20; with observed vocal fold abnormalities eg. subjects 2, 3 and 14, and some improvement in such measures with voice therapy e.g. subjects 10, 12, 14 (Appendix 3 A/B).

To enable use of data already collected for the Pilot study, no major changes were introduced in the assessments for the Main Study. Longer speech and reading passages were, however, analysed to improve the representativity of the samples. The new reading passage was the 'Rainbow Passage' (Fairbanks, 1960) (Appendix 7), which is slightly longer than the one used with the Pilot subjects and often quoted in American normative studies of voice fundamental frequency.

Some modifications of the questionnaire were also introduced (Appendix 1 b) on the basis of comments that subjects had made on the original version.

It was clear that some normative ELG data from a group of age and sex matched subjects with no history of significant voice or hearing problems was needed for comparison with the irradiated speakers. The study by Kramer (1989) was designed to allow such comparisons. She used the same instrumentation in recording and analysing a speech sample, the reading of the Rainbow Passage and calculated %TS for both reading and speech for a group of 'Normal' speakers.

Kramer's study (1989) compared Fx and %TS parameters of male smokers (N=9), non-smokers (N=5) and ex-smokers (N=9), aged 50-75, an age range chosen to correspond to our irradiated subjects'. They had no significant history of voice problems as judged from an extensive questionnaire interview carried out by the researcher. A trend was found of non-smokers having more regular vocal fold vibration as measured by the proportion of Tx sample carryover into second order (%TS), than either smokers or previous smokers (Kramer, 1989).

As suggested by Barry et al, (1990) too stringent criteria for selecting a 'normal' group of speakers will make it extremely difficult to collect a sufficiently large sample. Our group of irradiated speakers came from all walks of life. Most had been heavy smokers immediately before or years before the diagnosis of laryngeal carcinoma was made. Several had had voice problems, sometimes for years before the diagnosis was made e.g subjects 2, 7, 12, 18, 20. It seemed not in-appropriate to compare them to a group of 'Normal' speakers, who were likely to come from as wide a sociocultural, smoking, health and linguistic background, but without a significant history of voice problems.

The answers to five questions were sought in the Main study:

1. Do patients irradiated for T1 tumours have better voice quality than patients irradiated for T2 tumours?

2. Do patients irradiated in 5 fractions per week have better voice quality than those treated in 3 fractions per week?

3. What is the relationship between objective, self rated and perceptual measures of voice quality used in this study?

4. What happens to voice quality features with increasing time after radiotherapy?

5. Do voice quality measures improve after voice therapy in irradiated laryngeal cancer patients?

ii Subjects.

Subjects referred to the Main Study were to be male native English speakers (with one exception, subject 35) referred before, during and after radiotherapy, but only for early vocal cord tumours, stages T1 and T2:

a) Patients who had received treatment in 3 fractions per week (3F/wk) (Hyperfractionation).

b) Those who had received treatment in 5 fractions per week (5F/wk) (Hypofractionation).

Medical staff were made aware that the study was to continue and were invited to refer any patient who fulfilled the criteria. Subjects referred from February 1988, including subject 22, were accepted for 'The Main Study'. The last subject admitted to the study, subject 40, was referred in March 1993 by which time another 19 subjects had been assessed and followed up. The group consisted of 13 subjects with T1 tumours and 6 with T2 (see Appendix 2). Among them were five who had been treated in 3 fractions per week, subjects 22, 23, 33, 34 and 38. All other subjects referred to the Main study had been treated in 5 daily fractions spread over one month.

Re-assessment of some of the T1 and T2 subjects from the Pilot study was carried out to see whether any major changes in fundamental frequency or self-ratings had occurred in patients who were now several years post radiotherapy e.g. subjects 2, 4, 5, 7, 8, 10, 12, 14, 16 and 21. All these subjects had been treated with three fractions per week (3 F/wk). They will be described in later sections as 'Pilot subjects revisited'.

Appendices 8-10 show plots of the total collected fundamental frequency and %TS data as a function of time, for all T1 and T2 subjects, for both speech and reading including previous assessments for T1 and T2 Pilot study subjects. As there were only four T2 subjects who had received 5 fractions per week, (the other eight had been treated with 3F/week), they all appear in the same plots, Appendix 9.

Subjects' self assessments of degree of 'hoarseness', their voice posing a 'Problem' and their 'Mean P-scores' also as a function of time, can be seen in Appendix 6 A/C.

This is a study of changes in voice quality measures in the long term. The number of subjects referred is still not large, nor were they randomly selected from a total population of irradiated T1 and

T2 laryngeal cancer patients. Statistical analysis of the total body of data is therefore not attempted.

A random draw is, however, made of ten irradiated T1 subjects, and ten 'Normal speakers' on whose objective measurements, and, in the case of irradiated subjects, perceptual voice evaluation and self rated voice quality parameters, a number of statistical tests are carried out. Comparisons will be made and tentative conclusions drawn regarding the total sample with reference to such smaller studies within the study to investigate:

a) repeatability and agreement of measurements using different analysis programs and instrumentation (Chapter IX),

b) significance of fundamental frequency differences between 'Normal' speakers and irradiated T1 subjects (Chapter X)

c) significance of differences between speech and reading fundamental frequency parameters (Chapter X)

d) correlations between acoustic, perceptual and self rated voice quality measurements (Chapter XI).

A descriptive approach is then employed regarding the whole group of T1 and T2 subjects' with reference to these studies (Chapter XII).

CHAPTER IX - REPEATABILITY OF AND AGREEMENT BETWEEN SOME Lx BASED MEASUREMENTS OF FUNDAMENTAL FREQUENCY, REGULARITY AND IRREGULARITY.

A i Assessing the agreement between two analysis programs.

If two analysis systems are applied to the same body of recorded Tx data, the resulting Fx parameters should be almost identical. Two alternative software systems were used for fundamental frequency (Fx) analysis of recorded Lx data in the early stages of this study. One was the Dx ROM, which allowed a quick and easy way to get a graphic display with some statistical information of the recorded Fx data (Fig. 61 a). The Dx ROM was originally used in the voice clinic for immediate feedback of a patient's speech or reading Fx. The fundamental frequency histograms (Dx) and scatterplots (Cx) were printed and kept in the notes for future comparisons.

The second analysis program was on disc and called the TPS system. It allowed a more detailed and elaborate analysis of the recorded Lx data, which, however, had to be 'acquired', processed and stored on a second disc for analysis purposes. It was more time consuming but gave more detailed graphic and statistical information (Fig. 61 b). This is the program that has been used to analyse the vast majority of the speech and reading samples of the subjects in this study.

The Dx ROM displayed the 1st and 2nd order Dx histograms and the scatterplot and employed the same 128 'bins' logarithmically equally spaced along the frequency axis, as the TPS system.

We were interested to find out how closely Fx parameters agreed, calculated on the same body of data for each subject on the same occasion, using the two different systems. As the ROM analysis was so much quicker to carry out could the two systems be used interchangeably? If there was good agreement and less detail needed, the Dx ROM program would be the analysis method of choice. Agreement between analysis methods and programs is important when comparing fundamental frequency data arrived at using different measurement techniques and analysis methods (cf. Baken, 1987)

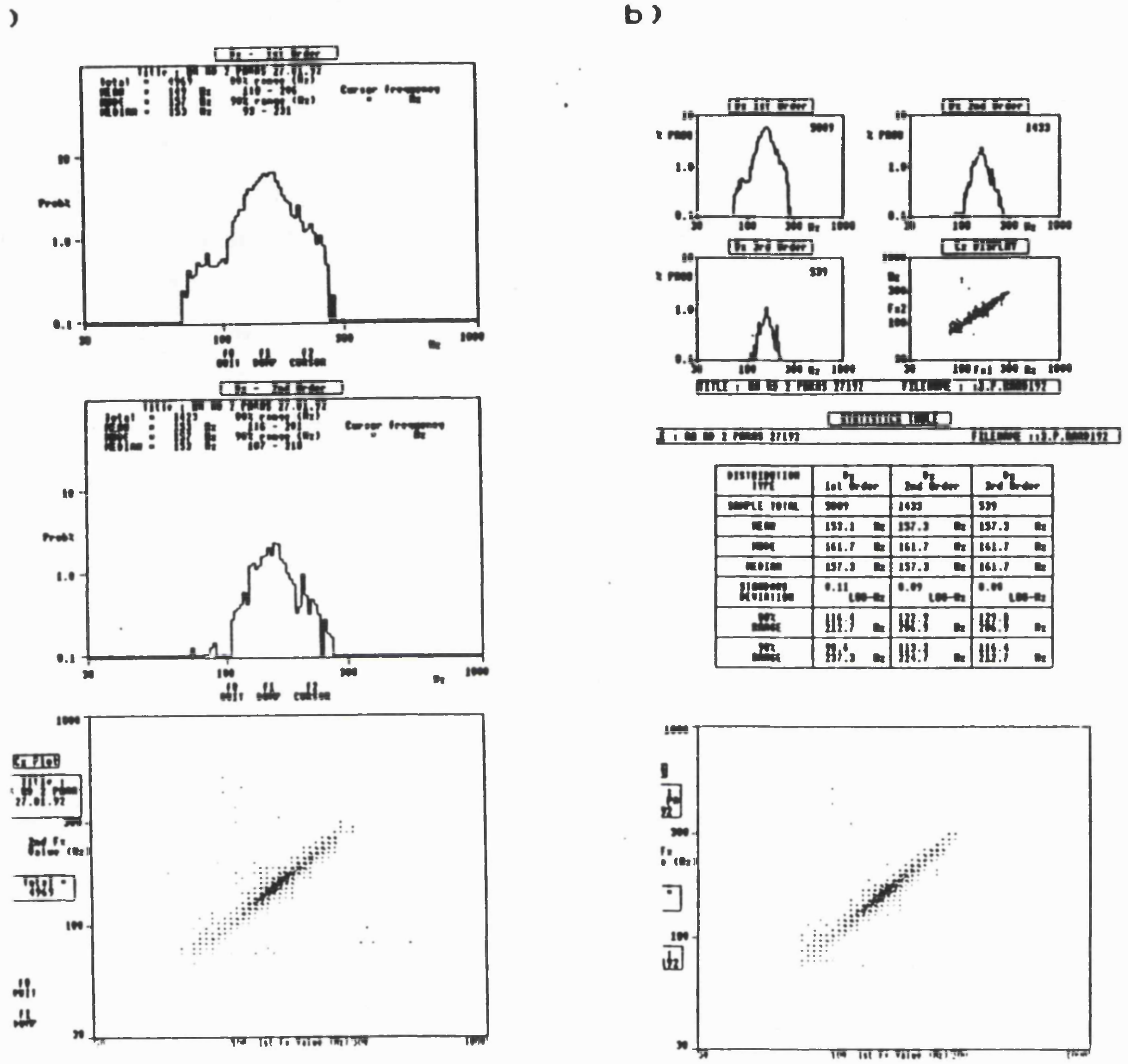


Figure 61 - Dx ROM (a) and TPS (b) software.
Fx analysis of the same recorded
reading passage. Normal speaker.

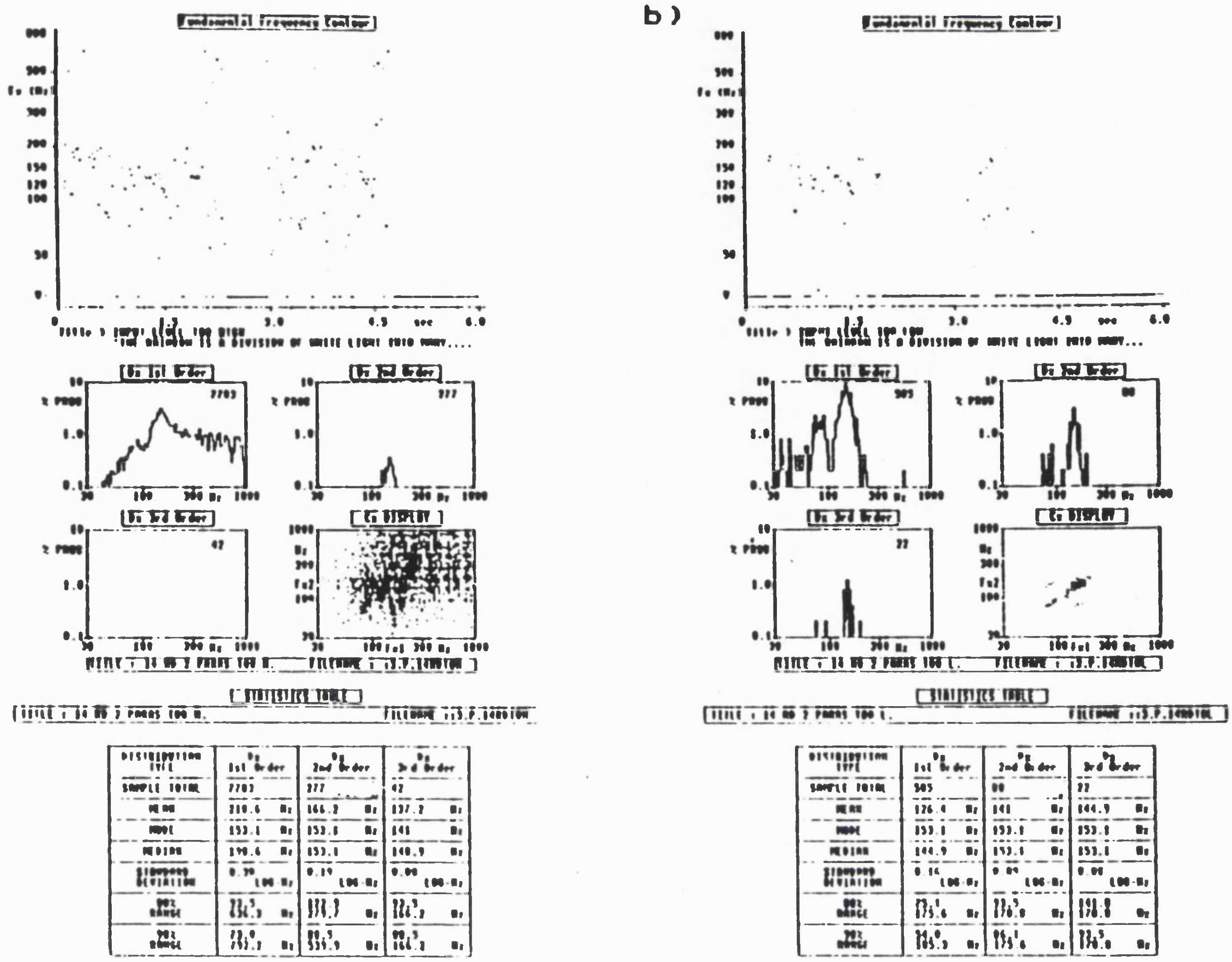


Figure 62 - Comparing the effect of different gain settings on analysis of a recorded reading passage. T1 subject, 14.

a) setting too high b) setting too low

Another aim was to find out how repeatable and reliable the measurements were using the two systems. Would the same values be obtained if the same recording were analysed again using either the TPS or ROM system? This aim was achieved by analysing the same recorded reading samples on two different occasions. Repeatability is particularly important when comparing fundamental frequency statistics of the same subject's voice on different occasions. It is also important when comparing data collected and analysed by two different observers using the same instrument. Different gain settings on analysis may affect measurements (Fig. 62) and consistency is needed on the part of the observer/analyst.

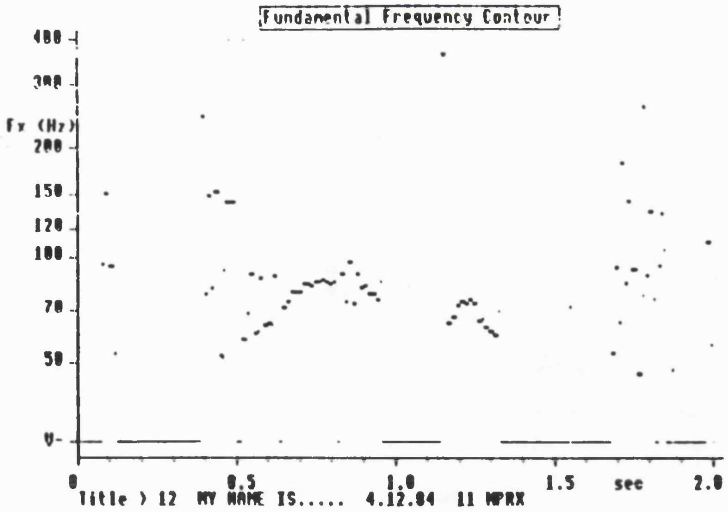
There is one known difference between the programs, namely that the central bin values of the ROM are expressed as whole numbers, whereas the TPS uses fractions. Presumably this reflects a slightly more gross division of bin widths and analysis of the Tx data. This would mean that a difference between Fx measurments of +/- 1 Hz should be ignored.

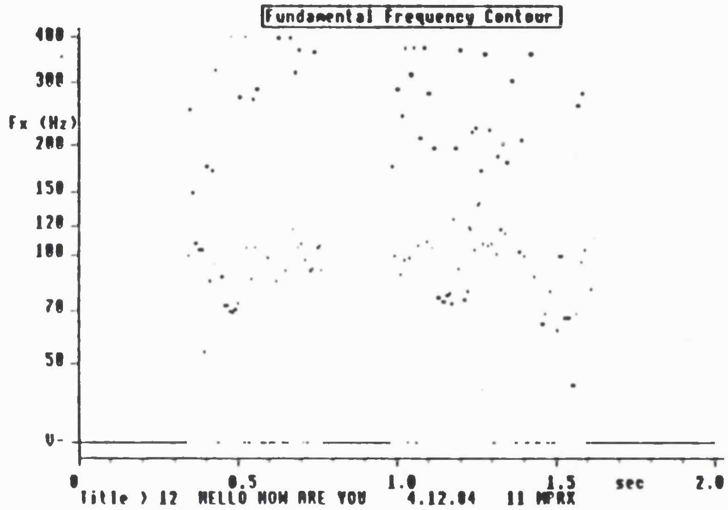
ii Equipment

In the first agreement and repeatability study, the Laryngograph Processor, a development from the original Laryngograph and Voiscope, was linked to a BBC Master microcomputer, a Cumana dual disc drive and an Epson FX 80 printer. A Telequipment S 61 oscilloscope was used to show the Lx waveform in real time. The inputs from the laryngeal surface electrodes and an RS professional dynamic microphone, were recorded on a Sony TC 144CS stereo cassette tape recorder . The equipment was used to record and analyse fundamental frequency (Fx) parameters of vocal fold vibratory function in two groups of subjects.

iii Calibration of input level.

The Fx ROM was used to calibrate the input level from the tape to ensure all voiced segments were displayed with little evidence of extraneous 'noise' in the signal (Figure 52, p. 224). There are some factors which may interfere with the strength and clarity of the Lx signal e.g. a lot of subcutaneous fat overlying the thyroid





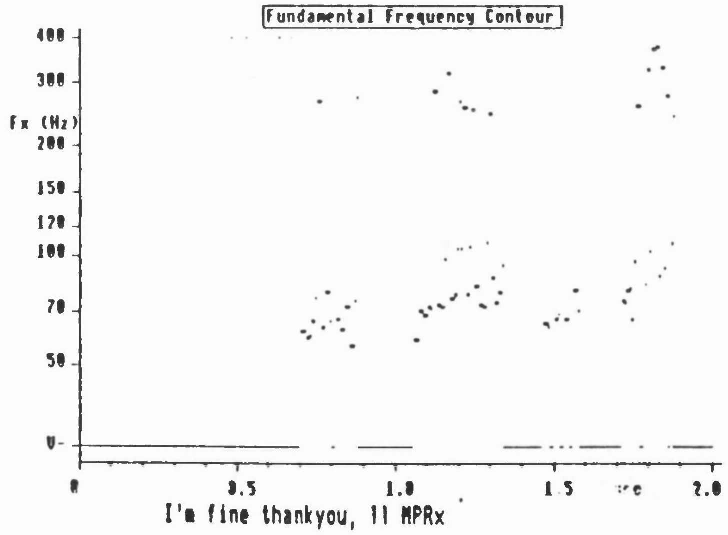


Figure 63 - Subject 12, Fx contours for three sentences, 11 MPRx.

cartilage, difficult access to the thyroid alae in some older men with short necks, a lot of laryngeal movement during speech and a very wide angle between the thyroid alae (Titze, 1990). Another factor important in the group of radiotherapy subjects, is excessive dryness of vocal fold mucosa, which leads to poor vocal fold contact, poor vibrational periodicity and effortful voice production with excessive laryngeal movement.

All these factors may lead to a 'noisy' Fx contour (Fig. 63) resulting from spurious Tx samples which are not reflecting vocal fold contact patterns but which may be included in the eventual Dx and Cx plots and calculations of Fx parameters. When calibrating the input level to get a representative Tx sample to be analysed, a balance has to be struck between how much of the recorded true vocal fold vibratory information one is prepared to sacrifice, for the benefit of not having extraneous 'noise' blurring the measurements, or whether some 'noise' is tolerable for the benefit of getting as large a Tx sample as possible on which to carry out the analysis. Examples of the effect of different input levels on Fx contours and the resulting Dx, Cx and statistical calculations for the same voice sample are shown in Fig. 62, p. 256. Visual monitoring of Fx on the screen is the easiest way to ensure consistency between the levels of one input and another.

For each subject the reading sample to be analysed was identified on the tape. The Fx ROM was set to scroll through the reading passage continuously, while the replay level on the tape recorder and the gain control on the Laryngograph Processor were adjusted to show the 'cleanest' and most representative Fx contour possible, for each speaker.

iv The subjects.

One group of ten subjects was drawn at random from the total number of subjects referred with irradiated T1 tumours. This group will be called the 'T1-subjects' (See Appendix 11).

A control group was made up of ten 'normal' male subjects, who had taken part in an ELG study of the effects of smoking on voice fundamental frequency (Kramer, 1989). The sample consisted of three non-smokers, four smokers and three ex-smokers (Table 12). This group will be referred to as the 'Normal speakers'.

Table 12

	Normal speakers (N=10)	T1 subjects (N=10)
Non smokers (N)	3	2
Ex smokers (N)	3	7
Smokers (N)	4	1

Table showing the smoking habits among the randomised samples of speakers used in the repeatability study.

Kramer defined her 'Ex smokers' as those who reported having smoked at least one cigarette or pipe per day, one cigar per week or any combination of these for more than one year. They must have stopped at least one year before the interview. Nine subjects in her study fell into this category four of whom had stopped 1-10 years before the interview. Five had stopped more than ten years previously. They all reported having smoked for more than ten years before giving up. Six for more than 20 years.

A 'Smoker' was defined as someone who still smoked at the rate described for the ex smokers. In the total sample there were 9 subjects, who fell into this category. One had smoked between 10-20 years, eight had smoked more than 20 years.

The sample of irradiated T1 subjects recorded for this study could not be described as selected at random. They had been referred from the joint ENT/Radiotherapy clinic at any time post radiotherapy and recorded at various times post treatment as they attended for reviews. For the purpose of the repeatability study, a random selection was therefore made of the last reading sample recorded of ten T1 subjects, two months or more after the end of radiotherapy. Their respective Months-Post-Radiotherapy (MPRx) are shown in Appendix 11. The range is extremely wide: between 4 and 168 MPRx, with a Mean of 65 MPRx. The Median is 50 months i.e. 4 years and two months and is the better central measurement here as the range is so wide.

As heavy smoking is one of the main causative factors in laryngeal carcinoma (Shaw, 1979, Gori and Bock, 1980, Herity et al, 1981, Flanders and Rothman, 1982, De Stefani et al, 1987, Guenel et al, 1988, Maier, 1990) it is not surprising to find that among the ten T1 subjects selected at random for the repeatability study, there were seven ex-smokers, one current smoker and, surprisingly, two who had never smoked (Table 12).

An attempt was made in the Kramer study (1989) to record a sample of 'Normal' male speakers, within approximately the same age range as the T1 subjects. The age range aimed at was 50-75 years. However, as the decision was made to select the last reading sample recorded out of the otherwise very large potential sample of recordings of T1 subjects at 2 months or more post radiotherapy, this resulted in the randomised sample of T1 subjects being significantly older than the 'normal' sample ($p < 0.03$). This does not matter for the repeatability study but will be discussed in the comparison of fundamental frequency parameters between the two groups. Individual Normal speakers' and T1 subjects' ages and smoking habits are shown in Appendix 11.

v Voice samples.

The recorded Tx data from the reading of a standard text was analysed. Pre-recorded stereo cassette tapes with the ELG (Lx) signal on one channel and the audio signal on the other, were used. Most subjects were reading the first two paragraphs of the 'Rainbow Passage' (Fairbanks, 1960 Appendix 7) as the standard task to be compared for fundamental frequency (Fx) parameters derived from Lx. The Rainbow passage was read by all the 'Normal' subjects and most of the T1 subjects. It had not been available in the very early days of the Pilot study, when the 'North wind and the Sun' (Appendix 4) was used as the standard text. This does not matter for this part of the study, however, as the same text is analysed on two occasions and the resulting pairs of measurements compared with each other.

The input level to the microcomputer for the analysis of the taped reading samples, was calibrated for each speaker, as described above. Once a satisfactory replay level was arrived at, using the controls on the tape recorder and the Laryngograph Processor, the setting was left constant for analysis of each reading sample by means of the Dx ROM and the TPS system on the same occasion (occasion 1 or occasion 2) referred to in the text as TPS 1/2 and ROM 1/2 respectively. To avoid any systematic variation in the replay levels from the first analysis to the second on each occasion, the order in which the TPS and ROM analyses were carried out was varied at random. This part of the study was then used to compare the **agreement** between the Fx parameters using the two systems of analysis, i.e. TPS 1 vs ROM 1 or TPS 2 vs ROM 2.

To assess the **repeatability** of the measurements from occasion 1 to occasion 2, by each method, TPS or ROM, each subject's recorded reading sample was analysed again six weeks later, using the same manner of calibrating the gain of the input signal. These analyses are referred to as TPS 1 vs TPS 2 and ROM 1 vs ROM 2. The same experimenter carried out both analyses and the same means of calibrating the input level was used.

vi Statistical methods.

Bland and Altman (1986) and Altman (1991) describe a method of testing the repeatability of and degree of agreement between different methods designed to measure the same variables. They suggest that the common use of correlation coefficients is not appropriate in these cases as they only describe the strength of a relation between two variables, not the degree of agreement between them *'....it would be amazing if two methods designed to measure the same quantity were not related'*.

In much clinical research, indirect methods of measuring physiological function have to be used. One example is the laryngeal electrodes used in ELG for measuring fundamental frequency and regularity of vibration of the vocal folds. Such measurements often vary slightly from one occasion to another due to uncontrollable variables, as suggested above, that may affect the Lx input signal. In this chapter we try to assess

> ***the size of the resulting differences between two measurements and their distribution around zero (Fig. 64-67). The 'Average' of two measurements of the same parameter is used as the best estimate of an unknown 'true' value (Altman, 1991).***

There were two early systems for analysis of the recorded Lx data as described above, the Dx ROM which is quick and the TPS system which is slower but gives more detail. *'..We want to know by how much the new method is likely to differ from the old. If this is not enough to cause problems in clinical interpretation we can replace the old method by the new or use the two interchangeably.'* (Bland and Altman, 1986).

Only clinical judgement and what we know about the variability of fundamental frequency measures productively and perceptually, within individuals, can decide what are 'acceptable' differences between repeated measurements of the same parameters in the same body of data. Barry et al (1990) found fluctuations in Fo within consecutive two minute sections of the same 15 minute reading task of up to 18 Hz in men and 23 Hz in women. This amounted to a 15% variation within an individual on the same task. Coleman and Markham (1991) using more subjects, found that speakers could be expected to be within 3 semitones of their "true" speaking fundamental frequency 90% of the time. They suggest there needs to be at least a 10-15% difference in Fo measures for the same individual on different occasions for there to be a significant difference. (For a detailed account of intra - individual differences see Chapter II).

The acceptable difference between repeated measurements for an individual on the same recording of a speech task must obviously be considerably less than the quoted variation within speakers. As the recording and analysis of the Lx input signal is used in the voice clinic and for research purposes, to store and analyse voice samples of the same speakers before and after intervention, it is, however, important that true, actual changes in fundamental frequency are registered reliably. An arbitrary 'Mean difference' (d) of +/- 5 Hz between measurement analyses of the same speech sample was considered acceptable in this study.

%TS is a fairly gross estimate of voice regularity and expected to be sensitive to the level of the input signal because of the way the analysis distributes Tx samples into 'bins' of certain frequency widths (Fig. 62). In second order only those instances where two adjacent Tx samples fall into the same bin are registered and reflected in the 2nd order Total Sample size. A weak input signal may not be too problematic but a too strong one with many frequency bins incremented to a very low level in 1st order due to noise, would result in a very low %TS in second order (Fig. 62 a). As previously mentioned, normal male speakers tend to carry over 30-50% of the original sample into second order (Kramer, 1989). An arbitrary 'Mean difference' (d) between measurements on the same recording of +/- 5 percent was deemed acceptable in this study.

Data analysis

Tables of each T1 subject's and each 'Normal' subject's 1st and 2nd order Means, Modes and %TS, were drawn up after each occasion, 1 and 2, for each analysis method, TFS and ROM (Appendix 12, Tables 1-20). The differences between and averages of each pair of Means, Modes and %TS, from one analysis to the other, for each speaker, were tabulated as shown in Table 13. The pairs of values were then plotted against each other as shown in Figs. 64-67.

The 'Mean difference' (d) was calculated for each group of ten speakers, and the standard deviation of the differences (SD_{diff}). The 'Mean difference' gives an estimate of the average bias of one method compared to another. The standard deviation gives an idea of how well two methods agree for an individual within the sample (Altman, 1991).

TABLE 13

TPS 1 VS ROM 1

NORMALS, Reading

Subject	1st order Mean (Hz)				2nd order Mean (Hz)			
	TPS 1	ROM 1	diff.	Average	TPS 1	ROM 1	diff.	Average
X	93,5	91,0	2,5	92,3	96,1	93,0	3,1	94,6
JSt	88,5	86,0	2,5	87,3	96,1	93,0	3,1	94,6
PBr	113,2	113,0	0,2	113,1	119,6	119,0	0,6	119,3
RA	153,1	149,0	4,1	151,1	157,3	153,0	4,3	155,2
FC	119,6	116,0	3,6	117,8	122,9	119,0	3,9	121,0
H	129,9	126,0	3,9	128,0	133,5	129,0	4,5	131,3
FW	86,1	83,0	3,1	84,6	88,5	88,0	0,5	88,3
TVM	137,2	133,0	4,2	135,1	141,0	137,0	4,0	139,0
DTH	116,4	113,0	3,4	114,7	116,4	113,0	3,4	114,7
RDD	126,4	123,0	3,4	124,7	126,4	123,0	3,4	124,7
Σx			30,9	1148,6			30,8	1182,7
d			3,1	114,9			3,1	118,3
Σx^2			107,9				112,9	
S_{diff}			1,2				1,4	

Example of a table drawn up for calculation of differences and averages of 1st and 2nd order Mean Fx values arrived at using two different software systems, TPS and ROM, on the same reading passage. Mean difference (d), standard deviation of the difference, S_{diff}.

The Mean difference 'd' and S_{diff} are used to calculate the 'Limits of agreement' (L.A.) which give the values between which 95% of the difference values for each group of subjects' fall, i.e. the range of agreement between individual pairs of measurements shown within +/- 2SD of the 'Mean difference'.

The 'Mean difference' (d), SD_{diff} and 'Limits of agreement' are thus used to illustrate the degree to which two methods agree, can be used interchangeably and/or are sufficiently repeatable to allow reliable comparisons from one occasion to another to be made; alternatively, to compare results from different experiments using the same analysis method.

viii Results - Agreement.

Tables 14 and 15 show the results of the comparison of the TPS and the ROM systems for their degree of agreement (TPS1 vs ROM 1 and TPS2 vs ROM 2). The same gain setting was maintained and the analysis carried out on the same data, on the same occasion (1 or 2), using each system. Inspection of the Mean differences (d) of both the 1st order Means and of the 2nd order Means (Table 14 and 15, Fig. 64 and 65, p. 268) shows a consistent reduction from the TPS to the ROM analysis illustrated by the positive Mean differences (d) of around 3 Hz. The Standard deviation and the limits of agreement are not large but there are no instances of either Normal speakers nor T1 subjects showing a 0 - difference (Appendix 12, Tables 1 and 3, 5 and 7) and there is a systematic bias towards the ROM giving approximately 3 Hz lower Mean values (Table 14 and 15).

Had this bias been greater, it would have warranted further investigation, to find out which system gives the 'true' value, but it is so small as to be insignificant and well within our stated acceptable differences of 5 Hz within the same analysis system, also, due to the ROM statistics not calculating fractions, +/- 1 Hz of this difference can be accounted for by this lack of precision.

Comparing the %TS between methods (Table 14 and 15, Fig. 64 b and 65 b) there is very good agreement resulting in small Mean differences (d), SD_{diff} and narrow Limits of Agreement (Appendix 12, Table 2 and 4, 6 and 8). The incrementation of the logarithmically spaced frequency bins during analysis, evidently happens at much the same rate with TPS and ROM, provided the gain setting is the same.

TABLE 14
TPS1 vs ROM 1

READING

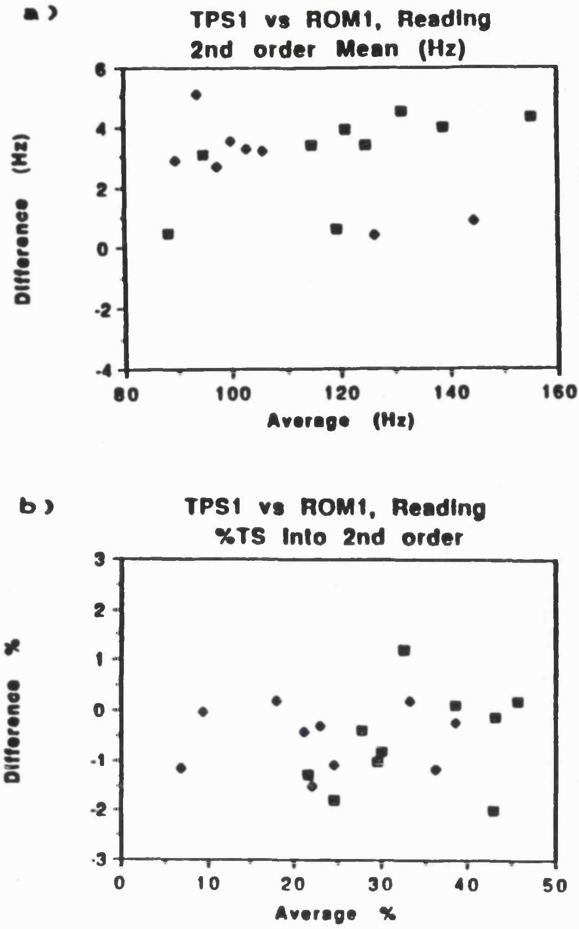
		Normal Speakers	T1 subjects
1st order Mean (Hz)	d	3,1	3,6
	SD_{diff}	1,2	2,3
	L.A.	5,5 to -0,7	8,2 to -1,0
2nd order Mean (Hz)	d	3,1	2,8
	SD_{diff}	1,4	1,3
	L.A.	5,9 to -0,3	5,4 to -0,2
%TS	d	-0,6	-0,6
	SD_{diff}	1,1	0,6
	L.A.	1,6 to -2,8	0,6 to -1,8

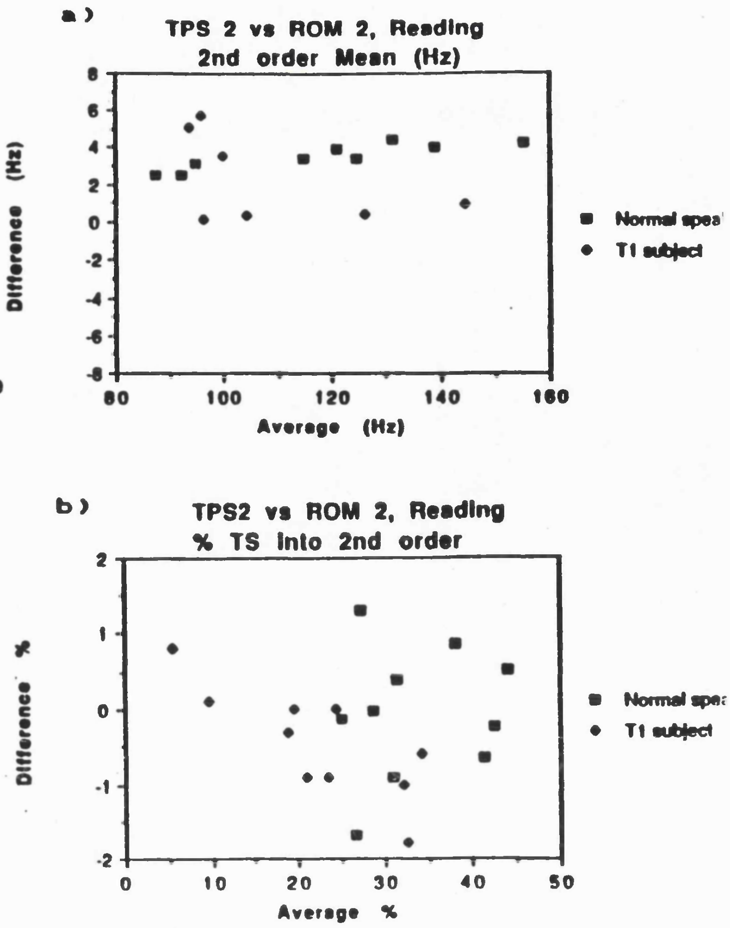
TABLE 15
TPS 2 vs ROM 2

READING

		Normal Speakers	T1 subjects
1st order Mean (Hz)	d	3,5	3,4
	SD_{diff}	0,5	0,9
	L.A.	4,5 to -2,5	5,2 to -1,6
2nd order Mean (Hz)	d	3,6	2,5
	SD_{diff}	0,7	2,0
	L.A.	5,0 to -2,2	6,5 to -1,5
%TS	d	-0,1	-0,5
	SD_{diff}	0,8	0,7
	L.A.	1,6 to -1,8	1,0 to -1,9

Tables showing agreement between Fx measurements on the same recorded reading task using the TPS and the ROM analysis systems on occasions 1 and 2. Mean differences (d), Standard deviation of these differences (Sdiff) and Limits of agreement (L.A.).





Agreement between the TPS and ROM systems
on the first and second occasion, 1 and 2.
READING a) 2nd order Mean Fx b) %TS

Figure 64

Figure 65

ix Results - Repeatability.

Comparing frequency and regularity measurements arrived at by repeating the analysis of the same reading samples, by means of TPS and ROM, on a second occasion and this time also looking at 1st and 2nd order Modal values, there is very good repeatability of the measurements within both systems (Tables 16 and 17, Figures 66 and 67 a-c).

Slightly different input gain settings on the two occasions have not unduly affected the measurements except in the case of 1st order Mode for T1 subjects (Table 16 and 17). Inspection of the raw data in Appendix 12 Tables 13 a and 19, reveals that the discrepancy is on account of one speaker, subject 38, whose 1st order Mode measured on the second occasion is considerably lower than on the first occasion. Different gain settings, from occasion 1 to occasion 2, have affected the first order Mode more than any other measure.

The Mode being more sensitive to different gain settings than the Mean, may be on account of the fact that it reflects the most frequently registered fundamental frequency in the sample irrespective of frequency range or intonation pattern. It is not an average of a set of values like the Mean. In the case of subject 38, these values are produced by a very abnormal and poorly controlled voice after radiotherapy, also reflected in a very poor Lx signal, and carries over a very small %TS into 2nd order (Appendix 12, Table 14). The removal of extreme irregularity from the analysed samples in 2nd order improves repeatability, however, which is well within our accepted limits. 2nd order Mode does however, show greater SD_{diff} values than the Mean and thereby wider Limits of agreement (Tables 16 and 17).

In both TPS and ROM analyses the majority of subjects show a 0-difference between Means from one occasion to the other (Figs. 66 a,b and 67 a,b). The widest 'Limits of agreement' 4.3 to -5.7, are shown by the T1 subjects' 1st order Means measured by TPS (Table 16). These limits are, however, just acceptable and either method shows good repeatability as far as Fx Means are concerned, provided the

TABLE 16

TPS 1 vs TPS 2

		READING	
		Normal Speakers	T1 subjects
1st order Mean (Hz)	d	0,02	-0,7
	SD_{diff}	1,9	2,5
	L.A.	3,8 to -3,8	4,3 to - 5,7
2nd order Mean (Hz)	d	-0,1	1,1
	SD_{diff}	1,4	1,4
	L.A.	2,7 to - 2,9	3,9 to - 1,7
1st order Mode (Hz)	d	-0,4	9,4
	SD_{diff}	1,1	27,3
	L.A.	1,9 to - 2,6	64,0 to - 45,2
2nd order Mode (Hz)	d	1,7	-1,0
	SD_{diff}	4,4	3,1
	L.A.	10,5 to - 7,1	7,1 to - 5,2
%TS	d	-0,2	1,2
	SD_{diff}	2,2	2,3
	L.A.	4,2 to - 4,6	5,8 to - 3,4

Repeatability of Fx measurments on the same recorded reading task using the TPS system on two different occasions 1 and 2. Mean differences (d), Standard deviation of these differences (SD_{diff}) and Limits of agreement (L.A.).

TABLE 17

ROM 1 vs ROM 2

READING

		Normal Speakers	T1 subjects
1st order Mean (Hz)	d	0,4	-1,2
	SD_{diff}	1,5	2,2
	L,A,	3,4 to - 2,6	3,2 to - 5,6
2nd order Mean (Hz)	d	0,4	1,1
	SD_{diff}	0,8	1,5
	L,A,	2,0 to - 1,2	4,1 to - 1,9
1st order Mode (Hz)	d	-0,6	8,9
	SD_{diff}	2,4	26,5
	L,A,	4,2 to - 5,4	61,9 to - 44,1
2nd order Mode (Hz)	d	1,5	0,3
	SD_{diff}	5,0	1,0
	L,A,	11,1 to - 8,9	2,3 to - 1,7
%TS	d	0,06	1,3
	SD_{diff}	1,8	2,0
	L,A,	3,7 to - 3,5	5,3 to - 2,7

Repeatability of Fx measurements on the same recorded reading task using the ROM system on two different occasions 1 and 2, Mean differences (d), Standard deviation of these differences (SD_{diff}) and Limits of agreement (L,A,),

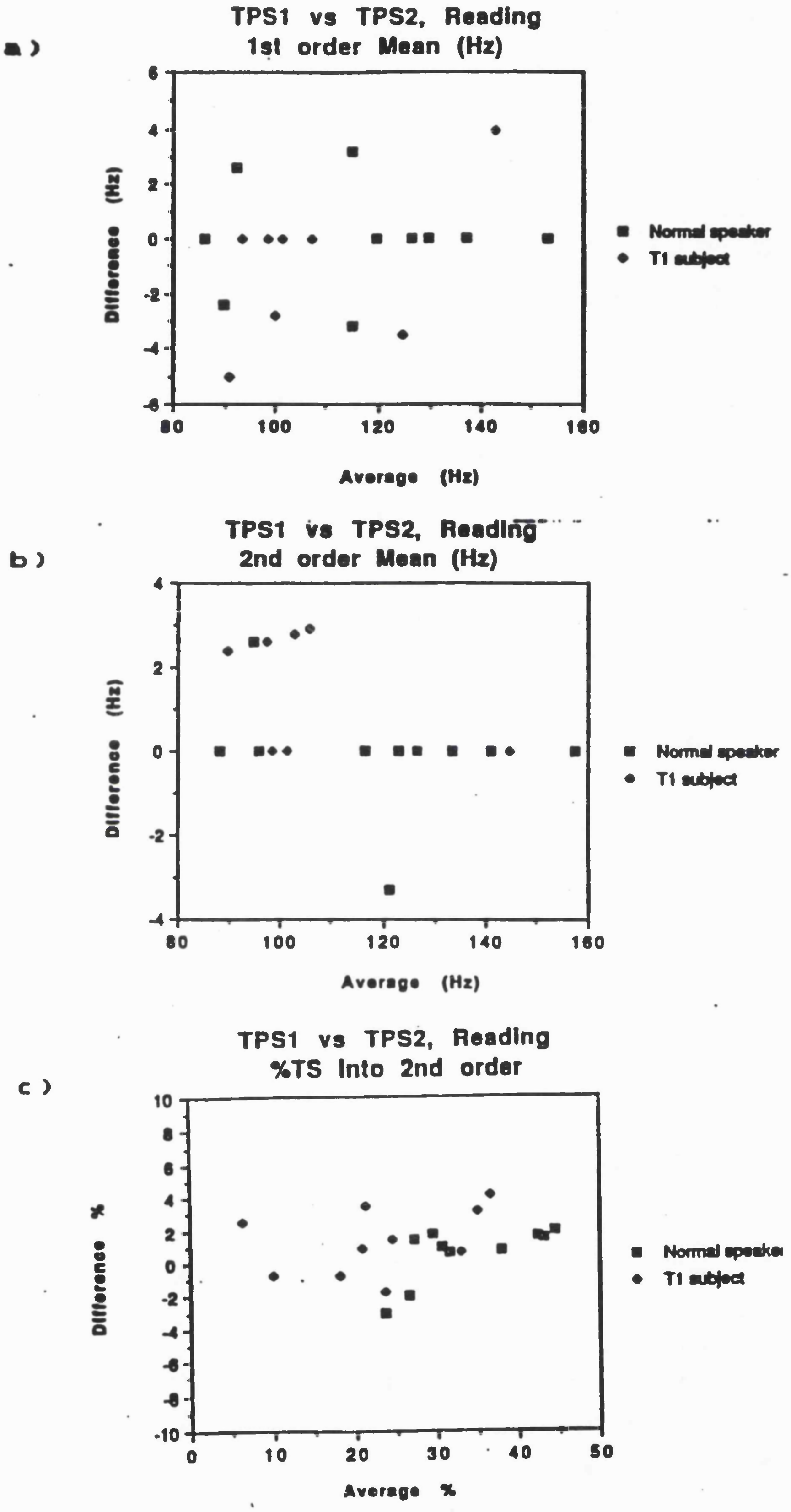


Figure 66 - Repeatability, TPS 1 vs TPS 2, Reading
a) 1st order Mean Fx b) 2nd order Mean Fx
c) %TS

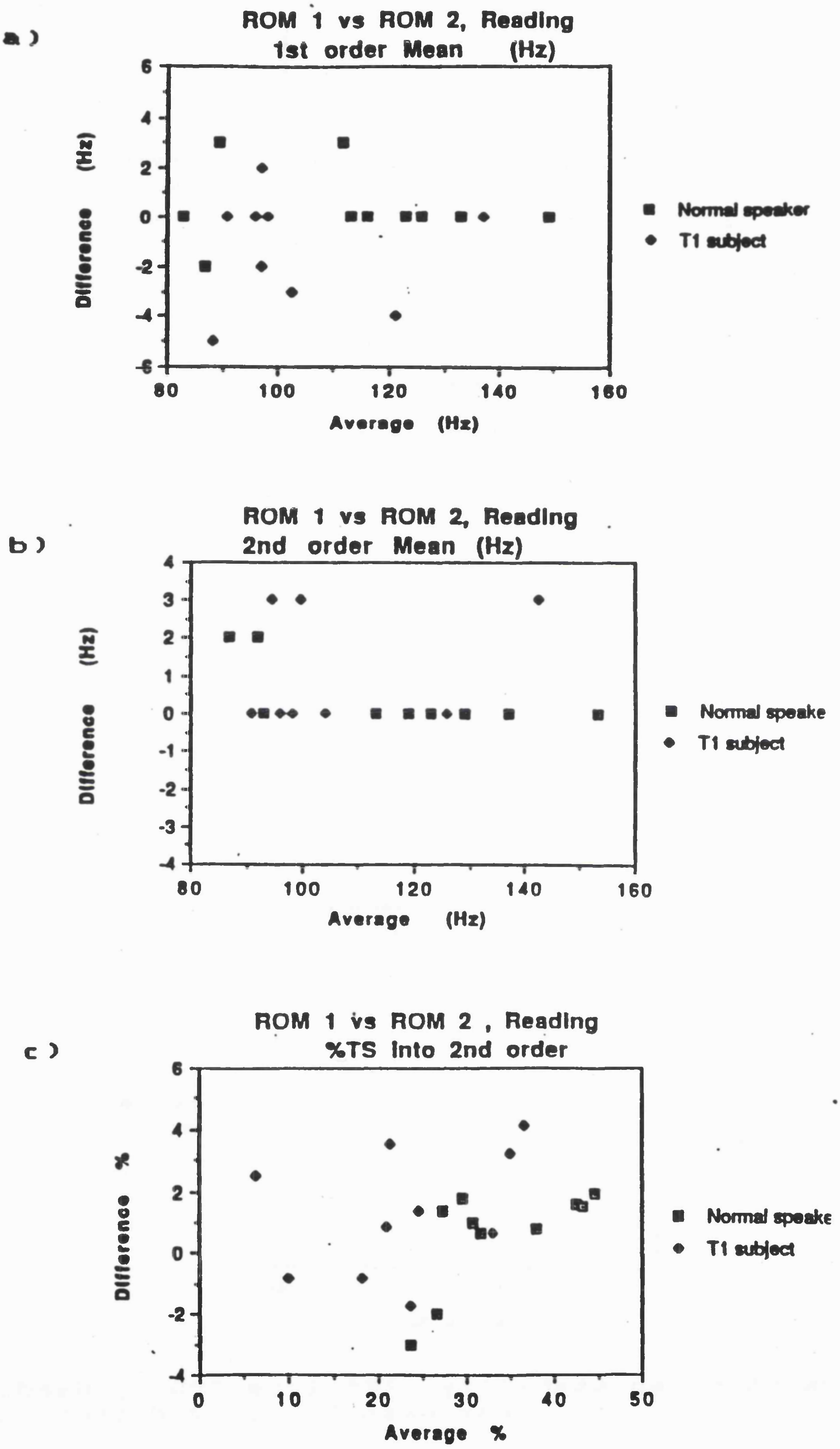


Figure 67 - Repeatability, ROM 1 vs ROM 2, Reading
a) 1st order Mean Fx b) 2nd order Mea
c) %TS

input gain is carefully calibrated on each occasion before analysis is begun.

Regularity of vocal fold vibration %TS, shows very good repeatability within both systems (Table 16 and 17). The figures 66 c and 67 c illustrate at a glance a tendency for T1 subjects' average %TS to be lower than the Normals'. This difference will be shown to be highly significant. There is less regularity of vocal fold vibration in the irradiated subjects' voices.

x Conclusion

The Agreement between the TPS and ROM systems is acceptable (Table 14 and 15). Provided care is taken in the calibration of the input gain level before analysis, either system can be used to give reliable and comparable Mean and %TS values. The ROM analysis, due to its more gross calculation of frequency values, shows consistently lower Mean values by 3 Hz. The difference is insignificant, however, and does not invalidate the use of the Dx ROM for clinical comparisons. For research purposes, however, the greater precision and detail offered by the TPS system seemed an advantage.

For Normal speakers both 1st and 2nd order Mean and Mode show good repeatability irrespective of which analysis system is used (Table 16 and 17). For T1 subjects, some of whom have very abnormal voice quality (e.g. Subjects 14 and 38), 1st order Modes show poor repeatability compared to 1st order Means (Table 16 and 17). 2nd order Mode shows much improved repeatability for T1 subjects but not as good as 2nd order Means using the TPS system (Table 16).

We conclude that the central measures least sensitive to different recording and analysis artefacts are 1st and 2nd order Means, with 2nd order showing a slight advantage over 1st order Mean.

Our regularity measure %TS, shows good agreement between systems (Table 14 and 15 Fig. 64 b and 65 b) and good repeatabiltiy within systems (Table 16 and 17, Figure 66 c and 67 c).

B i Determining the agreement between TPS and PCLx measurements and Repeatability of PCLx.

In the final stages of the work reported in this dissertation, new instrumentation and a new analysis system became available: the most recent version of the Laryngograph processor, a PC and new software for analysing Lx data, PCLx, as well as new recording equipment were acquired. This replaced the old instrumentation and software that had been used for almost ten years. Data collection of radiotherapy subjects' voices continued after the acquisition of the new equipment. It therefore seemed prudent to check the agreement between TPS and PCLx and the repeatability of PCLx, if comparison of data resulting from analyses using both PCLx and the TPS program was going to be made.

With a total change of both recording equipment, Lx processor and software for Fx analysis, there may be several sources of possible measurement error and differences. The aim was to find out if there was any evidence of small but consistent bias as was shown by the comparison of Fx Means between the TPS and the ROM programs.

The task now was to find out the degree of agreement between measurements using the TPS and the PCLx programs on the same recorded data as had been used for the comparison of TPS and the ROM previously described, and the repeatability of the PCLx analysis on the same data on two occasions.

ii Equipment

A new Laryngograph Processor was linked to a Sony stereo cassette deck TC - WR 570 used to record speech and reading samples; a Sony digital audio tape deck DTC - 77 ES was used to record and play back Lx waveforms for analysis. A personal computer 'Positive Logic 386 DX PC' replaced the BBC Master previously used, and enabled data analysis by means of the PCLx 'Pitch' and 'Wave' programs. A Laser Jet III was used for printing hard copy for storage of print outs in subjects' notes.

iii Subjects, voice samples and calibration of input level.

It was not possible to carry out repeated analyses using both systems again, on the same and different occasions, as the old equipment had been 'decommissioned' and was no longer in use. To determine the degree of agreement between the systems it was therefore decided to make use of the data from the TPS 1 analysis of the reading samples, and this time also spontaneous speech, produced by the same randomised 'Normal speakers' and T1 subjects described previously. The TPS measurements would be compared to measurements arrived at using the PCLx system (PCLx 1). To assess the repeatability of the PCLx analysis, the speech and reading samples were analysed again on a second occasion (PCLx 2).

It was decided to compare both spontaneous speech and reading samples, to assess the effect on measurements using exactly the same data sample as in reading, and slightly different and usually larger samples, as was the case in conversational speech. We used speech data analysed by the experimenter on an earlier occasion using TPS, and aimed to analyse not fewer than 6000 Tx samples from the speech recording on PCLx; this could be judged approximately from the level of the memory store on the PC. There was no way of knowing exactly what part of the speech sample had been analysed on TPS, thereby it was not possible to choose exactly the same sample for PCLx analysis. The size of the individual speech samples can be found in Appendix 12, Tables 32 and 36.

Calibration of the input level for the PCLx analysis was performed as described above in the section comparing the ROM and the TPS systems. The fundamental frequency contours displayed on the VDU screen in the 'PC pitch' program were used to achieve a level of input that gave the most representative Tx sample for analysis, avoiding excessive 'noise' in the signal. Once a satisfactory level was achieved, the settings on the Laryngograph processor were left alone and the recorded sample was analysed by PCLx.

As the TPS and PCLx analyses were not carried out on the same occasion, as had been the TPS 1 and ROM 1 and the TPS 2 and ROM 2 analyses, using the same gain settings for both analyses, we expected to find greater differences between TPS and PCLx measurements.

iv Known differences between the TPS and the PCLx programs.

There is one major known difference between the TPS program and PCLx. It is the way they select Fx samples to be carried over into second order. In the earlier programs, only those samples were selected where two adjacent ones fell into the same frequency 'bin' along the logarithmic frequency scale. A second order histogram illustrates the amount of regularity in a voice (Fig. 44, p. 203), but the use of a logarithmic frequency scale negatively biased a low pitched voice, and tended to carry over a smaller proportion of the total sample into the second order distribution than for a high pitched voice, although the low pitched voice did not necessarily sound harsh, rough or abnormal. This was described in the section on %TS as a measure of voice regularity (Chapter VII).

The PCLx program calculates a measure of 'IRREGULARITY', expressed as the proportion of the total recorded Fx samples that differ from their neighbours by more than ten percent. These are excluded from 2nd order distributions (Ball, Faulkner and Fourcin, 1990). It is expressed as a percentage and is not, like %TS, dependent on 'bin widths' (Chapter VII).

As %TS has been used throughout this study as a measure of REGULARITY, we will continue to calculate and refer to this measure. It will be demonstrated how the new way of admitting Fx samples into second order results in slightly increased %TS.

RESULTS

v Agreement between TPS and PCLx - 1st and 2nd order Means.

Table 18 shows the agreement between TPS and PCLx for 1st and 2nd order Means. The raw data on which these calculations are based is found in Appendix 12, Table 21 and 25. It is important to remember in the following account that a positive mean difference (d) means a reduction in values from the TPS to the PCLx analysis, a negative 'd' means an increase from one analysis to the next (see Table 13, p. 265).

TABLE 18

A) TPS vs PCLx, READING

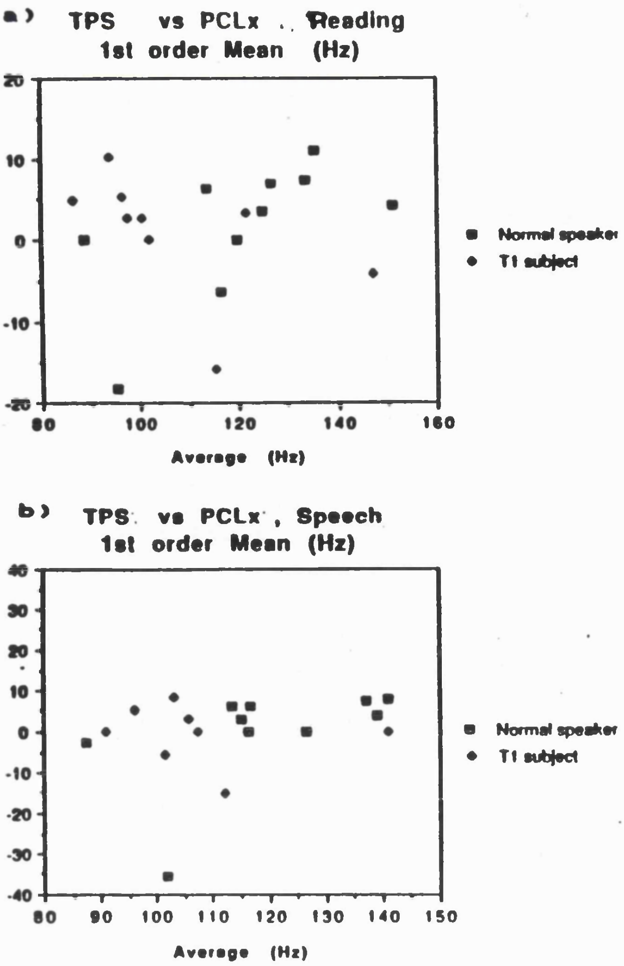
		Normal Speakers	T1 subjects
1st order Mean (Hz)	d	1.4	1.0
	SD_{diff}	8.5	7.4
	L.A.	18.3 to -15.5	15.7 to -13.8
2nd order Mean (Hz)	d	3.2	4.3
	SD_{diff}	5.0	4.0
	L.A.	13.1 to -6.7	12.2 to -3.7

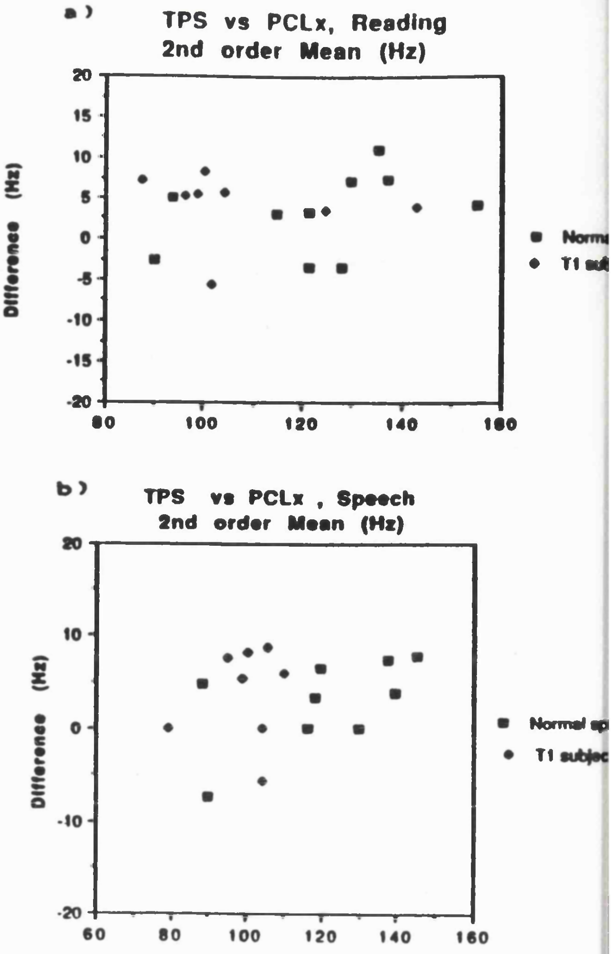
B) TPS vs PCLx, SPEECH

		Normal Speakers	T1 subjects
1st order Mean (Hz)	d	-0.4	-0.1
	SD_{diff}	12.9	6.8
	L.A.	25.5 to -26.2	13.5 to -13.7
2nd order Mean (Hz)	d	2.5	4.1
	SD_{diff}	5.2	4.7
	L.A.	12.9 to -7.9	13.5 to -5.3

Agreement between Fx measurements on the same recorded tasks using the TPS and the PCLx system of analysis on occasion 1. Mean differences (d), Standard deviation of these differences (SD_{diff}) and Limits of agreement (L.A.).

1st order Means for Reading and Speech (Table 18 A and B) show very small mean differences (d) between TPS and PCLx for both groups indicating there is no particular bias and generally good agreement between the systems, despite the speech samples analysed being slightly different. The differences are actually smaller in speech than in reading (Table 18), which may be due to the larger speech samples but also to two particularly large differences from one analysis to the other, in the reading task (Appendix 12, Tables 21 and 25). This applies to both Normal speakers and the T1 subjects





Agreement between the TPS and PCLx systems
1st (left) and 2nd order (right) Mean Fx.

a) Reading (top)
b) Speech (bottom)

Figure 68

Figure 69

(Fig.68 a) and for one speaker from each group in the speech task (Fig. 68 b) (Appendix 12, Tables 29 and 33).

This has the effect of substantially increasing the standard deviation of the differences SD_{diff} i.e the degree to which the methods agree for an individual within a sample (Altman, 1991) and the Limits of agreement (Table 18).

However, looking at the figures 68 a and b, it is evident that most 1st order Means tend towards being lower in the PCLx analysis, as shown by the tendency to positive differences between means, but with one extreme negative difference registered for both the 'Normal speakers' and the T1 subjects in reading and in speech.

2nd order Means (Fig. 69 a and b) show larger, systematically positive mean differences (d) than first order, both in reading (Table 18 A) and in speech (B) indicating a tendency for 2nd order Means also to be lower in PCLx analyses than in TPS (Appendix 12, Table 21, 25, 29, 33). The different way of admitting samples into second order employed by the PCLx system seems to result in an increase in low frequency samples being carried over, reducing the fundamental frequency of the 2nd order Means (Mean difference=d) for both groups, by approximately 3 Hz in the Normal speakers and by 4 Hz in the T1 speakers (Table 18, Fig. 69 a and b). This is not surprising as T1 speakers tend towards lower pitched voices. The different sample sizes between reading and speech do not seem to make any difference to second order Means, nor the fact that the speech samples analysed by TPS and PCLx would have been slightly different.

The standard deviation around the mean differences and Limits of Agreement (L.A.) (Table 18) are considerably reduced in 2nd compared to 1st order. The frequency range within the Limits of Agreement (+/- 2SD) are correspondingly narrower, +/- 10 Hz and +/- 8-9 Hz respectively for Normal and T1 speakers. This may be used as an indication that second order measures are more reliable for comparisons of measurements carried out using different analysis systems than first order measures. Second order seems to have got rid of the extreme differences encountered in first order (Fig. 69 a and b).

vi Agreement - 1st and 2nd order Modes.

Table 19 shows the agreement between TPS and PCLX analysing 1st and 2nd order Modes. It demonstrates a greater sensitivity of the Mode to extremely abnormal voice quality as demonstrated by T1 subject 38 particularly in 1st order (Appendix 12, Table 26). This results in an extremely large mean difference (d) and SD for T1 subjects both in 1st and 2nd order Modal values. The mean difference (d) between the systems is also larger for Normal speakers in 1st order Mode than in 1st order Mean (cf. Table 18). The removal of irregularity in 2nd order helps to reduce the mean difference (d) in second order Mode for T1 speakers (Table 19) but both 'd' and SD_{diff} are larger than 5 Hz, which was our arbitrary acceptable limit for differences between measurements. The situation is worse in the Speech sample, when subject 38 distorts the values in 2nd order (Table 19)(Appendix 12, Table 34). 1st order Mode in the Speech sample, however, shows smaller SD than 1st order Mean (Table 18).

Apart from being more sensitive to extremes in voice quality, the Mode seems more sensitive than the Mean to different input levels and slightly different voice samples analysed on different occasions as is the case in this comparison.

TABLE 19

A) TPS vs PCLx, READING

		Normal Speakers	T1 subjects
1st order Mode (Hz)	d	4,7	16,4
	SD_{diff}	3,8	28,4
	L.A.	12,3 to -2,9	73,2 to -40,4
2nd order Mode (Hz)	d	5,7	6,9
	SD_{diff}	65	7,6
	L.A.	12.2 to -0.8	22,1 to -8,3

B) TPS vs PCLx, SPEECH

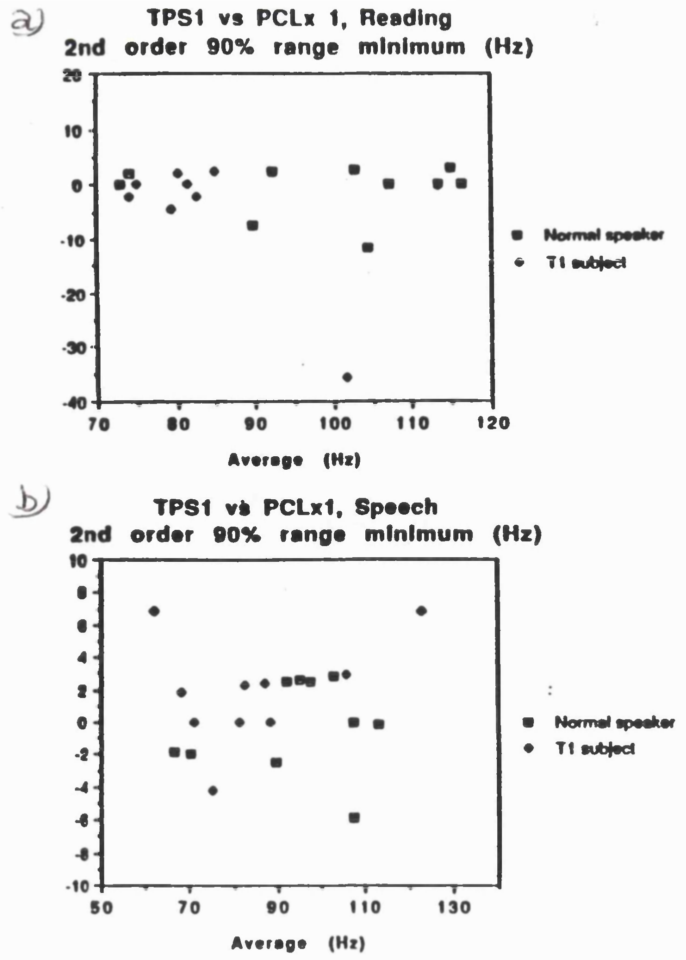
		Normal Speakers	T1 subjects
1st order Mode (Hz)	d	1,8	3,9
	SD_{diff}	8,1	5,3
	L.A.	18,0 to -14,4	14,5 to -6,7
2nd order Mode (Hz)	d	2,0	-3,3
	SD_{diff}	6,1	31,7
	L.A.	14,2 to -10,2	60,1 to -66,7

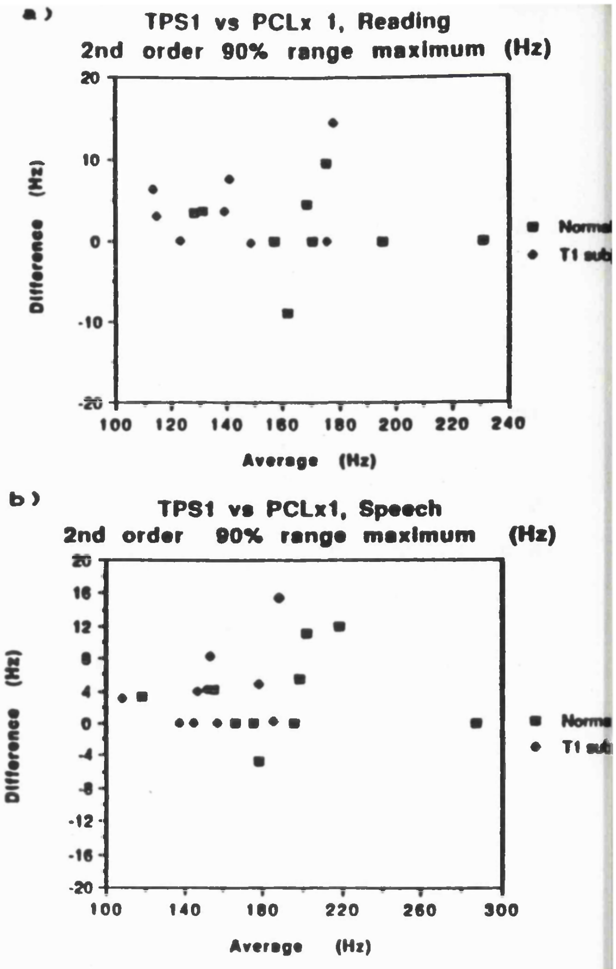
Agreement between Fx measurements on the same recorded tasks using the TPS and the PCLx system of analysis on occasion 1. Mean differences (d), Standard deviation of these differences (SD_{diff}) and Limits of agreement (L.A.).

vii Agreement - 2nd order 90 % range minima and maxima.

As we have found a slight advantage of 2nd order central measurements over 1st order in the agreement between analysis systems, we decided to test the agreement between 2nd order 90 % Fx range measurements.

Second order 90% range minimum comparisons (Table 20, Fig. 70 a and b) show very small negative mean differences for both reading and speech in Normal speakers, but almost 4 Hz higher range minimum in T1 subjects using the PCLx system analysing the reading sample (Table 20), very large SD and wide limits of agreement. The reason is one extremely deviant voice and resulting Fx values in one T1 subject (Subject 14)(Fig. 70 a) (Appendix 12, Table 27). Considerably smaller





Agreement between the TPS and PCLx systems
2nd order 90 % Range, Minimum range Fx (left)
and Maximum range Fx (right).

a) Reading (top)
b) Speech (bottom)

Figure 70

Figure 71

SD_{diff} and narrower LA are arrived at for both normal and T1 speakers in analysis of the speech sample (Table 20) (Appendix 12, Table 31 and 35).

The PCLx analysis seems to lower the range maximum slightly (Table 20). As in the range minimum calculations, the very deviant values of one particular T1 subject reading aloud distorts the figures (Fig. 71 a) (Appendix 12, Table 27). The speech samples show similar values for both groups, namely a lowering of the range maximum of 3 and 4 Hz (Table 20).

TABLE 20

A) TPS vs PCLx, READING

		Normal Speakers	T1 subjects
2nd order 90 % range Minimum (Hz)	d	-0,8	-4,0
	SD_{diff}	4,8	11,9
	L.A.	8,8 to -10,5	19,9 to -27,9
2nd order 90 % range Maximum (Hz)	d	1,2	3,8
	SD_{diff}	4,7	4,7
	L.A.	10,7 to -8,2	13,3 to -5,6

B) TPS vs PCLx, SPEECH

		Normal Speakers	T1 subjects
2nd order 90 % range Minimum (Hz)	d	-0,2	1,9
	SD_{diff}	2,9	2,5
	L.A.	5,6 to -5,8	6,9 to -3,1
2nd order 90 % range Maximum (Hz)	d	3,0	4,0
	SD_{diff}	5,3	4,9
	L.A.	13,6 to -7,5	13,8 to -5,9

Agreement between Fx measurements on the same recorded tasks using the TPS and the PCLx system of analysis on occasion 1. Mean differences (d), Standard deviation of these differences (SD_{diff}) and Limits of agreement (L.A.).

viii Agreement - %TS

The negative mean differences between TPS and PCLx %TS calculations (Table 21, Fig. 72 a and b) demonstrate the expected slight increase in the proportion of the Total sample carried into second order through inclusion of all samples within ten percent frequency of their adjacent neighbours, as is used in the PCLx system. The increase is almost 6% for normal speakers reading aloud, but only around 2 % for T1 readers (Table 21 A) and for both groups' speech samples (B). This is the result of some considerably higher %TS values arrived at in the reading task in the group of Normal speakers using the PCLx system (Fig.72 a). Their more regular vocal fold vibration and higher pitch in reading than in conversation results in much increased %TS (Appendix 12, Tables 24 and 32). This explains the larger standard deviation among the Normal speakers and corresponding wider limits of agreement and range of %TS values within the Limits of Agreement (= +/-2 SD of the Mean difference, d)which is +/- 16-17 %, almost twice that of the T1 speakers, +/- 9-10 %, in reading aloud (Table 21 a).

TABLE 21

A) TPS vs PCLx, READING

		Normal Speakers	T1 subjects
%TS	d	-5.9	-2.4
	SD_{diff}	8.5	4.5
	L.A.	11.2 to -23.0	6.6 to -11.5

B) TPS vs PCLx, SPEECH

		Normal Speakers	T1 subjects
%TS	d	-2.1	-2.1
	SD_{diff}	8.2	5.3
	L.A.	14.1 to -18.5	8.5 to -12.7

Agreement between Fx measurements on the same recorded tasks using the TPS and the PCLx system of analysis on occasion 1. Mean differences (d), Standard deviation of these differences (SD_{diff}) and Limits of agreement (L.A.).

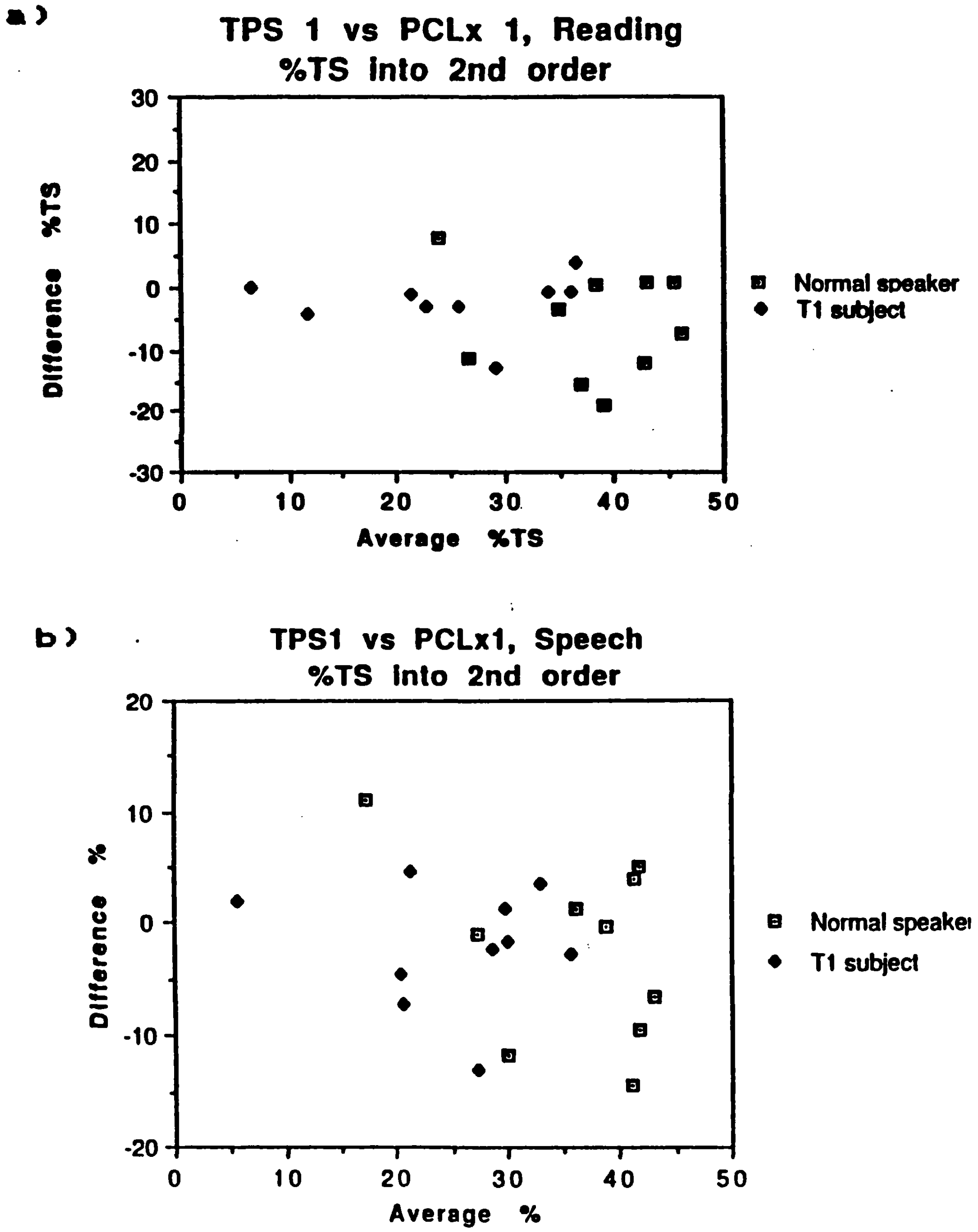


Figure 72 – Agreement TPS vs PCLx, %TS
a) Reading b) Speech

This amount of difference between the systems, of %TS carried into second order (Table 21) indicates that some caution must be exercised when interpreting %TS measurements as indicators of degree of regularity of vocal fold vibration, arrived at using different analysis systems. It seems however to be less of a problem when comparing the T1 subjects' voices than the Normal speakers', who show considerably greater variation in %TS between the systems.

%TS is still a useful measure, as it indicates the proportion of the original sample carried into second order on which second order Fx measurements are based.

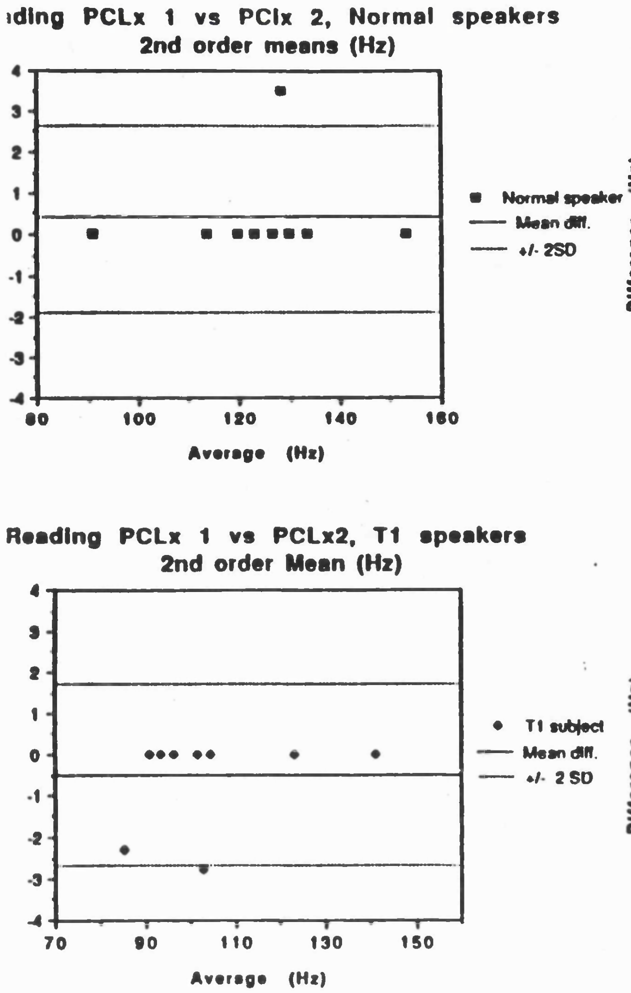
ix Repeatability of PCLx measures.

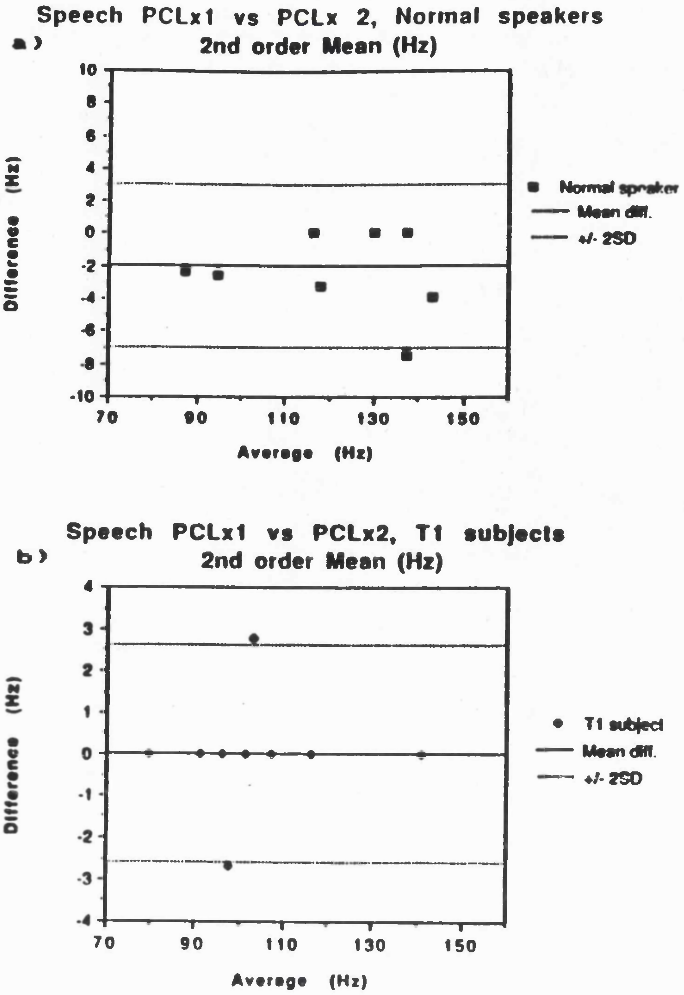
Repeated analyses using PCLx, of the same reading samples and approximately the same spontaneous speech samples, on two occasions, reveal very small Mean differences (d) for all parameters (Table 22-24 a and b), except in 2nd order 90% range maximum in the Speech sample (Table 23).

1st order Means show a zero difference on both occasions for the majority (7-8/10) of Normal speakers (Fig. 73 a and 74 a) and for 4-5/10 T1 subjects for both Reading and Speech (Fig. 73 b and 74 b).

2nd order Means show a zero difference between occasions for 8/10 of T1 subjects in both Reading and Speech (Fig. 75 b and 76 b), and for 9/10 Normal speakers in reading (Fig. 75 a) and 5/10 in speech (Fig. 76 a) (only 8 speakers are shown but there are 3 at '0 difference' and 116.4 Hz). This may be due to the slightly different speech sample used on the second occasion, but this was also the case for the T1 subjects, and it does not seem to have affected their measurements. On the other hand, the Normal speakers may use a wider range generally, which may influence their Means more when the samples are not exactly the same. This is confirmed later in the comparison of 90% range Maxima.

Comparing the repeatability of the Reading Means and Modes in 1st and 2nd order, there are larger, but still minimal mean differences



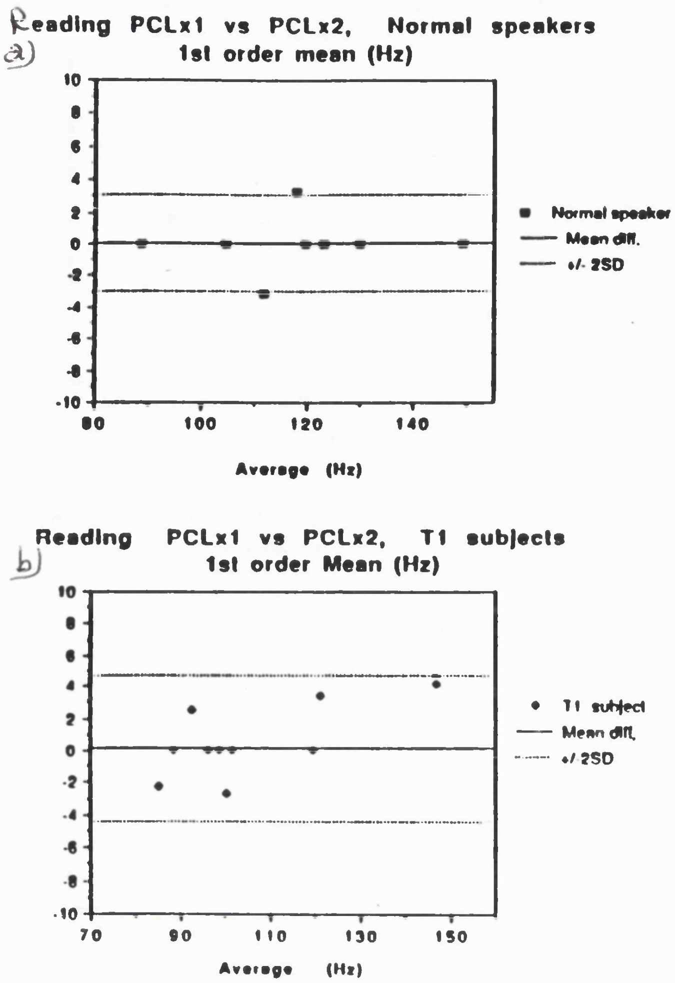


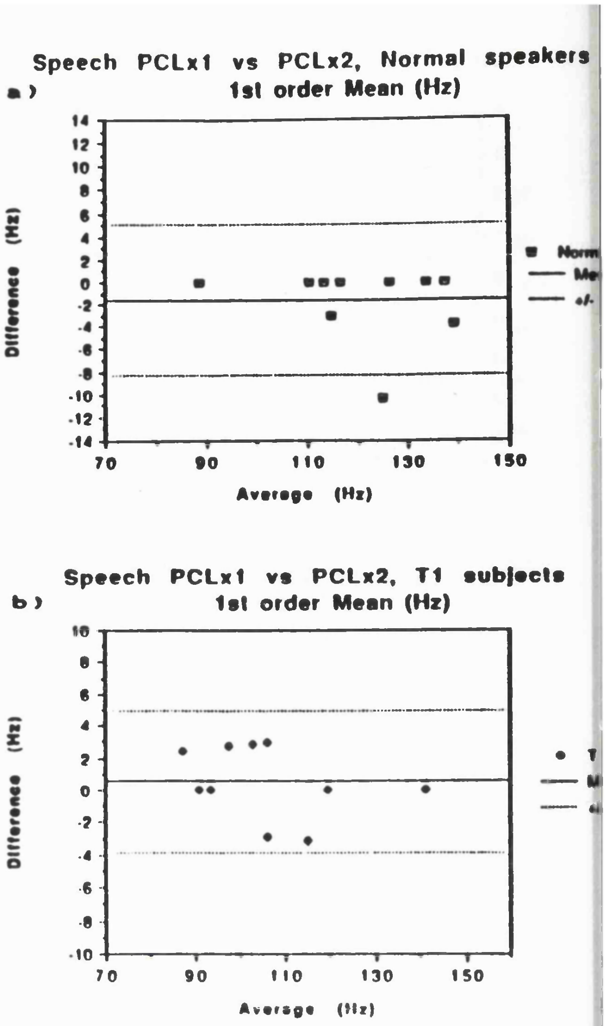
Repeatability of PCLx analyses 2nd order Mean
Reading (left), Speech (right).

a) Normal speakers (top)
b) T1 subjects (bottom)

Figure 75

Figure 76





Repeatability of PCLx analyses 1st order Mean
Reading (left), Speech (right).

a) Normal speakers (top)
b) T1 subjects (bottom)

Figure 73

Figure 74

(d) and SD_{diff} for 1st and 2nd order Modes than Means for Normal speakers (Table 22). T1 subjects however, show very small Mean differences and SD_{diff} both in 1st and 2nd order (Table 22). For the Speech samples both the 1st and 2nd order Modes show very good repeatability for both groups of speakers (Table 23).

The 2nd order 90 % range minimum shows small differences between the two analysis occasions (Table 22 and 23, Fig. 77 and 78).

The slightly different samples used for the two speech analyses seem to influence the 90 % range maximum more than any other parameter. Only 2/10 Normal speakers show a zero difference from one analysis to the other. They tend to achieve higher range maxima (shown by negative differences) on the second analysis (Fig. 80 a), whereas for the T1 subjects, although a majority (6/10) show a zero difference between the two analyses, 4/10 show a tendency towards reduced range maximum in the second speech analysis (shown by positive differences) (Fig. 80 b). Most of the differences are small however, although extreme differences are shown in both Normal and T1 subjects' plots resulting in larger Mean differences, SD_{diff} and wider limits of agreement for the 90% range Maxima than other parameters (Table 22 and 23) (Appendix 12, Tables 31. 35, 39 and 43).

The 90 % range maximum is slightly problematic in comparisons of different speech and reading samples, as it is the most sensitive measure when considering whether the analysed samples are exactly the same. Even where they are, as in Reading, (Fig. 79 a and b, Table 22), the Normal speakers show their largest standard deviation and Mean difference.

TABLE 22

PCLx 1 vs PCLx 2

READING

		Normal Speakers	TI subjects
1st order Mean (Hz)	d	0,01	0,09
	SD_{diff}	1,5	2,3
	L.A.	3,0 to -3,0	4,7 to -4,5
2nd order Mean (Hz)	d	0,4	-0,5
	SD_{diff}	1,1	1,1
	L.A.	2,6 to -1,9	1,7 to -2,7
1st order Mode (Hz)	d	-1,2	-0,3
	SD_{diff}	3,8	1,01
	L.A.	6,4 to -8,8	1,7 to -2,3
2nd order Mode (Hz)	d	-0,2	-0,2
	SD_{diff}	1,7	0,7
	L.A.	3,2 to -3,6	1,3 to -1,7
2nd order 90 % range Minimum (Hz)	d	0,5	0,2
	SD_{diff}	1,8	2,0
	L.A.	4,1 to -3,1	4,2 to -3,8
2nd order 90 % range Maximum (Hz)	d	-3,0	-0,6
	SD_{diff}	5,3	2,3
	L.A.	7,7 to -13,7	4,0 to -5,2

Repeatability of Fx measurements on the same recorded reading task using the PCLx system on two different occasions 1 and 2. Mean differences (d), Standard deviation of these differences (SD_{diff}) and Limits of agreement (L.A.).

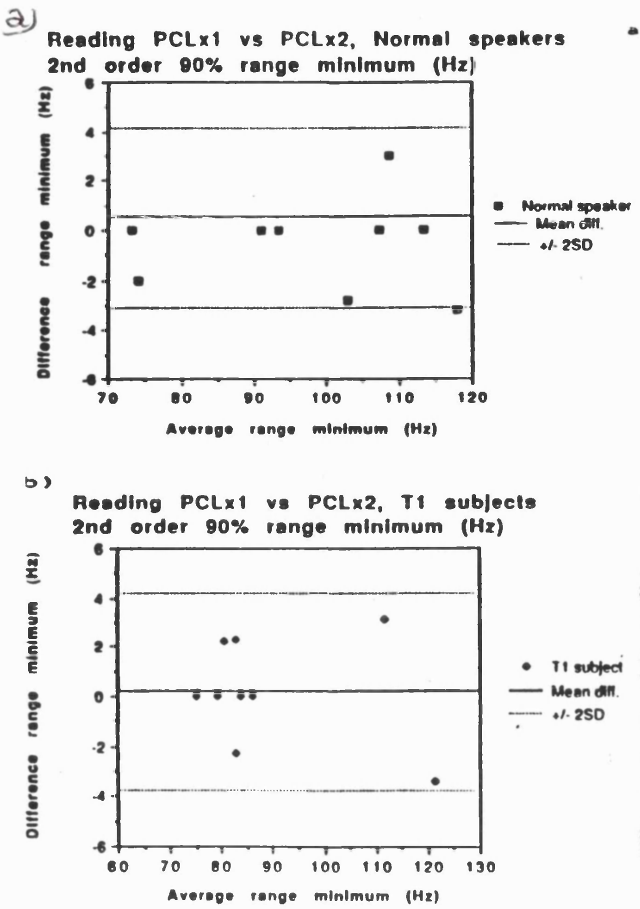
TABLE 23

PCLx 1 vs PCLx 2

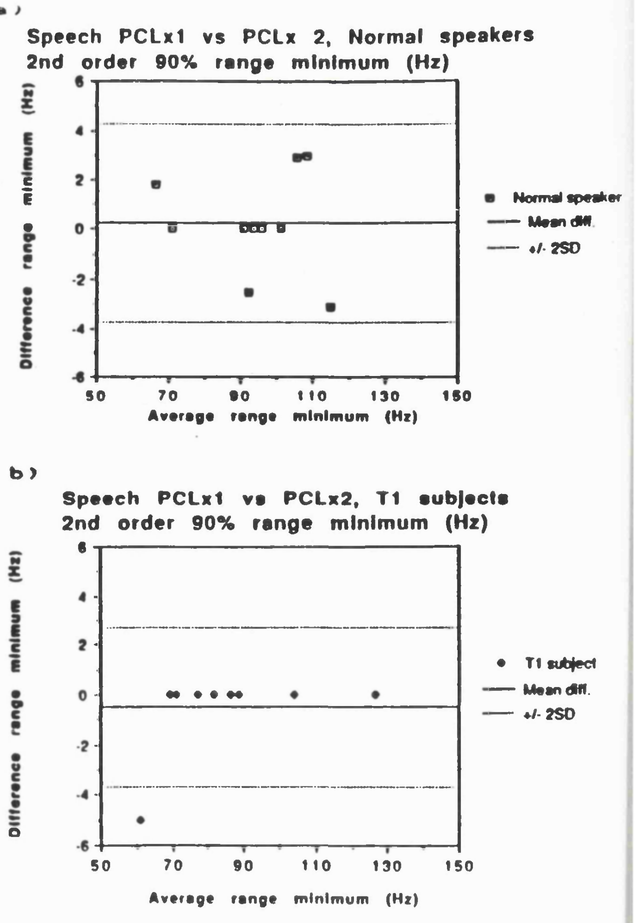
SPEECH

		Normal Speakers	T1 subjects
1st order Mean (Hz)	d	-1,7	0,5
	SD_{diff}	3,3	2,2
	L.A.	5,0 to -8,4	4,9 to -3,9
2nd order Mean (Hz)	d	-2,0	0,01
	SD_{diff}	2,5	1,3
	L.A.	3,0 to -7,0	2,6 to -2,6
1st order Mode (Hz)	d	1,3	0,8
	SD_{diff}	2,4	2,2
	L.A.	6,1 to -3,5	5,2 to -3,6
2nd order Mode (Hz)	d	1,3	-1,3
	SD_{diff}	3,2	2,2
	L.A.	7,7 to -5,1	3,1 to -5,7
2nd order 90 % range Minimum (Hz)	d	0,2	-0,5
	SD_{diff}	2,0	1,6
	L.A.	4,2 to -3,8	2,7 to -3,7
2nd order 90 % range Maximum (Hz)	d	-6,8	3,4
	SD_{diff}	9,1	5,4
	L.A.	11,4 to -25,0	14,2 to -7,4

Repeatability of Fx measurments on approximately the same speech task using the PCLx system on two different occasions 1 and 2. Mean differences (d), Standard deviation of these differences (SD_{diff}) and Limits of agreement (L.A.).


a)
Reading PCLx1 vs PCLx2, Normal speakers
2nd order 90% range minimum (Hz)

b)
Reading PCLx1 vs PCLx2, T1 subjects
2nd order 90% range minimum (Hz)

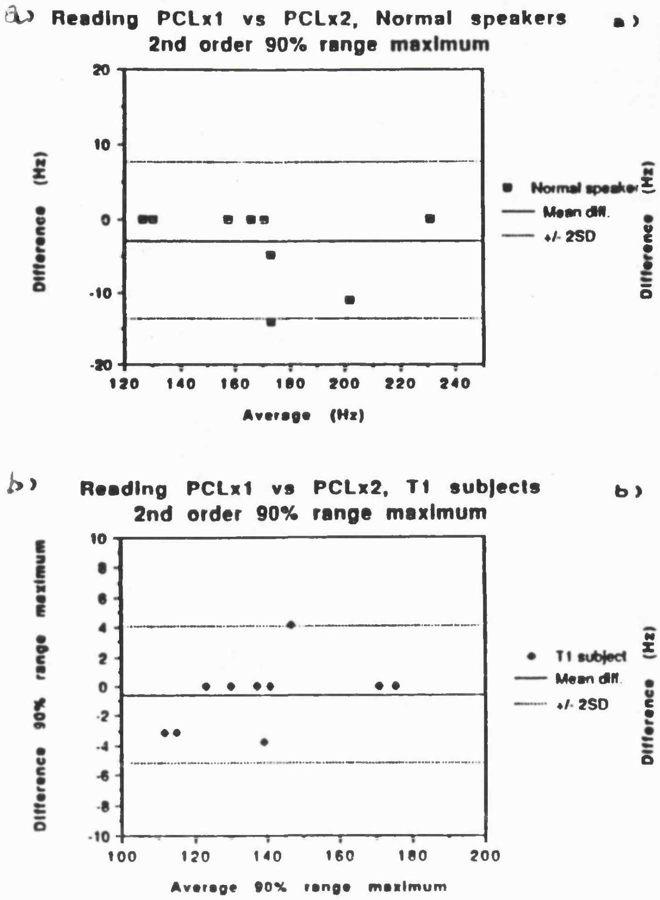

a)
Speech PCLx1 vs PCLx 2, Normal speakers
2nd order 90% range minimum (Hz)

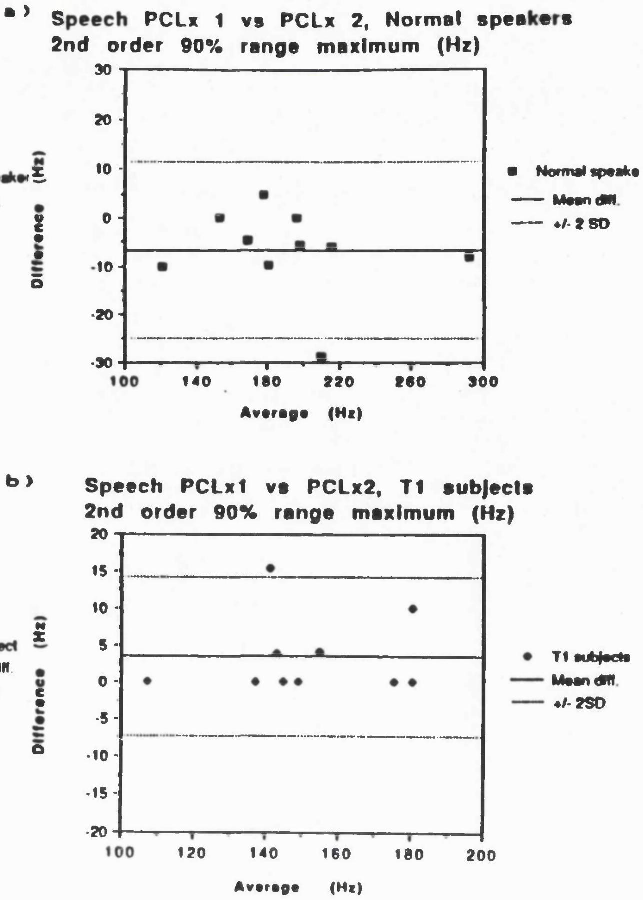
b)
Speech PCLx1 vs PCLx2, T1 subjects
2nd order 90% range minimum (Hz)

Repeatability of PCLx analyses, 90 % range Minimum Fx
Reading (left), Speech (right).

a) Normal speakers (top)
b) T1 subjects (bottom)

Figure 77 Figure 78





Repeatability of PCLx analyses, 90 % range Maximum Fx
Reading (left), Speech (right).

a) Normal speakers (top)
b) T1 subjects (bottom)

Figure 79

Figure 80

There is very little difference among all the speakers in the %TS carried into 2nd order on the two occasions, for both Reading and Speech (Table 24 A and B) as evidenced by the small Mean differences and narrow Limits of agreement (+/- 2SD), showing that this measure is quite robust, provided the same analysis program has been used.

Finally, the % Irregularity measure shows good repeatability. Mean differences between occasions are close to zero and the standard deviations and corresponding Limits of Agreement for these differences very small (Table 24 A and B). Not having used this measure before, we have no norms to judge what would be acceptable differences between measurements on the same sample from one analysis to another.

TABLE 24 A

PCLx 1 vs PCLx 2, READING

		Normal Speakers	T1 subjects
%TS	d	0,1	0,1
	SD_{diff}	1,7	0,9
	L.A.	3,5 to -3,3	1,9 to -1,8
% Irregularity	d	-0,1	0,02
	SD_{diff}	2,8	1,4
	L.A.	5,5 to -5,7	2,7 to -2,7

Repeatability of Regularity, %TS, and % Irregularity measurements on the same recorded reading task using the PCLx system on two different occasions, 1 and 2, Mean differences (d), Standard deviation of these differences (SD_{diff}) and Limits of agreement (L.A.)

Average % Irregularity for the group of Normal speakers are 13.9 % for Reading and 18.4 % for Speech. Corresponding % Irregularity for the T1 subjects are 29 % and 31.7 % respectively (Appendix 12, Table 44 and 52)

Having allowed a difference of 5% between repeated analyses for the % TS measure, this would be excessive for % Irregularity , but Mean differences and SD_{diff} are well below this limit (Table 24 A and B). Only one Normal speaker, JSt, (the same in both Reading and Speech) shows a difference between analyses of 5 and 5.5 % with an average of 14.3 % Irregularity in Reading and 22.1 % in Speech (Appendix 12, Table 40 and 48).

There does not seem to be any particular bias towards higher or lower values on the second analysis, for either Speech or Reading, so % Irregularity does not seem to be sensitive to sample size or to the samples being exactly the same (Table 24 A and B).

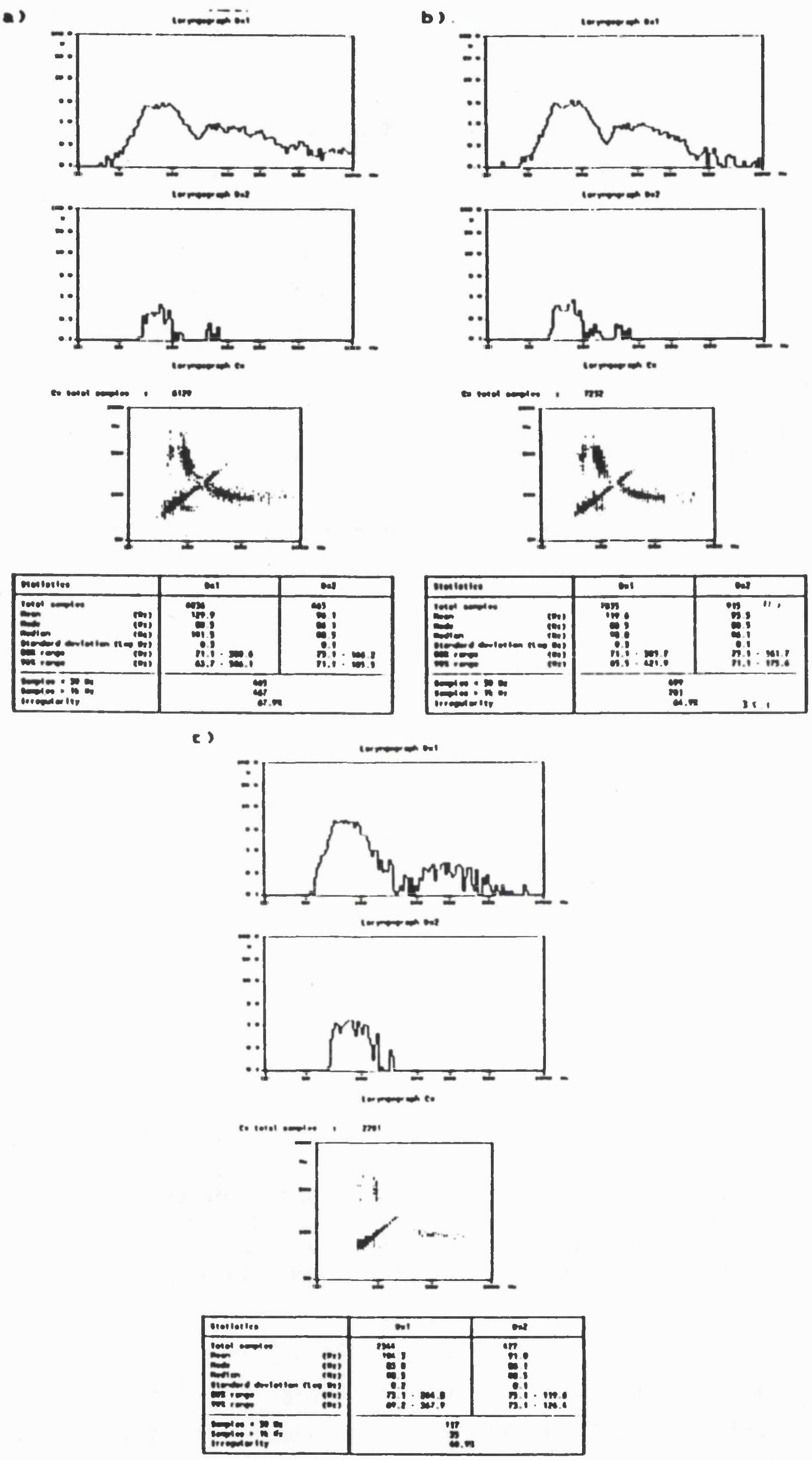
TABLE 24 B

PCLx 1 vs PCLx 2, SPEECH

		Normal Speakers	T1 subjects
%TS	d	0,2	0,5
	SD_{diff}	2,1	0,9
	L.A.	4,4 to -4,0	2,4 to -1,4
% Irregularity	d	-0,2	-1,2
	SD_{diff}	2,9	2,1
	L.A.	5,6 to -6,0	2,9 to -5,3

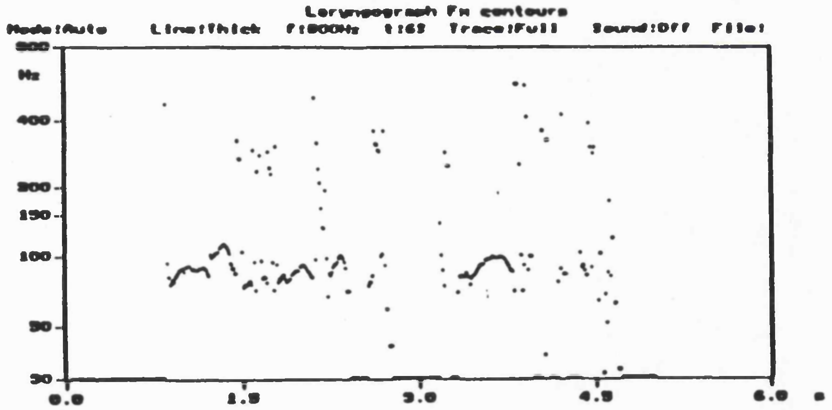
Repeatability of Regularity, %TS, and % Irregularity measurements on approximately the same recorded speech task using the PCLx system on two different occasions, 1 and 2. Mean differences (d), Standard deviation of these differences (SD_{diff}) and Limits of agreement (L.A.).

Some Normal speakers show a very small degree of Irregularity e.g. 4-6 %, whereas one, FW, shows very high degree of Irregularity, 40-60 % (Appendix 12, Table 40 and 48). The reason for this may have been the speaker's very thick neck, tendency to laryngeal movement during speech and a creaky voice. He produced a 'butterfly shaped' Cx plot (Fig. 81 a - c) illustrating alternating short and long period

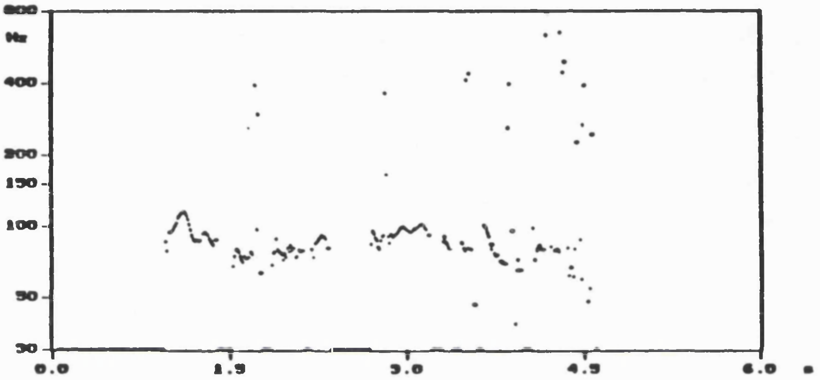


Normal speaker FW, Dx, Cx and statistics
a) PCLx 1 Speech b) PCLx 2 Speech c) PCLx 2 Reading

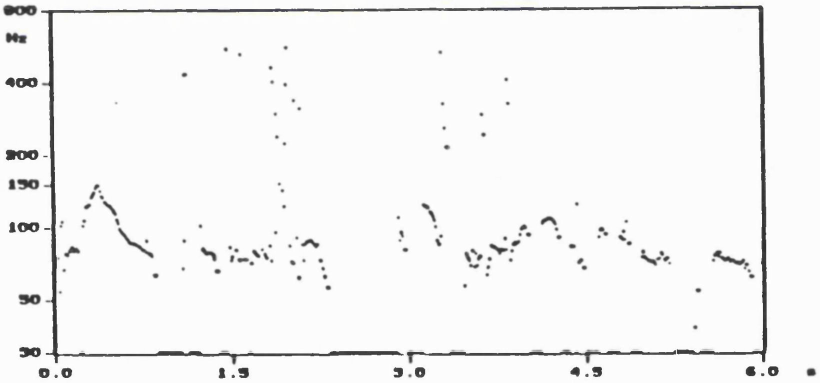
Figure 81



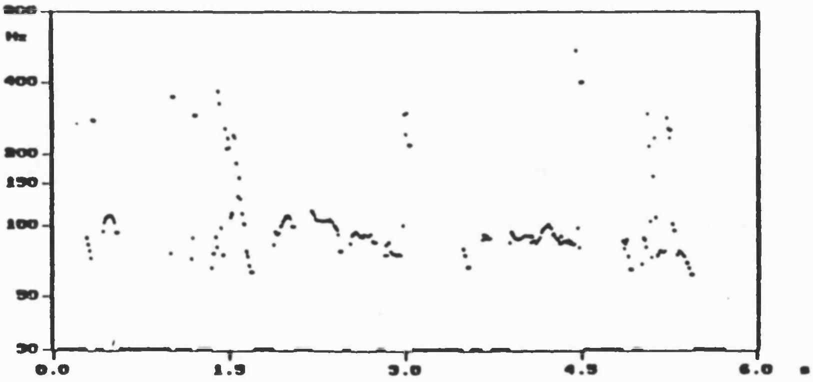
The rainbow is a division of white light into many beautiful colours.



There is, according to legend, a boiling pot of gold at one end.



When a man looks for something beyond his reach, his freinds say...



The Greeks used to say, it was a sign from the gods to foretell war
or heavy rain.

Normal speaker FW, Fx contours during reading of the Rainbow passage.

Fig 81 d

pulses during phonation at a high fundamental frequency. This pattern is obvious both in Speech and Reading aloud and may be due to a poor signal for the above mentioned reasons or the recurring pattern may also suggest that in this particular speaker there is synchronous vibration elsewhere in the region of the vocal folds which is picked up by the electrodes.

Figure 81 d illustrates the Fx traces produced by speaker FW during reading of some sentences from the Rainbow passage. Although the major part of the trace shows quite smooth Fx contours there is recurring high frequency 'noise' in the signal.

Despite the great amount of irregularity and distortion of Dx and Cx plots, speaker FW's Fx measurements give rise to remarkably small differences between analyses except for 1st order Mean (Appendix 12, Table 21, 29, 45, 53), 2nd order Mode (Appendix 12, Table 23, 30, 38) and %TS (Appendix 12, Table 24 and 32). This confirms the greater reliability of 2nd order Fx parameters, compared to 1st order, where such massive amounts of irregularity have been excluded.

x CONCLUSION.

The analysis of agreement between the old TPS system and the new PCLx system demonstrates that some extremely deviant voices distort the comparisons of fundamental frequency parameters, when such small groups (N=10) and few measurements are used. The changes in admittance of samples into second order in the new system, leads to greater differences between second order Fx values than were observed in the comparison of the TPS and the ROM systems. It may also be the case, that the new recording and analysis instrumentation is more sensitive than the old one. The mean differences (d) are however still within acceptable limits for the fundamental frequency parameters for the groups, and for 2nd order Means the standard deviation of the differences SD$_{diff}$, which indicate how well the analysis methods agree for an individual within the sample, are also within our stated acceptable limits of 5 Hz. 2nd order Means show good agreement between the systems (Table 18), whereas the Modes show such degrees of disagreement as to be difficult to compare (Table 19).

90% range minima show good agreement for large sample sizes i.e. our Speech samples (Table 20). Range Maxima do not seem to be sensitive to sample size (Table 20) and show just acceptable agreement between systems. As expected, %TS shows a systematic bias towards being greater in the PCLx analyses, albeit less so for T1 subjects than for Normal spakers and the tendency is reduced as the speech sample increases in size (Table 21).

The PCLx program shows good repeatability of measurements applied to the same voice recordings, where care has been taken to adjust the input gain before analysis to reduce extraneous 'noise' (Tables 22-24). Both the Means and the Modes show very good repeatability. Provided the same analysis program has been used, either can give a reliable estimate of central tendency in a voice sample. 2nd order 90% range Maximum is, not surprisingly, most sensitive to the analysed samples not being exactly the same, as in the Speech task, as it reflects any differences in fundamental frequency range used between samples.

The measures of % Irregularity and 'Regularity', %TS, show extremely good repeatability and different sample sizes seem only to have a very marginal effect. The Irregularity measure seems to be slightly less task and sample size sensitive than % TS. This is likely the result of % Irregularity being less sensitive to variation in fundamental frequency range than %TS and maybe reflecting more physiological bases for irregularity of vocal fold vibration and less the artefacts of the basis on which Tx samples are carried into 2nd order Dx distributions.

CHAPTER X - SIGNIFICANCE OF THE DIFFERENCES BETWEEN THE NORMAL SPEAKERS AND THE T1 SUBJECTS.

The randomized groups of ten T1 subjects and ten Normal speakers for the repeatability study, offered an opportunity to test the differences between their Fx and regularity measurements to find out whether these were significant.

The TPS program has been used for analysing the majority of the data in this study, but we chose to test both the reading and speech sample measurements from the TPS and from the PCLx repeatability study, as the latter analysis system is now in routine use. For comparison of the samples the average measurements from two occasions were used (Appendix 12, Tables 9-14 and 37-58). These are assumed to give the best estimate of any "true" value (Altman, 1990).

i Statistical Method.

The Mann-Whitney U-test for two independent samples was used to test the significance of the differences between the two groups' measurements. Our null hypothesis is that there is no significant difference between the measurements.

We choose to report the 95% confidence intervals (C.I.) for each parameter as recommended by Gardner and Altman (1989) as well as the level of significance of the differences between the Median values for the groups. They suggest that... *"It is more useful to present sample statistics as estimates of results that would be obtained if the total population were studied. The lack of precision of a sample statistic ...can be shown advantageously by a confidence interval."*

Our tables also contain a **'Point estimate'** of the differences between the group Medians. This is approximately the mid point of the confidence interval, and the **best estimate of the true difference between the two population Medians**. If the 95% confidence interval (C.I.) of the differences quoted in the tables below includes zero, this indicates that we cannot reject the null hypothesis.

The results of the Mann Whitney U test for the differences between the Medians of the two groups for each parameter using the TPS and the PCLx system are shown in Tables 25-30.

Table 25

TPS 1 vs TPS 2

READING

	Group	N	Median (Hz)	Point estimate	95 % Confidence Interval	Significance
1st order Mean (Hz)	Normal	10	117.2	14.00	-9.9 to 28.6	N.S. (p=0.4)
	Tl subj.	10	100.8			
2nd order Mean (Hz)	Normal	10	122.1	18.7	-6.7 to 32.0	N.S. (p=0.3)
	Tl subj.	10	100.1			
1st order Mode (Hz)	Normal	10	118.0	6.5	-14.9 to 34.7	N.S. (p=0.7)
	Tl subj.	10	96.8			
2nd order Mode (Hz)	Normal	10	123.9	17.7	-12.3 to 36.3	N.S. (p=0.4)
	Tl subj.	10	97.4			
2nd order 90% range Minimum	Normal	10	96.1	12.8	-5.1 to 27.5	N.S. (p=0.2)
	Tl subj.	10	81.0			
2nd order 90% range Maximum	Normal	10	161.9	23.1	-3.7 to 49.8	N.S. (p=0.06)
	Tl subj.	10	137.3			

Table showing the group Medians (Hz), the Point estimate of the differences between the population Medians, the 95% Confidence Interval and the significance level of the difference.

Table 26

PCLx 1 vs PCLx 2

READING

	Group	N	Median (Hz)	Point estimate	95% Confidence Interval	Signi-ficance
	Normal	10	121,3			
1st order				19,4	0,0 to	N,S,
Mean (Hz)					31,1	(p=0,054)
	T1 subj,	10	99,5			
	Normal	10	124,7			
2nd order				23,5	-2,5 to	N,S,
Mean (Hz)					33,8	(p=0,08)
	T1 subj,	10	98,8			
	Normal	10	115,7			
1st order				24,8	-5,3 to	N,S,
Mode (Hz)					37,9	(p=0,1)
	T1 subj,	10	89,8			
	Normal	10	121,4			
2nd order				23,5	-3,2 to	N,S,
Mode (Hz)					37,9	(p=0,09)
	T1 subj,	10	92,3			
	Normal	10	105,1			
2nd order				16,4	-6,3 to	N,S,
90% range					30,6	(p=0,2)
Minimum	T1 subj,	10	92,7			
	Normal	10	170,8			
2nd order				31,7	-0,00 to	Sign,
90% range					56,0	p < 0,04
Maximum	T1 subj,	10	138,2			

Table showing the group Medians (Hz), the Point estimate of the differences between the population Medians, the 95% Confidence Interval and the significance level of the difference,

ii Results.

Inspection of the tables above gives a clear indication of the T1 subjects' tendency to lower fundamental frequency measurements, and looking at tables 27 and 28 below, lower degree of voice regularity (%TS) and higher % Irregularity than the Normal speakers. This was also evident in figures 64-80 in the preceeding chapter.

Starting with the Reading measurements, all relating to the same voice samples in the TPS and the PCLx analysis, Tables 25 and 26 show consistently higher median 1st and 2nd order measurements for Normal speakers than for T1 subjects but no significant differences between Means or Modes (Appendix 12, Tables 9 and 12, 37 and 41). The slight increase in the Normal speakers' and decrease in T1 subjects' Means using the PCLx analysis (Table 26), brings the 1st order differences very close to significance at the 5 % level, however.

Tables 25 and 26 show the Minimum and Maximum 90% range measures in 2nd order for the Reading task (Appendix 12, Tables 10b and 13b, 39 and 43). We cannot reject the null hypothesis for either measure using TPS (Table 25). The Maxima are however close to significance using TPS and significant at the 4 % level in the PCLx analysis (Table 26).

The measures of Regularity (%TS) and Irregularity show highly significant differences between the groups (Appendix 12, Tables 11 and 23, 40 and 44). The median %TS measure comparing TPS and PCLx (Table 27 A and B) illustrates the effect of the changed admittance of Fx samples into 2nd order, with quite a marked increase of %TS for both Normal and T1 speakers in the Reading task using PCLx. This increase is much less marked for the larger Speech samples (Table 28).

The % Irregularity measure is only available on PCLx and shows slightly less irregularity in the T1 subjects' Reading samples (Table 27 B) than in their Speech (Table 28 B), whereas the Normal speakers show little difference between the tasks (Appendix 12, Tables 48 and 52).

Table 27

READING

A TPS1 vs TPS 2

	Group	N	Median	Point estimate	95 % Confidence Interval	Significance
% TS into 2nd order	Normal	10	31,1	9,5	2,2 to 20,8	Sign. p < 0,02
	T1 subj.	10	22,0			

B PCLx 1 vs PCLx 2

	Group	N	Median	Point estimate	95 % Confidence Interval	Significance
% TS into 2nd order	Normal	10	42,6	13,4	3,7 to 23,1	Sign. p < 0,009
	T1 subj.	10	30,6			
% Irregularity	Normal	10	13,5	- 9,95	-23,9 to - 0,61	Sign. p < 0,04
	T1 subj.	10	22,5			

Table showing the group Medians (Hz), the Point estimate of the differences between the population Medians, the 95% Confidence Interval and the significance level of the difference.

Table 28

SPEECH

	Group	N	Median	Point estimate	95 % Confidence Interval	Significance
A	<u>TPS1 vs TPS 2</u>					
% TS into 2nd order	Normal	9	36,9	12,0	2,9 to 20,8	Sign. p < 0,006
	T1 subj.	10	22,3			
B	<u>PCLX 1 vs PCLx 2</u>					
% TS into 2nd order	Normal	10	37,5	10,3	4,5 to 17,95	Sign. p < 0,01
	T1 subj.	10	29,5			
% Irregularity	Normal	10	13,7	-11,3	-22,6 to - 4,7	Sign. p < 0,006
	T1 subj.	10	25,5			

Table showing the group Medians (Hz), the Point estimate of the differences between the population Medians, the 95% Confidence Interval and the significance level of the difference.

Comparing the Speech samples (Tables 29 and 30 below) it is important to remember the likelihood of them reflecting slightly different portions of the speech recordings both between and within analysis methods. This said, there are no large differences either between or within methods among 1st or 2nd order Means (Appendix 12, Tables 45 and 49, 53 and 56). The tendency for T1 subjects' Means to be lower is obvious here as well, but the differences do not reach significance.

The 90% range Minimum (Table 29 and 30) shows quite small point estimates of the differences between the Medians and the measurements are remarkably similar for both analysis methods for both groups (Appendix 12, Tables 47 and 51, 54 and 57).

Table 29

TPS 1 vs TPS 2

SPEECH

	Group	N	Median (Hz)	Point estimate	95 % Confidence Interval	Significance
1st order Mean (Hz)	Normal	9	119.7	13.9	-3.2 to 29.8	N.S. (p=0.09)
	Tl subj.	10	105.8			
2nd order Mean (Hz)	Normal	9	124.7	18.2	-8.9 to 33.5	N.S. (p=0.1)
	Tl subj.	10	104.4			
1st order Mode (Hz)	Normal	9	107.8	12.0	-2.5 to 32.8	N.S. (p=0.09)
	Tl subj.	10	92.9			
2nd order 90% range Minimum	Normal	9	94.8	11.9	-10.6 to 27.5	N.S. (p=0.2)
	Tl subj.	10	81.0			
2nd order 90% range Maximum	Normal	9	180.5	27.1	2.5 to 64.2	Sign. $p < 0.02$
	Tl subj.	10	154.2			

Table showing the group Medians (Hz), the Point estimate of the differences between the population Medians, the 95% Confidence Interval and the significance level of the difference.

Finally, Table 30 shows the largest point estimate and a significant difference between the two groups of subjects in their 90% Maximum range measures using PCLx. This was also the case in the Reading task (Table 26) using the PCLx (Appendix 12, Tables 47 and 51, 54 and 57).

Table 30

PCLx 1 vs PCLx 2

SPEECH

	Group	N	Median (Hz)	Point estimate	95% Confidence Interval	Significance
1st order Mean (Hz)	Normal	10	120.6	18.2	-1.6 to 31.3	N.S. (p=0.06)
	Tl subj.	10	104.3			
2nd order Mean (Hz)	Normal	10	117.2	18.95	-1.3 to 35.7	N.S. (p=0.07)
	Tl subj.	10	102.2			
1st order Mode (Hz)	Normal	10	108.2	7.9	-17.0 to 28.8	N.S. (p=0.4)
	Tl subj.	10	98.9			
2nd order 90% range Minimum	Normal	10	94.8	12.5	-10.4 to 27.2	N.S. (p=0.1)
	Tl subj.	10	81.5			
2nd order 90% range Maximum	Normal	10	188.2	37.2	8.2 to 66.6	Sign. p < 0.02
	Tl subj.	10	147.0			

Table showing the group Medians (Hz), the Point estimate of the differences between the population Medians, the 95% Confidence Interval and the significance level of the difference.

iii Conclusion.

The above analyses of the differences between some fundamental frequency parameters and voice regularity measures obtained by analysing ELG data from two groups of subjects, indicates that 2nd order 90 % range Maximum, %TS and % Irregularity are most sensitive to changes in laryngeal tissue status, in this case as a result of irradiation, and might be useful indices of changed laryngeal dynamics.

The study by Lehman et al (1988) found significantly higher jitter values among a group of 20 men after radiotherapy for glottic carcinoma, compared to a group of normals. They used different measurement techniques and do not specify how jitter (i.e. frequency perturbation) was calculated. The measures were taken from sustained vowel phonation. As mentioned before, %TS is a gross measure of vocal fold vibrational regularity, and its relationship to jitter as described in the literature is a tentative one. However, it still must be affected by jitter in connected speech , as it is Fx based, and as shown here, it can be used to differentiate between normal and pathological speakers. Its advantage is that it is easy to calculate and reflects vocal fold vibratory regularity across longer speech and reading samples as opposed to less natural sustained vowel phonation. The same argument applies to the % Irregularity measure, which also differentiates between the two groups of speakers and is less sensitive to Fx range and sample size than %TS.

The likely cause of decreased regularity, %TS, i.e increased % Irregularity, of vocal fold vibration after irradiation, is mucosal dryness (Titze, 1980, Finkelhor, Titze and Durham, 1988), and the effect of radiotherapy on the vibrating portion of the vocal fold. Lehman et al (1988) found, under stroboscopy, reduced amplitude of vibration and reduced mucosal wave motion. They also noted relatively stiff vocal structures in the irradiated subjects.

The signifcant differences found between the groups in the Range Maxima in the PCLx analysis for both Reading and Speech, reflects a higher and wider range available to Normal speakers. Had it only been significant in the analysis of the Speech sample, one could have argued it was because the samples were not the same. The fact that the reading task also shows significant differences between the groups may indicate the explanation is not so simple, and the differences may be the result of physiological changes after radiotherapy as suggested above.

It is however important to consider, when looking at the 2nd order range maximum, that of all 2nd order measures, it must be the one most sensitive to the use of different intonation patterns occurring in different parts of a conversation or reading passage. Repeatability of this measure in Speech, especially for Normal speakers was poor (Table 22 and 23, Chapter IX). Comparison therefore, particularly of different speech samples, must be carried out with this in mind. Its effect, however, will always be reflected in the sample Mean value, here 2nd order Means, which, although showing consistently higher values for Normal speakers than for T1 subjects (Tables 25, 26, 29 and 30) do not reach significance.

Modal values, which do not reflect differences in fundamental frequency range, were found to be extremely sensitive to very deviant voice quality and to different gain settings in the repeatability study for the TPS system (Appendix 12, Tables 10 and 13). However, the same trend as for 1st and 2nd order Means is found, of Modes being lower in the irradiated group. The differences do not reach significance probably because of the very wide range of values within both groups.

The great similarity between analysis systems in the 90% range Minimum measurements, particularly in Speech (Table 29 and 30), despite the probability that slightly different speech samples have been used, may reflect a physiological 'lower limit' of vocal fold vibration: lower in T1 subjects than in Normal speakers. Barry et al (1990 b) also found the lower range limit less variable than the upper one. It was also less task dependent in their group of four normal speakers. With a large voice sample as we have in the Speech analysis (Appendix 12, Table 32 and 36), it seems a stable and repeatable measure irrespective of analysis method used. Range minima for Speech and Reading are very similar, although the trend is here as elsewhere of lower values among T1 subjects. 90% range Minima, however, show the smallest Point estimates of the differences between the populations of all the fundamental frequency comparisons.

The lack of significant differences achieved among the Means and Modes in 1st and 2nd order Fx distributions, despite the clear trends observed may be due to the small number of subjects used and no doubt to the great variability of values within both groups. This is easily observed in the Figures 64 to 80 in the previous chapter. The degree of imprecision of the Point estimates is reflected in the wide confidence intervals shown in the tables. If the groups were larger the precision of the Point estimates of the differences would improve and the width of the C.I. would be narrower.

The general trend of the differences observed throughout this study of lower fundamental frequency measures and greater irregularity of vocal fold vibration in irradiated speakers, does seem to suggest, that there are measurable differences between normal speakers' and irradiated speakers' voices. Here as in other studies however, there is no one objective fundamental frequency measure, that will on its own differentiate normal and abnormal voices.

Our findings seem to indicate that the PCLx system is slightly more sensitive and offers % Irregularity as an additional parameter, which reveals significant differences between normal and abnormal speakers and is less context and fundamental frequency bound than %TS.

iv Significance of differences between Speech and Reading samples.

In Chapter II (vi) we discuss stability of voice Fundamental frequency in normal speakers. Studies are referred to which show a tendency for Fx measurements to be lower in conversational speech compared to text readings (Saxman and Burk, 1968, Hollien and Jackson, 1973, Schulz-Conlon, 1975, Ramig and Ringel, 1983). Barry et al (1990 a) found significant differences in Mean and Modal Fx values derived from ELG between subjects recorded in a 'Free monologue' and reading the 'Environmental Passage'. Subjects were ten females and eight males aged 19-24.

In this study Fx data is collected of conversational speech and reading aloud. The majority of the data in the graphs Appendix 8 - 10 result from TPS analyses. Chapter IX, our study of agreement between TPS and PCLx derived measures, demonstrated that repeated 2nd order Means and %TS manage to fall within acceptable Limits of agreement and Mean differences. Data resulting from PCLx analyses is, however, indicated in the summary graphs of the total data collected.

Significant differences between Normal speakers and irradiated T1 subjects were found in %TS for both Speech and Reading (Tables 27 and 28) and for 2nd order 90 % maximum range measures in Speech (Table 29) using TPS. There were no significant differences found between Means or Modes between the groups.

The expected higher Mean Fx in Reading is not found using TPS (Tables 25 and 29) only in Normal speakers using the PCLx analysis system (Tables 26 and 30). Part of the reason for this may be that the statistical test used compares group Medians as indicated in the tables. The differences between Speech and Reading parameters within the two groups seem very small, however. A non-parametric test of the differences between Speech and Reading measurements within the group of T1 subjects and the group of Normal speakers was carried out using the data collected for the TPS - repeatability study (Appendix 12, Tables 9-14 and 53-58). The results are shown in Table 31 below:

Table 31

	Normal speakers			T1 subjects		
	N	P-value	95% C.I.	N	P-value	95% C.I.
1st Mean (Hz)	8	0.94	-13.8 to 14.9	10	0.76	-3.2 to 6.5
2nd Mean (Hz)	9	1.00	-10.7 to 11.3	10	0.31	-5.7 to 9.8
%TS	9	0.91	- 4.6 to 6.8	10	0.61	-4.0 to 1.8
90% Maximum range (Hz)	9	0.12	- 6.9 to 73.7	10	0.04 *	2.1 to 28.8
90% Minimum range (Hz)	9	0.29	- 8.8 to 2.5	10	1.00	-5.5 to 12.0

Result of Wilcoxon test for paired data, Speech vs Reading.

There is no evidence from this data to suggest that there are significant differences between Fx parameters derived from Speech and Reading. Particularly for Normal speakers there seems to be no difference between the tasks as far as 1st or 2nd order Means or %TS are concerned, nor between 90 % Minimum range measures for T1 subjects Their 90% maximum range, however, shows a significant difference between Speech and Reading, the Speech task showing higher maximum range Fx during conversational Speech than during Reading aloud (Table 25 and 29).

Because of the small numbers of subjects involved and the the wide range of values within the groups, the 95 % confidence intervals (C.I.) are wide.

On the basis of these findings, only measures from the Reading task will be used in the analyses, which follow comparing objective, perceptual and self-rated voice quality measures.

v **Significance of age differences between T1 and Normal subjects.**

In the first repeatability study it was found that the two groups of subjects differed significantly in age ($p < 0.03$) (Appendix 11). The T1 subjects' average age was 65.5.years with a Median age of 68.5

(Range 49-71). The Normal subjects' average age was 57.7 years with a Median age of 56.5 (Range 50-75).

Shipp and Hollien (1969) and Hollien and Shipp (1972) report a Mean speaking fundamental frequency (Mean SFF) of 118.4 Hz in a group of normal American male speakers with a mean age of 54.3 years (Range 50-59). Another group of males with the average age of 64.6 years (Range 60-69) had a Mean SFF of 112.2 Hz. The difference between the SFF values for the 50 and 60 year olds was not significant, however, (Hollien and Shipp, 1972). The task analysed was the first paragraph of the Rainbow Passage (Fairbanks,1960) the first two paragraphs of which were analysed here. They did find, however, that the tendency was for Mean SFF to increase in males with advancing age and the age group 70-79 showed a Mean SFF of 132 Hz. The difference between 60 and 70 year olds was significant. The increase in SFF is thought to be the result of senescent changes such as muscle atrophy, reduced vocal fold thickness and increasing stiffness of vocal fold tissues.

The age ranges in our two groups are similar but the random sample of T1 subjects are significantly older as a group. This may be expected to result in a tendency to higher Mean Fx values compared to the Normal group. This is not what is found however. Smoking, which has a known effect of lowering fundamental frequency (Abberton 1976, Sorensen and Horii, 1982, Murphy and Doyle, 1987) would not account for the differences as there were seven ex- and current smokers in the 'Normal' group of speakers in this study (Table 12, Chapter IX). It may be concluded that the differences noted between our two groups are not due to differences in age or smoking habits but may be put down to the effect of radiotherapy.

CHAPTER XI - PERCEPTUAL EVALUATION OF VOICE QUALITY.

i Inter- and Intra-rater agreement on Perceptual voice quality features of irradiated speakers.

As stated in the chapter on Perceptual evaluation of voice quality (Chapter III, ii), not all aspects of voice quality can be objectively tapped by e.g. Electrolaryngographic measurements of voice fundamental frequency and %TS. The voice source signal is filtered and modified in the vocal tract by its length and shape, the mucosal condition and habitual musculo-sceletal tension characteristics which vary from one individual to another.

To gain a more comprehensive picture of the voice quality of the irradiated subjects in this study, each subject's voice was therefore also judged according to Phonation Type and Laryngeal Tension characteristics using the Vocal Profile Analysis Scheme VPA (Laver et al, 1981), as described in Chapter III iv. The reason for selecting this limited number of parameters was economy of time, as a large number of voices were to be rated by two independent judges as well as by the experimenter.

An aim of this study was to examine the relationship between the chosen perceptual voice quality measures, irradiated speakers' self ratings and ELG based objective measurements. It was considered that the parameters chosen to be rated on the VPA, 'Harshness', 'Whisper', 'Creak' and 'Laryngeal Tension' were most directly related to laryngeal events which were objectively measured by ELG.

'Pharyngeal constriction' was also initially rated by the three judges, as it was considered to be relevant to our 'narrow' definition of voice quality. A tense and narrow oropharynx produces characteristic resonating qualities in the voice in that it dampens the source sound less than relaxed walls and a wide tract. Acoustically this results in a rise in the 1st formant and a lowering of 2nd formant frequencies (Fant, 1956). It also has the effect of narrowing formant bandwidths (Hardcastle, 1976).

The subjects were asked to fill in the Questionnaire and to rate their degree of 'Hoarseness' and functional limitations in voice usage, summarised in the 'Mean P-score' described in Chapter VIII, v.

The reason for asking two independent judges to rate the voices was to get some idea of the degree of inter- and, for the researcher, EC, intra- rater agreement on the parameters chosen. As this study was carried out over a very extended time span, and over 150 recordings were made of 40 subjects, EC had to be the main VPA judge. It was therefore important to find out to what extent her ratings agreed with other judges' and to what extent she was consistent within herself. Knowing the subjects, by sight and by sound, may have biased her perceptual ratings of their voices. It was therefore important to find out to what extent her ratings agreed with two independent judges, who did not have this knowledge.

Kreiman (1993) emphasises that for perceptual rating systems of voice quality to be meaningful, listeners must use the scales consistently (high intra-rater agreement), and for scales to be clinically useful raters must show high inter-rater agreement.

ii The making of the assessment tape, 20 random voice samples.

The researcher, EC, had been evaluating all the subjects' speaking voices on four of the parameters on the VPA, all except 'Pharyngeal Constriction', on each recording occasion (Appendix 5, Tables 1-3). She had been trained in the use of the VPA in 1986, and achieved satisfactory agreement with other raters at that time, but not had an opportunity to compare her ratings with other judges since. It was therefore felt that an interjudge agreement test needed to be run on the selected parameters to assess to what extent her ratings agreed with two independent judges, also trained on the VPA but more recently.

As described in the 'Pilot study' (Chapter VIII A) and 'Main study' (Chapter VIII, B) the original sample of irradiated subjects could not be described as random or neccessarily representative of the population. The selection criteria had not been stringent. We wanted to include as many subjects as possible irrespective of the time post radiotherapy, and to record them several times after the end of treatment if possible, to get a long term view of patterns of voice recovery after radiotherapy.

To make a tape representative of the wide range of voice quality represented among the irradiated T1 and T2 subjects, each recording occasion was assigned a letter suffix, if there were more than one for an individual subject. To select the recordings to be judged, a random draw was made of 15 reading samples from the total number of recordings. Five samples were selected at random from the group of 'Normal speakers' (Kramer, 1989) and copied in random order on the tape. The order of the 20 voice samples on the tape is shown in Figure 82 . Each reading was copied twice in succession, with a 4 second silent interval between samples.

Voice No	1	2	3	4	5	6	7	8	9	10
Speaker	FV	PBr	26	10	1	14	DTH	JSt	16	H
Voice No	11	12	13	14	15	16	17	18	19	20
Speaker	5	21	8	2	34	6	23	11	32	37

Order of the 20 voice samples on the test tape.
Letters denote 'Normal speaker', numbers refer to irradiated subjects.

FIG. 82

The reason for using a reading sample for the agreement test was an effort to 'standardise' the length and content of the voice samples to be evaluated on the VPA. However, as some of the speakers had been part of the Pilot study, some of them were reading 'The North Wind and the Sun' (Appendix 4) (Speakers No 5, 11, 13, 14, 16), the first two sentences of which were copied on to the assessment tape. The rest of the speakers were reading 'the Rainbow Passage' (Fairbanks, 1960) (Appendix 7), the second two sentences of which were copied.

The reason for this was that the word 'prism' appeared in the first sentence, on which most subjects hesitated.

Using two sentences of the reading passages, 45 and 32 words in length respectively, was an attempt at striking a balance between economy of time and neccessary length of sample to evaluate the parameters chosen. Laver et al (1981) suggest that *'Phonation type, audible in all phonetically voiced segments, can be judged over samples of only a few syllables, but settings which exert ther influence on a more limited number of susceptible segments.....will require much longer samples'*.

Two independent judges, AP and PR, Speech and Language Therapists, specialists in Voice and trained in VPA evaluation, were asked to rate the 20 randomly selected voices along five dimensions on the VPA; the degrees of 'Harshness', 'Whisper', 'Creak' , 'Laryngeal tension' and 'Pharyngeal constriction' (Figure 10, p.83). The researcher, EC also rated the voices on the same occasion. A separate 'practice tape' of five voices was used initially to help the judges to 'tune in' to the task in hand and to offer an opportunity to discuss and compare ratings before rating the test tape voices.

iii Statistical method.

The measure of agreement used is 'kappa' expressed as 'k'. This shows the agreement between judges in excess of the amount of agreement that could be expected by chance. If k=1 this indicates perfect agreement. k=0 indicates no agreement between judges. Altman (1991) suggests the following interpretation of k-values between 0 and 1 as there are no absolute definitions of 'k' (Fig. 83).

Value of 'k'	Strength of agreement
< 0.20	Poor
0.21-0.40	Fair
0.41-0.60	Moderate
0.61-0.80	Good
0.81-1.00	Very good

Interpretation of values for 'kappa' as proposed by Altman (1991).

Figure 83

Altman (1991) suggests, however, that *"In practise any value of 'k' much below 0.5 will indicate poor agreement, although the degree of acceptable agreement must depend on circumstances"*. They also suggest that inspection of the table of frequencies is essential as different frequency tables will yield similar values of 'k'. Our tables of frequencies are found in Appendix 13.

In the case of evaluation of voice quality parameters on the VPA Laver et al (1981) suggest that a difference between judges of +/- 1 scalar degree is acceptable, although the discrepancy, to be within acceptable limits, must fall either side and not include the 3 to 4 interval (see VPA form Figure 10, p. 83), which, as described earlier, differentiates between 'Normal' and 'Abnormal' degrees of a feature. We will not make any such distinction here, but use the scale as a 7-point equal appearing interval scale, including a '0' rating indicating 'absence' of a feature.

A problem arose in comparing agreement of ratings of 'Laryngeal tension'. Here there is an option for rating a voice either as laryngeally 'tense' or 'lax' (See VPA form, Figure 10, p.83). To resolve this, a convention was employed that any degree of rated 'laxity' was given a 'Laryngeal tension' rating of '0' in our frequency tables (Appendix 13, Tables 4, 9, 14). This was also the convention in all instances of ratings of 'Neutral' setting.

'Weighted kappa' recognises that "*observations near the diagonal* (in the frequency table), *representing a difference of only one* *category, are considered less serious than those where the discrepancy is two or three categories"* (Altman, 1991). Quadratic weighted kappa tends to be slightly higher than unweighted kappa described above, as the degree of disagreement is taken into account. This is the measure of agreement used here (Table 32).

iv **Results**

As in previous presentation of statistics in this study we follow Altman's (1991) advice and calculate the 95% Confidence Interval (C.I.= estimated kappa +/- 1.96 standard error (s.e.) of estimated kappa) for our quadratic weighted kappa and also the estimated standard error (e.s.e.), which gives a measure of the degree of imprecision of any given value of 'k' taking into account both the variation within the sample and the sample size. The interrater agreements between the three raters AP, PR and EC are found in Table 32 below.

TABLE 32

Table showing the degree of agreement 'k' greater than chance between three raters of twenty randomly selected voices along five voice quality dimensions on the VPA.

Voice parameter:		Harshness	Whisper	Creak	Pharyngeal Constrict.	Laryngeal Tension
Rater:						
AP vs PR	'k'	0,44	0,43	0,72	0,43	0,76
	e.s.e.	0,2	0,19	0,11	0,08	0,07
	95% C.I.	0,06 to 0,82	0,05 to 0,81	0,50 to 0,94	0,27 to 0,59	0,62 to 0,91
EC vs PR	'k'	0,24	0,37	0,37	0,26	0,37
	e.s.e.	0,13	0,10	0,18	0,12	0,10
	95% C.I.	-0,02 to 0,51	0,18 to 0,57	0,02 to 0,73	0,02 to 0,50	0,18 to 0,57
EC vs AP	'k'	0,46	0,71	0,32	0,11	0,42
	e.s.e.	0,15	0,12	0,24	0,05	0,12
	95% C.I.	0,17 to 0,75	0,46 to 0,95	-0,15 to 0,78	0,02 to 0,21	0,20 to 0,63

Inspection of the table reveals for the independent raters AP and PR, 'Moderate' agreement for the perceptual degrees of 'Harshness' (k=0.44), 'Whisper' (k=0.43) and 'Pharyngeal constriction' (k=0. 43). 'Good' agreement is achieved for 'Creak' (k=0.72) and 'Laryngeal tension' (k=0.76).

EC's agreement with PR is overall only 'Fair'. Her agreement is better with AP where the agreement is 'Moderate' for 'Harshness' (k=0.46) and Laryngeal Tension' (k=0.42) and 'Good' for 'Whisper' (k=0.71), 'Fair' and 'Poor' agreement is shown for 'Creak' (k=0.32) and Pharyngeal Constriction' (k=0.11) respectively.

Possible explanations for EC's rather poor agreement with the independent judges may be the long time since her training and lack of opportunity to test her ratings against others' in the meantime. All the speakers were also known to her, which may have biased her ratings. Inspection of frequency tables (Appendix 13, Tables 6-15) reveal a tendency for her to be less severe in her judgement of Pharyngeal Constriction and Laryngeal Tension and more severe on Harshness.

AP and PR had trained more recently. AP had a refresher course with a VPA tutor and AP and PR had carried out joint ratings on several occasions since. This is likely to account for their closer and more consistent agreement.

v <u>Interrater agreement on 37 irradiated voice samples.</u>

Having determined the interrater agreement on the twenty randomly selected voices including a few 'normal' ones, it was decided to repeat the exercise on reading samples for each of 31 subjects, for whom we had recorded reading samples, in the total sample of 35 irradiated T1 and T2 subjects in this study.

The first reading sample recorded for each subject 2 months or more after the end of radiotherapy, to allow for initial tissue recovery, was copied on to the assessemnt tape. To increase the sample size we also included, for subjects who had attended for a period of voice therapy, both the sample before therapy (a) and the one after the end of therapy (b) (Subjects 14 a,b, 16 a,b, 22 a,b, 25 a,b, 30 a,b). The 'b' samples were interspersed throughout the tape, which was copied as described previously with a 4 second gap between samples and each sample copied twice in succession. A total of 37 irradiated voice samples were recorded on the assessment tape.

A few reading samples had been erased so the next available recording was used (Subject 7, 8 and 21). A few subjects had not been recorded reading aloud due to severe difficulties with their voices (Subjects 13, 17 and 39). One subject (28) never brought his reading glasses so no record of his reading voice is available for assessment. The voice samples of two subjects, 14 and 30, were copied a third time after the end of irradiation and voice therapy (c), as they displayed severe deterioration in voice quality. Subject 14 as a result of late radiation changes, Subject 30, it is likely, because of persistent heavy smoking, vocal abuse and increasing deafness.

A table of the subjects, their ages, tumour stages and time post radiotherapy is shown in Fig. 84 The tape contained 37 recordings of 31 different subjects.

One VPA parameter was left out of the ratings of the 37 voices, namely 'Pharyngeal constriction'. The reason for this was the observation that although the independent raters AP and PR agreed quite well (Table 32), inspection of the frequency table (Appendix 13, Table 5) reveals that they rated this feature present in few of the voices and to a minimal degree for all except one. EC's agreement with both independent raters was extremely poor (Table 32).

Table 33 shows the 'Weighted kappas' resulting from the comparison of the three raters' judgements on the 37 irradiated voice samples. There is similar but improved agreement, compared to the ratings of the 20 voices, between the independent raters AP and PR of the 'Harshenss' (k=0.56), 'Whisper' (k=0.76) and 'Creak' (k=0.82) parameters and slight reduction in their agreement on 'Laryngeal

Voice No	1	2	3	4	5	6	7	8	9	10
Subject	2	3	4	5	6	7	8	10	11	12
Age	66	57	64	88	67	62	56	65	68	42
T-stage	T1	T2	T1	T1	T1	T1	T2	T1	T2	T2
MPRx	59	6	16	96	86	22	6	2	3	5

Voice No	11	12	13	14	15	16	17	18	19	20
Subject	14	16	20	21	22	23	24	14b	25	26
Age	63	58	57	39	48	61	65	64	52	59
T-stage	T1	T1	T1	T2	T1	T1	T2	T1	T1	T1
MPRx	63	61	4	94	5	3	11	65	9	2

Voice No	21	22	23	24	25	26	27	28	29	30
Subject	27	16b	29	30	31	32	25b	33	34	22b
Age	68	59	51	70	48	63	53	68	78	48
T-stage	T1	T1	T2	T1	T1	T1	T1	T2	T2	T1
MPRx	18	63	11	6	6	5	14	3	67	9

Voice No	31	32	33	34	35	36	37
Subject	35	36	37	30b	14c	38	30c
Age	64	65	55	70	70	72	71
T-stage	T2	T1	T1	T1	T1	T1	T1
MPRx	2	37	5	9	149	168	10

Figure showing 31 subjects' Ages, tumour stages and months post radiotherapy, in the order their voice recordings appeared on the tape.

Figure 84

Tension' (k=0.66) which remains 'Good' however. This is reflected in overall reduced e.s.e. and confidence intervals (Table 33).

TABLE 33

Table showing the degree of agreement 'k' greater than chance between three raters of 37 irradiated voices along four voice quality dimensions on the VPA.

Voice parameter:		Harshness	Whisper	Creak	Laryngeal Tension
Rater:					
AP vs PR	'k'	0,56	0,76	0,82	0,66
	e.s.e.	0,13	0,08	0,05	0,10
	95% C.I.	0,31 to 0,81	0,59 to 0,92	0,72 to 0,93	0,47 to 0,86
EC vs PR	'k'	0,43	0,52	0,52	0,35
	e.s.e.	0,12	0,09	0,16	0,08
	95% C.I.	0,20 to 0,66	0,35 to 0,70	0,21 to 0,82	0,20 to 0,50
EC vs AP	'k'	0,11	0,54	0,56	0,37
	e.s.e.	0,15	0,11	0,14	0,05
	95% C.I.	-0,18 to 0,39	0,33 to 0,74	0,29 to 0,82	0,27 to 0,48

EC now shows considerably improved agreement with PR, but dramatically reduced agreement on the 'Harshness' dimension with AP. It is difficult to speculate why this should be so, but maybe to do with the different sample of voices, these were all irradiated subjects well known to EC, many with quite high degrees of 'Harshness', 'Whisper', 'Creak' and 'Laryngeal tension'. AP was used to assess children with voice disorders, whereas EC and PR assessed mostly adult patients. This did not seem to make any difference in the first interrater agreement test, however, where EC and AP agreed as well as did AP and PR on 'Harshness' (Table 32).

The improved agreement between the judges may confirm Eskenazi, Childers and Hicks' (1990) finding that listeners found it easier to agree on degrees of abnormal voice quality parameters than aspects of 'normal' voice. Kreiman et al (1993) in their review of perceptual studies found that raters varied most widely in the middle of rating scales and agreed better at the 'Normal' and 'Abnormal' extremes. The 20 randomly selected voice samples referred to earlier contained five 'normal' voices. The irradiated sample was larger and representing mostly degrees of 'abnormal' voice quality.

The results confirm the need for thorough training and opportunities for refresher courses to maintain high degrees of agreement between users of perceptual voice evaluation systems as suggested by Hammarberg (1986).

Kreiman et al (1993) conclude, however: '*The present experimental results suggest that even highly experienced listeners frequently disagree completely about what they hear.*' They do, however, offer some hope for how to improve perceptual evaluation systems: '*..variability in voice quality ratings might be reduced by replacing listeners' idiosyncratic, unstable, internal standards with fixed external standards or 'reference voices' for different vocal qualities. A voice rating protocol using fixed reference voices would reduce listener related rating variability by providing all raters with a constant set of perceptual referents.*'

Laver et al (1981) provided a 'Graded reference tape' to be used in the training of VPA users to develop such perceptual referents. The voices represented, however, are reported to suffer from a number of problems that we have discussed earlier (Henton, 1983) e.g all speakers have Scottish accents, only 2/10 voices are female, ages of speakers are not provided. Not least would it seem daunting to have to develop listener agreement on the full range of rated Supralayngeal and other 'settings' provided for in the VPA protocol, considering the difficulties we and others have found in agreeing on just a few 'Phonation types' and Laryngeal tension settings.

vi Determining Intra-rater agreement.

Had rater agreement on voice quality parameters been a major aim of this study, we would have needed to run repeated ratings of the same voice samples by the three judges and also calculated intrarater agreement for AP and PR. Perceptual voice quality ratings were, however, supplementary to our objective measurements and subjects' self ratings and were carried out by EC for the total sample of voice recordings. It was therefore decided only to determine intra-rater agreement for EC, for the twenty randomly selected voices including some 'normal' ones, and for the 37 irradiated voice samples, particularly in view of the variable inter-rater agreement found.

Repeated ratings, denoted 'EC I' (the original rating used in the comparison with the independent judges) and 'EC II', took place with an interval of approximately six months and the same pre-recorded tapes were used. Intra-rater' agreement for EC is shown in Tables 34 and 35.

TABLE 34

Table showing intra-rater agreement 'k' greater than chance between repeated ratings by EC of 20 randomly selected voices along four voice quality dimensions on the VPA.

Voice parameter:		Harshness	Whisper	Creak	Laryngeal Tension
Rater:					
EC I vs EC II	'k'	0,70	0,52	0,50	0,58
	e.s.e.	0,10	0,12	0,16	0,10
	95% C.I.	0,50 to 0,91	0,28 to 0,75	0,20 to 0,81	0,38 to 0,78

EC seems at least to demonstrate some consistency in her own use of the VPA rating scales. 'Moderate' agreement is demonstrated in her repeated ratings of 'Whisper' (k=0.52) and 'Creak' (k=0.50). 'Good' agreement is shown for 'Laryngeal Tension' (k=0.58) and 'Harshness' (k=0.70). The agreement is illustrated in the frequency tables in Appendix 15 (Tables 1-4), where most of the paired ratings fall at least close to, if not on, the diagonal. There is, however, a discernable tendency to more severe ratings the second time.

As demonstrated in the inter-rater agreement earlier, the rating of more abnormal voice quality in the 37 irradiated voice samples resulted in improved agreement between raters. It seems to have had the same effect here. Table 35 shows EC's intra-rater agreement on repeated ratings of the 37 voice samples.

TABLE 35

Table showing intra-rater agreement 'k' greater than chance between repeated ratings by rater EC of 37 irradiated voice samples along four voice quality dimensions on the VPA

Voice parameter:		Harshness	Whisper	Creak	Laryngeal Tension
Rater:					
EC I vs EC II	'k'	0.69	0.70	0.86	0.76
	e.s.e.	0.09	0.09	0.04	0.08
	95% C.I.	0.51 to 0.87	0.51 to 0.88	0.78 to 0.95	0.60 to 0.91

'Whisper' (k=0.70), 'Creak' (k=0.86) and 'Laryngeal tension' (k=0.76) now show 'Good' or 'Very good' agreement, and 'Harshness' (k=0.69) virtually the same 'Good' agreement in the two groups. Estimated standard error (e.s.e) and Confidence intervals (C.I.) are smaller than in previous tables.

The frequency tables in Appendix 15 (Tables 5-8) show many more paired ratings on or close to the diagonal than in the repeated ratings of the 20 voices. There are instances of EC occasionally being more lenient on the second occasion and occasionally more severe. The tendency of improved agreement in rating more abnormal voices is again confirmed (Kreiman et al, 1993).

The researcher, EC, demonstrates a fairly consistent use of the VPA rating scales, even with a relatively long gap between ratings. The greatest consistency is shown in her rating of the irradiated voice samples, with which we are concerned in this study.

Kreiman et al (1993) mention three potential sources of variability in voice quality ratings:

1. Listeners' experience with voices shapes their internal standards for voice quality parameters being judged. Listeners have individual perceptual habits and biases and vary in how sensitive they are to the particular quality being judged. These factors would affect inter- rater agreement more than intra-rater agreement.

2. If the feature to be rated is poorly defined or 'lacks perceptual reality' this will affect listeners ability to rate it consistently.

3. There may be systematic interaction between the listeners and the task such as mismatches between listener sensitivity and scale resolution.

As factor analyses in other studies have demonstrated, voice quality features may be 'multidimensional' (Hammarberg, 1986, Sederholm et al, 1992). Kreiman, Gerratt and Berke (1992) suggest for instance a relationship between 'breathiness' and 'roughness'. They suggest that listeners may show different attention to the different aspects of such multidimensional features, which will provide another significant source of inter-rater disagreement.

It is likely that all these factors affected the agreement between the three judges in this study and it highlights the need for objective measures for evaluating voice quality.

In the case of irradiated subjects the intention was also to find out to what extent they considered themselves as being 'hoarse'. We asked them to rate their degree of 'hoarseness' and functional limitations in voice use, averaged and expressed as 'Mean P-score' (Chapter VIII,v). (Appendix 1b).

vii The relationship between acoustic, perceptual and subjects' self assessments of voice quality.

An aim of the study was to examine the relationship between ratings by trained judges of the voice quality features chosen from the VPA, irradiated subjects' self ratings and objective measures of voice fundamental frequency and regularity expressed as %TS, derived from ELG.

It was decided to relate perceptual ratings of an extract of the total reading passages to objective measures derived from the whole passage. The reading passages varied somewhat, as different passages had been used for the Pilot study and the Main study and the length of passages had been increased from one to three paragraphs of the Rainbow Passage (Fairbanks, 1960, Appendix 7) as the Main study progressed.

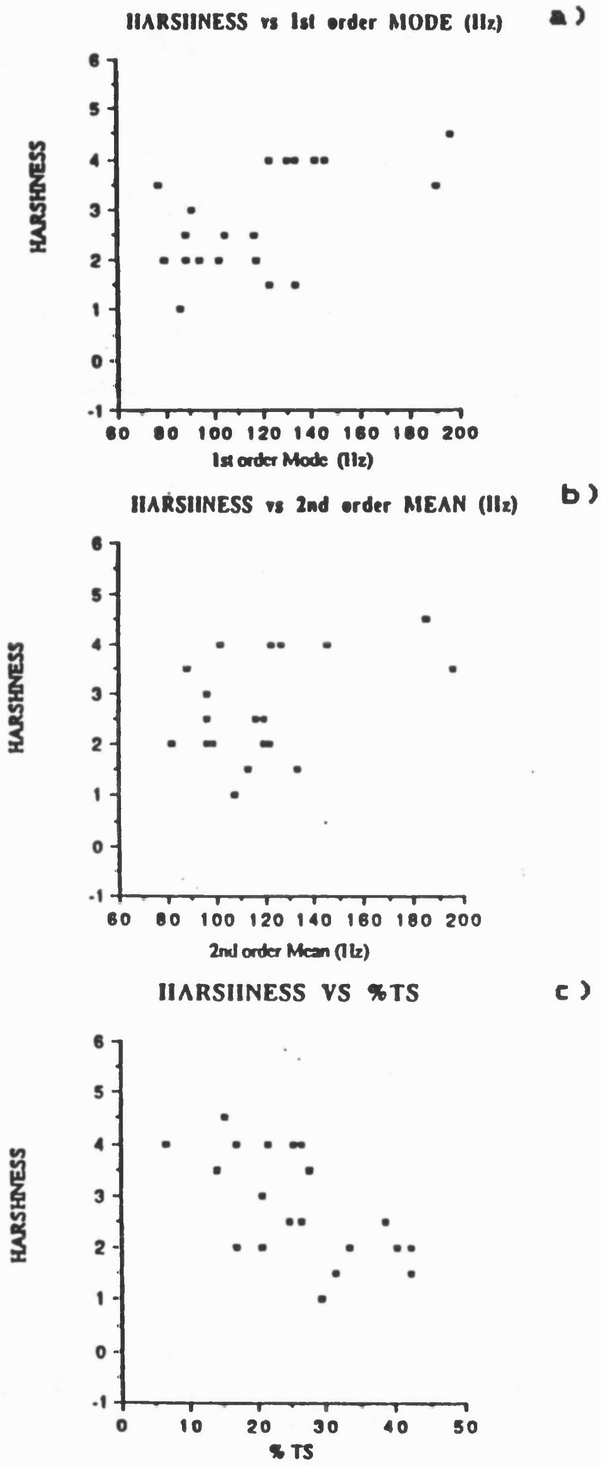
Laver and Hanson (1981) summarise evidence that suggests that 45-70 seconds of speech is necessary for the automatic abstraction of long term acoustic features by computer. Our objective measures were derived from the whole passage 'The North Wind and the Sun' (Appendix 4) and from between 1 and 3 paragraphs of the 'Rainbow Passage'. These passages were between 98 and 220 words long and took between 30 and 80 seconds to read for a fluent reader, longer for a less fluent one. According to Laver and Hanson (1981) this may be considered just adequate in length. Green, (1972), Hiller et al (1984) and Barry et al (1990 b) suggest however, that a two minute sample is needed.

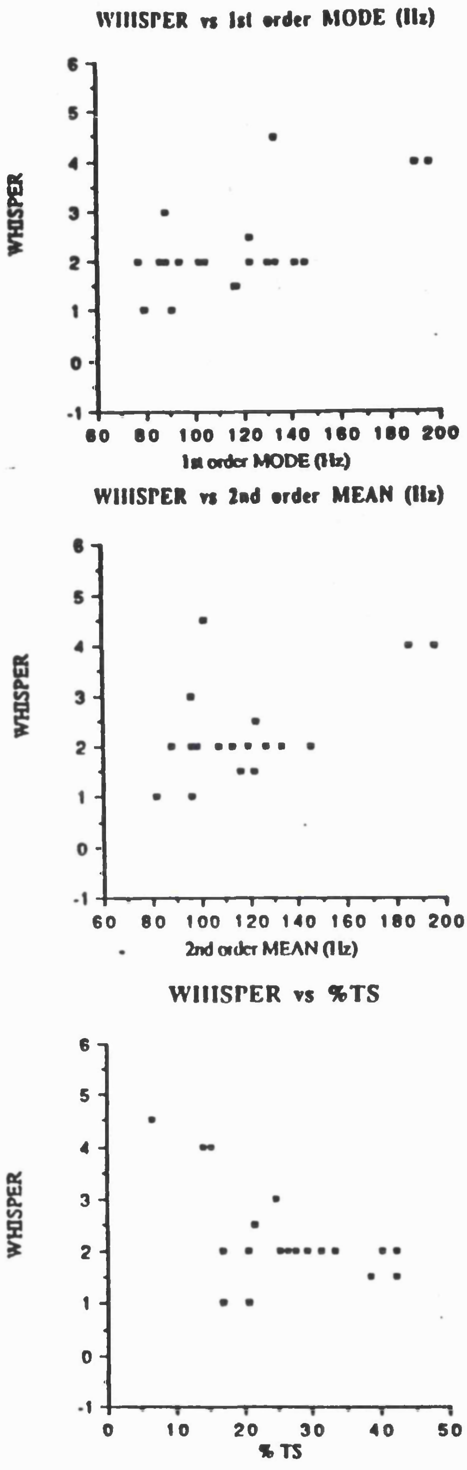
In our comparison of differences between speech and reading samples within our randomised groups of 'Normal' and T1 speakers (Chapter X), where the speech samples were at least 6000 Tx samples long, there were, however, no significant differences in frequency measurements within the groups between speech and reading although the latter samples were generally much shorter.

Plotting the average VPA ratings from the two independent judges of the twenty random voices described earlier (section iv) against some objective measures, suggested the following relationships (Fig. 85-88):

1. Tendency towards a negative correlation between 'Harshness' and regularity of vocal fold vibration expressed as %TS, i.e. the harsher the voice the less regular (Fig. 85 c).
2. Positive correlation between 'Harshness' and fundamental frequency measures, (2nd order Mean and 1st Mode) i.e. the harsher the voice the higher the fundamental frequency (Fig.85 a and b).
3. Negative correlation between 'Whisper' and regularity expressed as %TS (Fig.86 c).
4. Slight positive correlation between 'Whisper' and fundamental frequency measures (Fig.86 a and b).
5. No correlation between 'Creak' and %TS (Fig.87 c).
6. Slight negative correlation between 'Creak' and fundamental frequency i.e. the higher the measured fundamental frequency the less 'Creak' (Fig.87 a and b).
7. Negative correlation between %TS and Laryngeal tension, i.e. the more regular the voice the less the 'Laryngeal tension' was deemed to be (Fig.88 c).
8. Positive correlation between fundamental frequency and 'Laryngeal tension' (Fig.88 a and b).

As some of these trends seemed to make good sense in view of findings in other studies, we decided to formally test the correlations between our objective and perceptual measures. Rank order correlations between the average of the two independent raters' judgements of voice quality on the VPA, two fundamental frequency measures, 1st order Mode and 2nd order Mean and %TS from the reading passages in question are shown in Table 36. Baken (1987) suggests that the Modal fundamental frequency of the voice, as the frequency most used, will closely approximate an individual's 'habitual pitch'. This is confirmed by Barry et al (1990 b, 1991) who found modal values less task sensitive and more stable than the means. In our repeatability study, however, described in Chapter IX, the Mode was subject to much greater variation between recordings and analyses than the 2nd order Mean in our irradiated T1 subjects. This was the





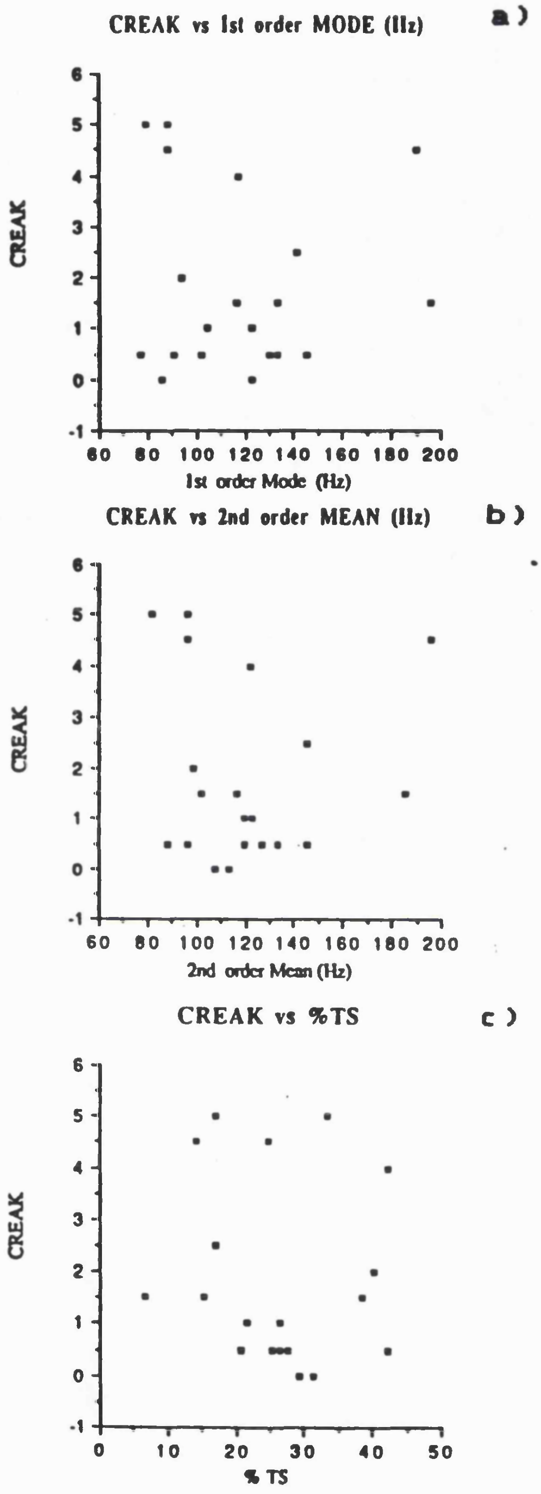
Plots showing the relationship between perceptual voice quality features and some objective measures derived from Lx.
a) 1st order Mode b) 2nd order Mean c) %TS

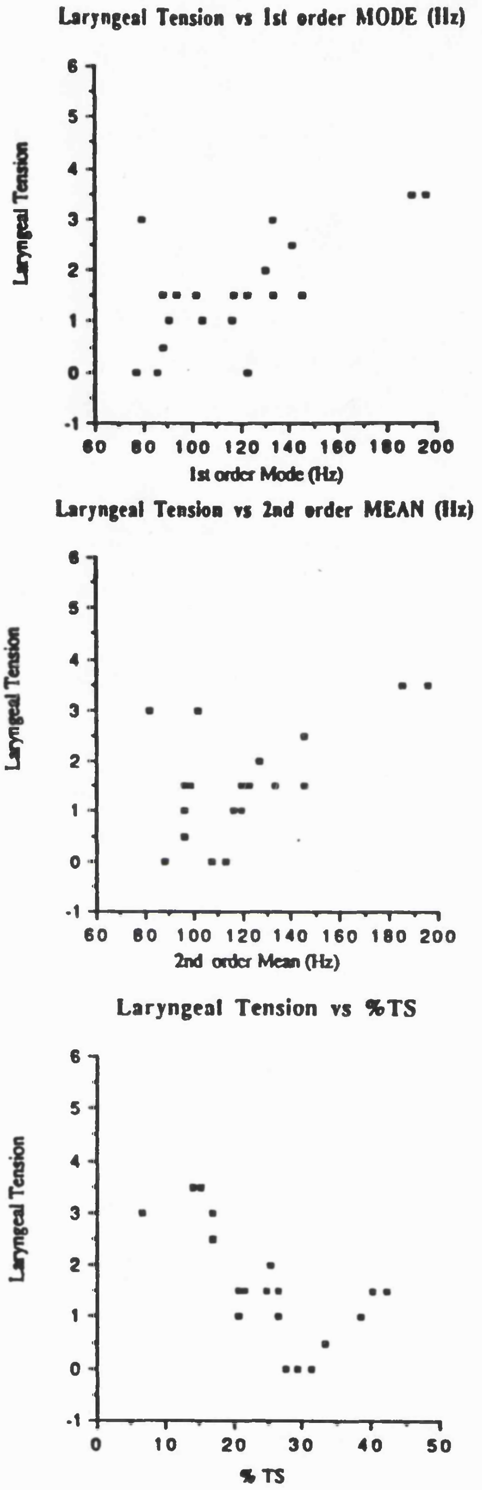
HARSHNESS

Figure 85

WHISPER

Figure 86





Plots showing the relationship between perceptual voice quality features and some objective measures derived from Lx.
a) 1st order Mode b) 2nd order Mean c) %TS

CREAK LARYNGEAL TENSION

Figure 87 Figure 88

reason why we chose to use the 2nd order Mean as descriptive of central tendency in our speech and reading samples.

%TS was chosen as indicative of vocal fold regularity of vibration and is related to pitch perturbation. In the repeatability and significance study it was found to be a very repeatable measure, which differentiated significantly between normal and irradiated speakers in both speech and reading.

Spearman rho rank order correlations were calculated, converted into z scores through Fisher's transformation (Gardner and Altman, 1989) and confidence intervals (C.I.) and levels of significance were calculated and are shown in Table 36.

TABLE 36

Correlations between perceptual and acoustic voice quality measurements in 20 randomly selected voices, 2 independent raters.

Acoustic measure:		1st order Mode (Hz)	2nd order Mean (Hz)	%TS
Perceptual Variable:				
HARSHNESS	r	0.54	0.41	-0.64
	C.I.	0.12 to 0.79	-0.04 to 0.72	-0.85 to -0.28
	p	0.01 *	0.08	0.002 *
WHISPER	r	0.47	0.36	-0.44
	C.I.	0.03 to 0.75	-0.10 to 0.69	-0.74 to -0.001
	p	0.04 *	0.1	0.05 *
CREAK	r	-0.03	-0.11	-0.20
	C.I.	-0.47 to 0.42	-0.53 to 0.35	-0.60 to 0.26
	p	0.9	0.7	0.4
LARYNGEAL TENSION	r	0.60	0.48	-0.65
	C.I.	0.21 to 0.82	0.05 to 0.76	-0.85 to -0.29
	p	0.005 *	0.03 *	0.002 *

Table showing the correlations between perceptual and objective measures for twenty randomly selected voices. r = Spearman rho correlation coefficient, C.I. = 95% confidence Interval p = level of significance. * indicates significance at the 5 % level.

Most of the correlations that were visibly apparent from the plots shown in Figs. 85-88 above were confirmed and found to be statistically significant. First order Mode seemed the more sensitive fundamental frequency measure, which correlated significantly with perceptions of 'Harshness' (r=0.54, p = 0.01), 'Whisper' (r=0.47, p = 0.04) and 'Laryngeal Tension' (r=0.60, p = 0.005). The latter also showed a significant correlation with 2nd order Mean (r=0.48, p = 0.03). This seems to confirm Laver's (1980) description of 'Tense voice' as being perceived as high pitched. Hirano (1981) found that 'strained' voice quality as rated on the GRBAS scale correlated with abnormally high fundamental frequency.

Our 'regularity' measure , %TS, proves a valuable objective measure of voice quality in that it shows highly significant negative correlations with 'Harshness' (r=-0.64, p = 0.002) and 'Laryngeal Tension' (r=-0.65, p = 0.02), and slightly less but still significant negative correlation with 'Whisper' (r=-0.44, p = 0.05) (Table 36, Fig. 85, 86 and 88 c).

'Harshness' is described by Laver (1980) as having greater amount of aperiodicity of vocal fold vibration than 'modal' voice (Wendahl, 1963, Zemlin, 1964, Michel, 1964). In Takahashi and Koike's (1975) study 'roughness' correlated with frequency perturbation measures. This was also the case in Hammarberg's (1986) study. 'Harshness' is described as produced with high degrees of laryngeal tension (Milisen, 1957, Kaplan, 1960, Zemlin, 1964, Laver, 1980), and so is 'Whisper' (Laver, 1980), which also seems to be reflected in these correlations.

In Laver's discussion of the difference between 'Breathy' and 'Whispery' voice quality, a distinction that has, however, been abandoned in the more recent version of the VPA, 'Whispery' voice would have less of a modal periodic component than a 'Breathy' voice, i.e. a 'Whispery' voice may have an aperiodic component which may be reflected in the negative correlation with %TS here. It may be the effect previously suggested both of increased laryngeal tension in 'hyperfunctional, whispery' voices and an artefact of how %TS is calculated as the percentage of adjacent Tx samples that fall in the same frequency 'bin' and are carried into the 2nd order Dx plot

(Chapter VII, iv). Occasionally a very 'Whispery' voice may be produced with erratic or missing vocal fold contact which would result in a correspondingly erratic Lx signal on which the Tx and %TS calculations would be based.

The reason why 'Creak' does not show any correlation with our objective variables (Fig. 87, Table 36) may be that this feature may only be present at certain points in the individual's speech, commonly at the end of utterances and may not have been perceived in most of the reading samples rated. The figure confirms, however, that voices with both high and low Modal and Mean fundamental frequency may be perceived as 'Creaky'.

The figure (87 c) illustrating the relationship between 'Creak' and regularity, %TS, shows the lack of an expected negative correlation between these two measures. This is surprising as 'Creak' is deemed to be the result of highly irregular glottal waveshapes, again, the reason for the lack of correlation may be that the rated short sample did not contain instances of 'Creak', whereas the total reading sample, on which the objective measures are based, did.

viii Correlations between Acoustic, Perceptual and Self-rated Voice quality features in 37 irradiated voice samples.

With confirmation of significant correlations between our perceptual and objective measures, the examination of the relationship between ratings and objective measurements was extended to the 37 samples of T1 and T2 radiotherapy subjects' voices described in the inter-rater agreement study described in section v. All these subjects had been asked to rate their own perception of degree of hoarseness and limitations in voice use (Mean P-score) (see Questionnaire Appendix 1 b). We could therefore also include their self ratings from the same occasion when the reading sample was recorded, in our correlations.

The average of three judges', (AP, PR and EC) ratings on the VPA were related to the other parameters. The average was used to reduce the effect of the variation between observers.

EC's ratings were included this time, despite her variable agreement with the independent raters, as her ratings were considered equally representative. In view of the lack of systematic retraining of judges on the VPA, current users are likely to show similar variability.

Laver (1980) proposed a relationship between degrees of 'Laryngeal tension', 'Harshness' and 'Whisper'. 'Harsh' voice would be produced with greater degree of 'Laryngeal Tension' than 'Whispery' voice. Table 37 below is based on the ratings of the three judges of the 37 irradiated voice samples and seems to confirm this relationship with a strong positive correlation (r=0.65, p = 0.001) between perceived 'Laryngeal Tension' and 'Harshness' but weaker, non-significant correlation between 'Whisper' and 'Laryngeal Tension'. There is, however, a strong positive correlation between 'Creak' and 'Laryngeal Tension' (r=0.47, p = 0.003) confirming Laver's suggestion that during 'creaky' voice production '...*the glottis is subjected to strong adductive tension and medial compression with vigorous ventricular involvement*.

TABLE 37

Correlations between voice quality parameters in the VPA as rated by three judges - 37 irradiated voices.

Perrceptual Variable:		WHISPER	CREAK	LARYNGEAL TENSION
HARSHNESS	r	0,27	0,32	0,65
	C.I.	-0,07 to 0,54	-0,004 to 0,59	0,41 to 0,80
	p	0,11	0,05	0,001**
WHISPER	r		-0,23	0,30
	C.I.	---	-0,52 to 0,10	-0,03 to 0,57
	p		0,16	0,07
CREAK	r			0,47
	C.I.	---	---	0,17 to 0,69
	p			0,003**

r = Spearman rho correlation coefficient, C.I. = 95% confidence Interval p = level of significance. * indicates significance at the 5 % level.

Table 38 reveals that some of the trends observed for the 20 voices rated by two independent judges remain and turn out to be highly significant for the group of irradiated subjects as well. The measure of voice regularity, %TS, still shows significant negative correlations with perceived degrees of 'Harshness' (r= -0.54, p = 0.0004), 'Whisper' (r= -0.45, p = 0.005) and 'Laryngeal Tension' (r= -0.54, p = 0.0004) (Table 38).

The next most important objective measure to influence judgements of perceived voice quality is the 1st order Modal fundamental frequency (Table 38). This measure correlates significantly with 'Harshness' (r= 0.40, p = 0.01) and 'Whisper' (r= 0.34, p = 0.01). 'Creak' interestingly, shows a negative correlation, not quite significant at the 5% level (r= -0.31, p = 0.06) with 1st order Mode. There is, however, a significant negative correlation with 2nd order Mean (r= -0.36, p = 0.03) (Table 38).

TABLE 38

Correlations between acoustic and perceptual voice qualitu ratings, in 37 irradiated voices,

Acoustic measure:		1st order Mode (Hz)	2nd order Mean (Hz)	%TS
Perceptual Variable				
HARSHNESS	r	0,40	0,23	-0,54
	C.I.	0,09 to 0,64	-0,10 to 0,52	-0,73 to-0,26
	p	0,01 *	0,2	0,0004 *
WHISPER	r	0,34	0,18	-0,45
	C.I.	0,02 to 0,60	-0,16 to 0,48	-0,68 to -0,15
	p	0,04 *	0,3	0,005 *
CREAK	r	-0,31	-0,36	-0,21
	C.I.	-0,57 to 0,02	-0,61 to - 0,04	-0,50 to 0,13
	p	0,06	0,03 *	0,2
LARYNGEAL TENSION	r	0,21	0,01	-0,54
	C.I.	-0,12 to 0,50	-0,31 to 0,34	-0,74 to -0,27
	p	0,2	0,9	0,0004 *

r = Spearman rho correlation coefficient, C.I. = 95% confidence Interval p = level of significance, * indicates significance at 5 % level,

Consequently, an increase in 1st order Mode (cf. 'habitual pitch', Baken 1987), seems to contribute to judgements of increased 'Harshness' and 'Whisper'. Judgements of an increase in 'Creaky' voice quality seem to be influenced by an overall reduction in frequency. This, however, only becomes obvious in 2nd order Mean measurements, which reflect the overall range of fundamental frequencies used. 'Creak' often only appears at the end of an utterance and not throughout a voice sample. 1st order Mode, as the frequency most commonly produced in the total voice sample is not influenced by extremes either end of the frequency range.

In our correlations calculated on the twenty randomly selected voices (Table 36), 'Creak' showed extremely low, negative, and non-significant correlations with our objective measures. 'Creaky' voice (Am. vocal fry) is closely related to overall low fundamental frequency measurements (Keidar, 1983), which we have found a tendency for in our irradiated subjects (Chapter IX). The greater number of subjects here and the fact that they were all irradiated patients, as opposed to a few normal speakers included among the twenty random voice samples, may account for the emergence of a significant negative correlation between 2nd order Mean fundamental frequency and degree of 'Creaky' voice quality.

High and signifcant correlations were found among the twenty randomly selected voices between 'Laryngeal Tension' and both 1st order Mode (r= 0.60, p = 0.005) and 2nd order Mean (r= 0.48, p = 0.03) (Table 36). In the correlations performed on the 37 irradiated voice samples this relationship is not found. The reason may be the converse of the appearance of a correlation in this group between 'Creak' and 2nd order Mean, a likely result of overall lower fundamental frequencies. The disappearance of correlations between 'Laryngeal Tension' and fundamental frequency may be to do with the lowering of fundamental frequency in this group. High degrees of Laryngeal Tension' are related to increase in fundamental frequency (Fig. 88 a and b)(Laver, 1980).

ix <u>The relationship between objective measures and Subjects' self-rated degrees of 'Hoarseness' and 'Mean P-score'.</u>

The measure of voice regularity, %TS, shows significant negative correlations with the irradiated subjects own perception of their degree of 'Hoarseness' (r= -0.43, p = 0.008) and Mean P-score' (r= -0.36, p = 0.03) (Table 39). All the correlations are negative. As the degree of voice 'regularity' increases, the lower the ratings of degrees of self-rated 'Hoarseness' and 'Mean P-score'.

Neither the self-ratings of 'Hoarseness' nor Mean P-score show significant correlations with either Modal or Mean fundamental frequency measurements. The subjects do not seem to consider a deviant pitch as indicative of abnormal voice quality, whereas it does seem to contribute to clinicans' judgements (Table 38).

TABLE 39

Correlations between acoustic measures and subjects' self ratings of voice quality - 37 irradiated voice samples.

Acoustic measure:		1st order Mode (Hz)	2nd order Mean (Hz)	%IS
Subjects' ratings:				
	r	0.07	-0.03	-0.43
HOARSENESS	C.I.	-0.26 to 0.38	-0.35 to 0.30	-0.66 to -0.12
	p	0.7	0.9	0.008**
MEAN	r	0.20	0.21	-0.36
P-SCORE	C.I.	-0.13 to 0.50	-0.12 to 0.50	-0.61 to -0.04
	p	0.2	0.2	0.03*

r = Spearman rho correlation coefficient, C.I. = 95% confidence Interval p = level of significance. * indicates significance at the 5 % level.

x The relationship between trained judges' voice quality ratings on the VPA and irradiated subjects' self ratings of voice quality.

Table 40 shows a positive correlation between clinicians'judgements of 'Whispery' voice quality and subjects rating of their perceived degree of 'hoarseness' (r= 0.49, p = 0.002) and of limitations in voice function, expressed as Mean P-score (r= 0.40, p = 0.01). Clinicians' ratings of degree of 'Laryngeal Tension' is also correlated with self ratings of degree of 'Hoarseness' (r= 0.37, p = 0.02).

TABLE 40

Correlations between clinicians' VPA ratings and subjects' self ratings of voice quality.

Self rated variable:		HOARSENESS	MEAN P-SCORE
Perceptual Variable			
	r	0.30	0.31
HARSHNESS	C.I.	-0.03 to 0.57	-0.02 to 0.57
	p	0.07	0.07
	r	0.49	0.40
WHISPER	C.I.	0.20 to 0.70	0.09 to 0.64
	p	0.002**	0.01*
	r	0.09	0.09
CREAK	C.I.	-0.29 to 0.40	-0.24 to 0.41
	p	0.6	0.3
	r	0.37	0.25
LARYNGEAL	C.I.	0.06 to 0.62	-0.09 to 0.53
TENSION	p	0.02*	0.1

r = Spearman rho correlation coefficient, C.I. = 95% confidence Interval, p = level of significance, * indicates significance 5 % level.

The positive relationship between perceived 'Laryngeal Tension' and irradiated subjects' rating of their degree of 'Hoarseness' may confirm Lehmann et al's (1988) findings of voices produced with greater than normal effort post radiotherapy. They bore this out with objective measures showing increased subglottal pressure used for voice production after radiotherapy. Laver (1980) suggests a 'Whispery' voice is produced with a greater degree of laryngeal effort than a 'Breathy' voice, and that glottal friction will be more prominent in a 'Whispery' voice due to greater 'Medial Compression'.

xi CONCLUSION.

Acoustic information derived from the speed and regularity of vocal fold vibration measured by ELG contributes to trained listeners' judgements of voice quality and laryngeal tension characteristics. Irradiated subjects' ratings of their degrees of hoarseness and limitations in vocal function are related to trained listeners' perception of degrees of 'Whisper' and 'Laryngeal Tension'.

The best objective measure that correlates with both sets of perceptual parameters, trained listeners' and irradiated subjects', is the measure of voice regularity, %TS. All the correlations are negative. As %TS increases, the ratings of degrees of perceived 'Harshness', 'Whisper', 'Laryngeal Tension', 'Hoarseness' and 'Mean P-score' decrease (Table 38 and 39)

xii SUMMARY OF FINDINGS - RANDOMISED GROUP STUDY.

The agreement between and repeatability of objective acoustic measures derived from ELG, using different analysis programs and instrumentation for recording purposes have been determined and found satisfactory in most instances. TPS was selected for use rather than Lx ROM, due to its more precise Fx calculations and more detailed statistical analysis.

TPS and PCLx also satisfy our levels of agreement and PCLx repeatability except for %TS, which is more affected by the different admittance in the PCLx system of Tx samples into 2nd order distributions, for Normal speakers than for the irradiated subjects. The additional % Irregularity measure offered by PCLx is less task and sample size dependent than %TS.

2nd order parameters showed slightly better agreement and repeatability than 1st order particularly using TPS with T1 subjects.

Significant differences were found between Normal speakers and irradiated subjects in % TS, % Irregularity and 2nd order 90 % maximum range measurements.

There were no significant differences between Speech and Reading fundamental frequency measurements, except among T1 subjects' 90 % maximum range measures.

We found a significant age difference between T1 subjects' and Normal speakers'. The former were significantly older and would have been expected to show higher Fx measurements. This does not emerge.

Instead, they demonstrate consistently lower Fx Means, Modes and Range measures compared to the younger Normal speakers.

In spite of variable, sometimes poor, inter-rater agreement between the experimenter, EC, and two independent judges on the VPA, EC's intra-rater agreement shows good consistency even with a six month interval between ratings.

Averaging the three raters judgements along four VPA dimensions and correlating them with irradiated subjects' objective voice quality measures and their self-ratings of voice quality and function, significant correlations emerged:

Laryngeal Tension and Harshness	(r=0.65, p=0.001)	(Table 37).
- " - Creak	(r=0.47, p=0.003)	(-"-)

Harshness and 1st order Mode	(r=0.40, p=0.01)	(Table 38)
- " - % TS	(r=-0.54, p=0.0004)	(-"-)
Whisper and 1st order Mode	(r=0.34, p=0.04)	(-"-)
- " - %TS	(r=-0.45, p=0.005)	(-"-)
Creak and 2nd order Mean	(r=-0.36, p=0.03)	(-"-)
Laryngeal Tension and %TS	(r=-0.54, p=0.0004)	(-"-)

Hoarseness and %TS	(r=-0.43, p=0.008)	(Table 39)
Mean P-score and %TS	(r=-0.36, p=0.03)	(-"-)

%TS seems to be the most useful objective measure of voice quality. Regularity, or lack of it, of vocal fold vibration is the most powerful parameter which influences both trained and untrained judges' ratings of voice quality.

CHAPTER XII - DESCRIPTION OF TRENDS IN THE TOTAL COLLECTED DATA.

1 Tumour stage and fractionation - Effect on measurements.

Two questions that the Main Study is trying to answer concern to what extent the size and spread of the vocal fold tumour affects post radiotherapy voice quality and the difference, if any, of treatment in 3 fractions as opposed to 5 fractions per week.

Appendix 2 A-C lists all the subjects according to tumour stage, fractionation, tumour dose, field size, histology and site of lesion. Despite some missing data due to inadequate information in patients' notes, and some notes being unavailable, T1 and T2 subjects seem to have received, on the whole, similar tumour doses. Only two of the latter, subjects 12 and 29, received more than 6000 cGy in split courses. The T2 subjects' seem to have had wider radiation fields, but there are still some T1 subjects who have had field sizes of up to 6.5 cm^2 eg. subjects 14, 16, 20 and 25.

As the 3F/week fractionation regime had been in use for much longer than the 5F/week, only a few 5F/week subjects were recorded at a maximum time post radiotherapy of 63 months. This is used as the Maximum follow - up time in this comparison of measurements and ratings for three groups; T1 subjects treated in 3F and 5F/week (Appendix 2A) and T2 subjects including 8 treated in 3 fractions and 2 treated in 5 fractions per week (Appendix 2 B). We exclude subject 35 (T2 treated in 5F/week) from these calculations as he was the only non-native speaker of English in the sample and spoke with a heavy Italian accent which may affect both his fundamental frequency measures and the perceptual voice evaluation. It was not meaningful to divide the T2 subjects according to fractionation because of the small numbers involved. The minimum time limit is set to 2MPRx, to allow for initial tissue recovery post radiotherapy.

We chose to take into account all assessment occasions of each subject within the stated time limit of 2-63 MPRx (Appendix 8, 9 and 10, A and B). The comparison does not take into account the time factor, except that assessments were caried out within the stated time span. To avoid individual subjects, who had been assessed on more occasions than others, unduly influencing the group values, the averages are calculated of each subjects' fundamental frequency and %TS measurements over the number of assessment occasions (Appendix 5, Table 1-3).

Table 41 shows the average of these averages of objective measurements, %TS and 2nd order Fx Mean for each group's conversational speech. (MPRx = Months post radiotherapy)

TABLE 41

Average Speech %TS and 2nd order Mean (Hz)
All T1 and T2 subjects 2 - 63 MPRx.

	T1 subjects		T2 subjects
	3 F/week	5 F/week	3F and 5F/week
$N_{subj.}$	9	10	10
$N_{occ.}$*	28	48	35
Median MPRx	12,5	17	11
Mean Age	59,1	60,5	57,3
Age Range	48-67	48-73	39-77
Average % TS	25,1	23,2	18,5
Average 2nd order Mean (Hz)	108,4	101,3	95,2
Smokers	1	4	3

* Number of recording occasions average is based on

There is a slight reduction in the % TS i.e. regularity of vocal fold vibration as we move from the T1 subjects treated in 3F/week, to those treated in 5F/week to T2 subjects. The Mean fundamental frequency for the speech task also seems to show a systematically decreasing trend from one group to the next, confirming the low average pitch used by speakers after radiotherapy.

The T2 subjects seem to speak with an overall lower pitch and more irregular vocal fold vibration than the T1 subjects. In our significance study, comparing Normal speakers and a group of T1 subjects (Chapter X) there was a significant difference between their voice regularity measures. Normal speakers' Median %TS was 36.9 % and T1 subjects' 22.3 % for Speech (Table 28, p. 306). There was a nonsignificant difference in their Fx Means, however, which showed Normal speakers' Mean of 124.7 Hz and T1 subjects' 104.4 Hz (Table 29, p. 307). Non significant differences in fundamental frequency were also found by Lehman et al (1988) and is likely due to the wide range of measurements among both 'Normal' and irradiated speakers.

Although the number of subjects who are smoking is very small in each group (Table 41), it seems as if subjects treated in 3F/week have heeded advice to stop more than other subjects but two out of the three smokers in the T2 group were treated in 3F/week. Non - smoking could conceivably influence the better voice quality as measured by the group Fx and %TS values in T1 subjects treated in 3F/week. Smoking has been found to lower fundamental frequency in speakers (Gilbert and Weismer, 1974, Abberton, 1976, Stoicheff, 1981, Sorensen and Horii, 1982, Comins, 1988) and to give rise to significantly poorer voice quality after radiotherapy (Benninger, Gillen, Thieme et al, 1994).

Kramer (1989) found a trend among non-smokers to have more regular vocal fold vibration as measured by %TS than either smokers or ex-smokers. It did not reach significance, however, due to small group numbers. Comins, also using ELG measurements, found significantly greater variance in fundamental frequency measures with increasing measures of exhaled Carbon Monoxide in her sample of 'Old' (45-60) Smokers compared to non-smokers. It was thought to be due to more years spent smoking.

EC's VPA evaluation of voice quality for the same groups of subjects, for the same task, Speech, and on the same occasions are shown in Table 42. Individual VPA ratings for subjects on each of these occasions are shown in Appendix 5, Tables 1-3. However, for comparison purposes of such a large number of perceptual ratings, we show the proportion of judgements within 'Normal' scalar limits 1-3, 'Abnormal' Limits 4-6 and occasions where there were 'None' perceived' of each of three 'Phonation types' and 'Laryngeal tension' (Laver et al, 1980).

There seems to be an overall tendency for T2 subjects to be perceived as having more severe degrees of harshness, whisper, creak and laryngeal tension than T1 subjects. Among T1 subjects those treated in 3F/week seem on the whole to show marginally less severe degrees of harshness and laryngeal tension than those treated in 5F/week. This would be contrary to assumptions that 3F/week would give rise to more severe late tissue reactions than 5F/week. Nor was this assumption confirmed in the BIR study reported in Chapter VI (Table 6, p. 182). Smoking may be a factor as it is found to have a detrimental effect on irradiated tissue recovery and more T1 5F/week subjects smoked (Table 41) (Rugg et al, 1990, Whittet et al, 1991, Benninger et al, 1994).

That T2 subjects show less good voice quality is no surprise as their original lesions were more extensive (Benninger et al, 1994) and, as pointed out above, they tend on the whole to have had larger radiation field sizes which has been found to result in more persistent oedema post radiotherapy (Inoue et al, 1992) (Appendix 2 B). The group includes three subjects whose tumours have subsequently recurred; subject 3, who has undergone a hemilaryngectomy, subject 24, who had a total laryngectomy and subject 29, who is refusing further treatment. Both the latter subjects had tumours that extended to the posterior commissure and into the subglottis with resulting worse prognosis. McIlwain (1991) suggests the posterior glottis should be considered as part of the subglottis for tumour staging purposes as it is so intimately related to the subglottis. Both subject 24 and 29 were treated in 5F/week. Among T1 subjects there have been no recurrences to date.

TABLE 42
Vocal Profile Analysis
All T1 and T2 subjects 2 - 63 months after the end of radiotherapy.

	T1 subjects		T2 subjects
	3 Fractions/week	5 Fractions/week	3F and 5F/week
$N_{subj.}$	9	11	10
$N_{occ.}$	30	52	40
Median MPRx	10	16	11
VPA ratings: HARSHNESS			
1-3	46,7%	17,3%	19,4%
None	26,7%	32,7%	16,7%
4-6	26,7%	50,0%	63,9%
WHISPER			
1-3	46,7%	38,5%	27,8%
None	3,3%	3,8%	2,8%
4-6	50,0%	57,7%	69,4%
CREAK			
1-3	46,7%	15,4%	11,1%
None	6,7%	30,8%	30,6%
4-6	46,7%	53,8%	58,3%
LARYNGEAL TENSION			
1-3	58,3%	34,7%	17,1%
None	6 lax or variable	3 lax or variable	1 lax
4-6	41,7%	65,3%	82,8%

Table showing proportion of perceptual voice quality ratings within'Normal' (1-3), 'None perceived' and 'Abnormal' (4-6) limits on the VPA for T1 and T2 subjects.

The smaller number of subjects (N_{subj}) and occasions (N_{occ}) in table 41 than in Table 42 is the result of a few subjects' ELG recordings being of such poor quality as to not be analysable i.e. early recordings of T1 subjects 28 and 39 in the 5F/week group, subject 13 in the T2 group and two early erased recordings of T1 subject 16 in the 3F/week group. This also has a small effect on the Median time post radiotherapy (MPRx). The average ages are similar in the groups but there is a tendency for the T2 subjects to be somewhat younger than the T1 subjects.

To avoid any one subject's VPA ratings unduly influencing the group values due to him having been assessed on a greater number of occasions, which may be the case in Table 42, the Median scalar degrees for each subject were calculated for the occasions represented in Tables 41 and 42 (Appendix 2, Table 1-3). Table 43 shows the Medians of the Medians for the groups:

TABLE 43
Vocal Profile Analysis
All T1 and T2 subjects 2 - 63 months after the end of radiotherapy.
Median scalar degrees 1-6

	T1 subjects		T2 subjects
	3 Fractions/week	5 Fractions/week	3F and 5F/week
N_{subj}.	9	11	10
N_{occ}.	30	52	36
Median MPRx	10	16	11
Harshness	3,0	3,5	4,0
Whisper	3,3	4,3	4,0
Creak	3,5	3,8	4,0
Laryngeal Tension	3,0	4,0	4,8

Although we cannot test if these differences are significant, the trend of best voice quality perceived in the group of T1 subjects treated in 3 fractions per week and worst voice quality in the group of subjects treated for T2 tumours remains.

In our correlations of perceptual and objective measures of voice quality (Chapter XI) there are significant negative correlations between % TS and perceived degrees of 'Harshness' ($r = -0.54$, $p < 0.0004$), 'Whisper' ($r = -0.45$, $p < 0.005$) and 'Laryngeal tension ($r = -0.54$, $p < 0.0004$). 2nd order Mean is found to correlate significantly with the degree of perceived 'Creak' ($r = -0.36$, $p < 0.03$) (Table 38).

In Table 41, 42 and 43 both objective voice quality measures and VPA ratings seem to agree that T1 subjects treated in 3F/week have less severe degrees of Harshness, Whisper and Laryngeal Tension, higher degree of voice regularity (%TS) and speak with slightly higher pitch than either T1 subjects treated in 5F/week or T2 subjects. Patients with more extensive T2 tumours have worse voice quality as measured by fundamental frequncy, % TS and perceived degrees of Harshness, Whisper and Laryngeal Tension than those with T1 tumours.

The degree of creaky voice quality (Table 42) also seems to show a clear distinction between the groups and a tendency for T1 subjects treated with 3F/week to show a higher proportion of ratings within 'Normal' scalar degrees 1-3 than the other two groups. On the other hand the other groups show a higher degree of 'None' perceived. The groups' 'abnormal' 4-6, ratings show an increasing trend from one group to another. The generally low pitched voices of speakers in all the groups seems to give rise to the perception of creaky voice quality, supporting Keidar's (1983) finding that creaky voice is mainly perceived as a function of fundamental frequency and our finding of a negative correlation between perception of creaky voice and 2nd order Mean fundamental frequency (Table 38, p. 337), i.e the higher the Mean Fx the less 'Creak' is perceived.

'Lax or variable' tension features (see VPA form Fig. 10, p. 83) are included among the 'None Perceived'. It is a feature of a few subjects' voice quality, often as a response to advice to reduce vocal effort. It is observed with a decreasing trend from T1 3F/wk to the other two groups.

Table 44 shows the subjects' own ratings of voice quality and vocal limitations on most of the occasions represented in tables 41, 42 and 43. Individual ratings have been averaged over a number of occasions as described above. The ratings in the table are averages of these average ratings.

Table 44

Subjects' self ratings of voice quality
All T1 and T2 subjects 2 - 63 months after the end of radiotherapy.
Averages of average ratings 1-7

	T1 subjects		T2 subjects
	3 Fractions/week	5 Fractions/week	3F and 5F/week
$N_{subj.}$	9	11	10
$N_{occ.}$	20	47	29
Median MPRx	14,5	17,5	8,0
Hoarseness?	2,2	3,3	3,4
Voice use now?	4,7	4,4	3,9
Voice a problem?	2,0	2,2	2,3
*Voice tires?	3,4	3,2	3,6
*Difficulty singing?	4,5	3,8	3,7 (N=9)
*Difficulty shouting?	2,6	3,2	3,9
*Difficulty speaking over noise?	3,1 (N=8)	3,5	3,6
*Difficulty using the telephone?	2,3	1,8	2,8
Dry throat?	4,3 (N=2)	3,5	3,1 (N=3)
Contributes to problem?	4,3 (N=2)	2,9	4,0 (N=3)

* item included in Mean P-score

The table reveals a tendency for T2 subjects to have the highest degree of hoarseness, the voice being a problem, voice tiring after a lot of use and difficulty shouting and using the telephone. This is not surprising as they have had more extensive vocal fold lesions in the first place. Nor are the small differences between the groups surprising, given the gross seven point rating scale used and reported average ratings. The only major difference is in subjects' reported degree of hoarseness and difficulty shouting, which is least in the T1 subjects treated in 3F/week and greatest in T2 subjects.

The study of correlations between objective and perceptual measurements reported earlier (Chapter XI), shows some significant correlations between irradiated subjects' objective and self rated measures (Table 39). Negative correlations were found between %TS, i.e regularity of vocal fold vibration and self rated 'Hoarseness' (r= -0.43, p < 0.008) and also with Mean P-score, which is an average of an individual's rating of particular voice function questions 6-10 (Appendix 1b)(r=-0.36, p < 0.03). These questions, whose ratings are averaged and expressed as Mean P-core, are indicated in Table 44 by an asterisk. The correlations were performed on objective measurements from the first reading sample of each irradiated T1 and T2 subject at least 2MPRx and their replies to the questionnaire on the same occasion.

The gradual increase in ratings of Hoarseness from one group to the next and the slightly higher ratings on questions 6-10 (all included in Mean P-score) seems to correspond to the same tendency observed in Table 41 of decreasing %TS, or regularity of vocal fold vibration from one group to the next. There is, however, a surprisingly small difference between the groups in their perception of their vocal limitations as a problem (Table 44). T1 subjects treated in 3F/week seem to have greater problem with singing than both the other groups and T2 subjects seem to rate their daily voice usage slightly less than the other groups.

ii <u>The effect of continued employment on measured and rated voice quality.</u>

Riska and Lauerma (1966) found a tendency among patients who went back to work after radiotherapy for laryngeal carcinoma to be less satisfied with their voices than those who did not.

Table 45

Table showing averages of average individual Speech %TS and 2nd order Mean
All T1 and T2 subjects, WORKERS AND RETIRED/UNEMPLOYED, 2 - 63 MPRx.

	T1 subjects 3 Fractions/week	5 Fractions/week	T2 subjects 3F and 5F/week	All Subjects
WORKERS	$N_{subj}=5$ $N_{occ}=18$	$N_{subj}=6$ $N_{occ}=27$	$N_{subj}=4$ $N_{occ}=12$	$N_{subj}=15$ $N_{occ}=57$
Mean age	56,0	54,8	56,2	55,5
Smokers	0	2	1	3
%TS	26,7	22,7	19,6	23,2
2nd order Mean (Hz)	100,3	96,7	97,9	98,3
RETIRED AND UNEMPLOYED	$N_{subj}=4$ $N_{occ}=10$	$N_{subj}=4$ $N_{occ}=21$	$N_{subj}=6$ $N_{occ}=23$	$N_{subj}=14$ $N_{occ}=54$
Mean age	64,7	67,9	57,9	63,0
Smokers	1	2	2	5
%TS	23,1	23,9	17,8	21,0
2nd order Mean (Hz)	118,5	111,6	93,3	105,7

Table 45 illustrates the average objective measurements taken from all individual Speech recordings also featured in Table 41 but here grouped according to whether subjects are still working or not. The T2 subjects still stand out with the lowest degree of regularity of vocal fold vibration, %TS; and among workers T1 subjects treated in 3F/week have the highest degree of regularity. The difference in %TS has disappeared in the retired T1 groups but the fundamental frequency measurements have increased. The retired T1 subjects are the oldest and the higher fundamental frequency could be due to this (Mysak, 1959, Hollien and Shipp, 1972, Pegoraro-Krook, 1986). However, one subject in each T1 'retired' group speaks with an exceptionally high pitch, Subject 23, (Mean 153.3 Hz, T1 3F/week) and subject 30 (Mean 136.41 Hz, T1 5F/week), (See also Appendix 8 A and 10 A) and distort the values in Table 45 where so few subjects (N=4) are included in each group. Both are also smoking again.

Our T2 group contains most of our youngest subjects (Appendix, 2 B), which is reflected in the lower average age in this group. Table 45 also reveals that the T2 group also includes a greater proportion of retired and long term unemployed subjects (6/10). This may explain their ratings of lower daily voice usage than the T1 groups and their relatively low reported degrees of voice problems (Table 44) compared to what objective and VPA ratings might suggest. It may also reflect a general attitude towards their laryngeal malignancy, possibly contributing to why they had more extensive tumours when first diagnosed.

There are greater differences in objective measures between the different fractionations and tumor stages, as found in the previous section (i), than between workers and non-workers.

Table 46 shows averages of average self ratings of all subjects who were working in each group. Table 47 refers to the subjects who were retired or unemployed. Table 46 confirms the higher ratings of voice usage among workers, but only a marginally higher degree of hoarseness and limitations in vocal function than among the non-workers. The lowest hoarseness ratings are among subjects, working and not working, treated in 3F/week. The greater daily voice usage among workers seems, not surprisingly, to lead to greater experience of the voice tiring after a lot of use. Shouting, speaking over noise and using the telephone is also rated slightly more difficult among workers. The differences are small, however, between the latter two functions. There is only a marginally higher rating of voice as a 'Problem' among workers (Table 46 and 47).

Table 46

Working Subjects' average self ratings (1-7) of voice quality.
All T1 and T2 subjects 2 - 63 months after the end of radiotherapy.

	T1 subjects		T2 subjects	
	3 Fractions/week	5 Fractions/week	3F and 5F/week	All workers
$N_{subj.}$	5	6	4	15
$N_{occ.}$	12	26	10	48
Hoarseness?	2,5	3,5	3,2	3,1
Voice use now?	5,1	5,3	4,3	4,9
Voice a problem?	2,2	2,1	2,6	2,3
*Voice tires?	4,2	3,2	3,8	3,7
*Difficulty singing?	4,4	3,8	2,4	3,7
*Difficulty shouting?	3,7	3,4	4,0	3,7
*Difficulty speaking over noise?	3,5	3,6	3,5	3,5
*Difficulty using the telephone?	2,4	1,9	3,0	2,4
Dry throat?	4,3 (N=2)	3,7 (N=6)	1,0 (N=1)	2,1 (N=9)
Contributes to problem?	4,3 (N=2)	3,2 (N=6)	1,0 (N=1)	1,9 (N=9)

* Item included in Mean P-score

Riska and Lauerma's (1966) finding of workers considering their voices more of a problem than non-workers does not seem to be confirmed here. This is despite the deterioration in voice function as a result of irradiation being slightly more acutely felt by those who use their voices more, confirmed by our working subjects' self ratings of greater daily voice use compared to those who are retired (Rating of '4' indicating 'Normal use', see Questionnaire , Appendix 1 b).

Table 47

Retired or Unemployed Subjects' average self ratings (1-7) of voice quality.
All T1 and T2 subjects 2 - 63 MPRx.

	T1 subjects		T2 subjects	
	3 Fractions/week	5 Fractions/week	3F and 5F/week	All retired
$N_{subj.}$	4	5	6	15
$N_{occ.}$	8	21	19	48
Hoarseness?	1,9	3,0	3,6	3,0
Voice use now?	4,8	3,3	3,7	3,8
Voice a problem?	1,8	2,4	2,0	2,1
*Voice tires?	2,5	2,7	3,5	3,1
*Difficulty singing?	4,6	4,0	4,2	4,2
*Difficulty shouting?	1,3	3,0	3,9	2,9
*Difficulty speaking over noise?	3,6	3,3	3,6	3,3
*Difficulty using the telephone?	2,2	1,7	2,6	2,2
Dry throat?	-	3,2 (N=5)	4,2 (N=2)	3,5 (N=7)
Contributes to problem?	-	2,5 (N=5)	5,1 (N=2)	3,2 (N=7)

* item included in Mean P-score

'Ability to sing or 'hum a tune' seems to be more of a problem among the retired and unemployed (Table 47). Several subjects in all the groups did not bother to rate their singing ability on some occasions, or commented they had never been able to sing. In retrospect this question could have been left out of the final questionnaire, as it does not only reflect difficulties due to the effects of irradiation but no doubt in some subjects also to a pre-morbid reluctance or inability to sing. However, in the T2 group, subject 12 had been singing semi professionally, and subject 33 had been a keen amateur tenor.

The rating of degree of throat dryness, and to what extent this is experienced as contributing to the voice problem, was added to the revised questionnaire (Appendix 1 b) after comments by some Pilot study subjects that they experienced this as a major problem. This has therefore only been rated by a few T1 and T2 subjects treated in 3F/week, but by all subjects treated in 5F/week. Due to this imbalance in the data, the differences are impossible to interpret at this stage.

iii <u>'Pilot subjects revisited' and Main study subjects - The effect on voice quality measures as a function of time after</u> radiotherapy.

The graphs in Appendix 8,9 and 10 A and B illustrate developments or stabilisation in individual subjects' objective measurements over time. Among T1 subjects treated in 3F/week (Appendix 8 A and B) there is a tendency towards a lowering and then stabilisation in 2nd order Fx Mean. %TS seems to gradually increase during the first year after radiotherapy. Among T1 subjects treated in 5F/week and among T2 subjects (Appendix 9 and 10 A and B) there does not seem to be much increase in %TS but a tendency to a decrease in many subjects and on the whole the Fx measures seem to stay slightly lower than in the T1/3F/week group. These observations were confirmed in Table 41 (previous section ii) where measurements representing averages of all individuals' %TS and Fx values, plotted in the graphs, showed a tendency towards a lower Fx and %TS as we moved from T1/3F/week to T1/5F/week to T2/3 and 5F/week.

To find out whether a medium to long interval of several years between perceptual and self-ratings, reveals differences in voice quality and function irrespective of tumour stage, two groups are examined:

The first group, referred to as 'Pilot subjects revisited', consists of eleven T1 (N=8) and T2 (N=3) subjects (8, 12 and 21), all treated in 3 fractions per week, who had been recorded for the Pilot study on one or more occasions and then been reassessed for the Main study several years later. Details of the times of each subjects' first recording for the Pilot study, two months or more after radiotherapy (1st occ.), and their first recording for the Main study (2nd occ.), expressed in MPRx (Months post radiotherapy) and their ages on these occasions, are found in Table 48. Smoking habits and occupations for this group of subjects are shown in Appendix 16, Table 1. Six of the eleven subjects were working on the first assessment occasion but only two on the first recording occasion for the Main study, which took place after an interval of between four and seven years.

TABLE 48

Assessment occasions, T-stage and ages of 'Pilot subjects revisited'.

		1st occasion		2nd occasion	
Subject No	T-stage	Age	MPRx	Age	MPRx
2	T1	66	53	68	112
4	T1	64	16	71	107
7	T1	63	16	70	100
8	T2	56	5	62	71
10	T1	65	2	70	61
12	T2	42	2	48	62
14	T1	64	61	71	149
16	T1	54	2	61	61
20	T1	57	2	65	92
21	T2	39	5	43	55

The Pilot subjects 'revisited' were compared to a group of subjects referred to the Main Study where a follow up occasion at an interval of at least two years had been recorded. Intervals in this group varied between 26 and 57 months. Details of assessment occasions and subjects' ages are found in Table 49; their occupations and smoking habits on the last assessment occasion are found in Appendix 16, Table 2. The group consisted of 12 subjects three of whom were treated for T2 tumours; all except two subjects (22 and 33) in this group were treated in 5 fractions per week. Six were working on the first occasion and four on the second occasion. Two were still smoking on reassessment.

TABLE 49

Assessment occasions, T-stage and ages of Main study subjects.

Subject No	T-stage	1st occasion Age	1st occasion MPRx	2nd occasion Age	2nd occasion MPRx
22	T1	49	5	53	57
24	T2	65	11	68	44
25	T1	52	6	56	50
26	T1	60	4	64	61
27	T1	68	18	72	63
28	T1	58	3	63	58
29	T2	50	2	53	42
30	T1	70	6	73	40
31	T1	47	4	51	37
32	T1	63	5	66	48
33	T2	68	3	70	31
36	T1	65	37	67	63

iv Self-ratings.

Inspection of Tables 50 and 51 showing average self ratings, 1-7 (See Questionnaire, Appendix 1 a, b) of vocal symptoms and limitations in voice function, reveals very similar degrees of 'Hoarseness', daily 'Voice use' and experience of the voice posing a 'Problem' in the two groups on the first assessment occasion (Median MPRx = 5 for both groups). Pilot subjects seem to experience slightly more limitations in vocal function summarised in Mean P-score on this occasion compared to the Main study subjects (Table 50 and 51).

On the second assessment occasion ('First occasion Main study'), in the case of the Pilot subjects between four and seven years later (Table 46), there is a slight reduction (i.e. improvement) in most of their ratings (Table 50, Fig. 89). Two subjects did not return their questionnaires on this occasion, subject 2, who had severe problems with his voice, and subject 10, whose voice had returned to almost normal. The reduction in 'Voice use now' and experienced functional limitations expressed in Mean P-score (Table 50) may also reflect the fact that only 2/9 subjects were still working, subjects 16 and 21.

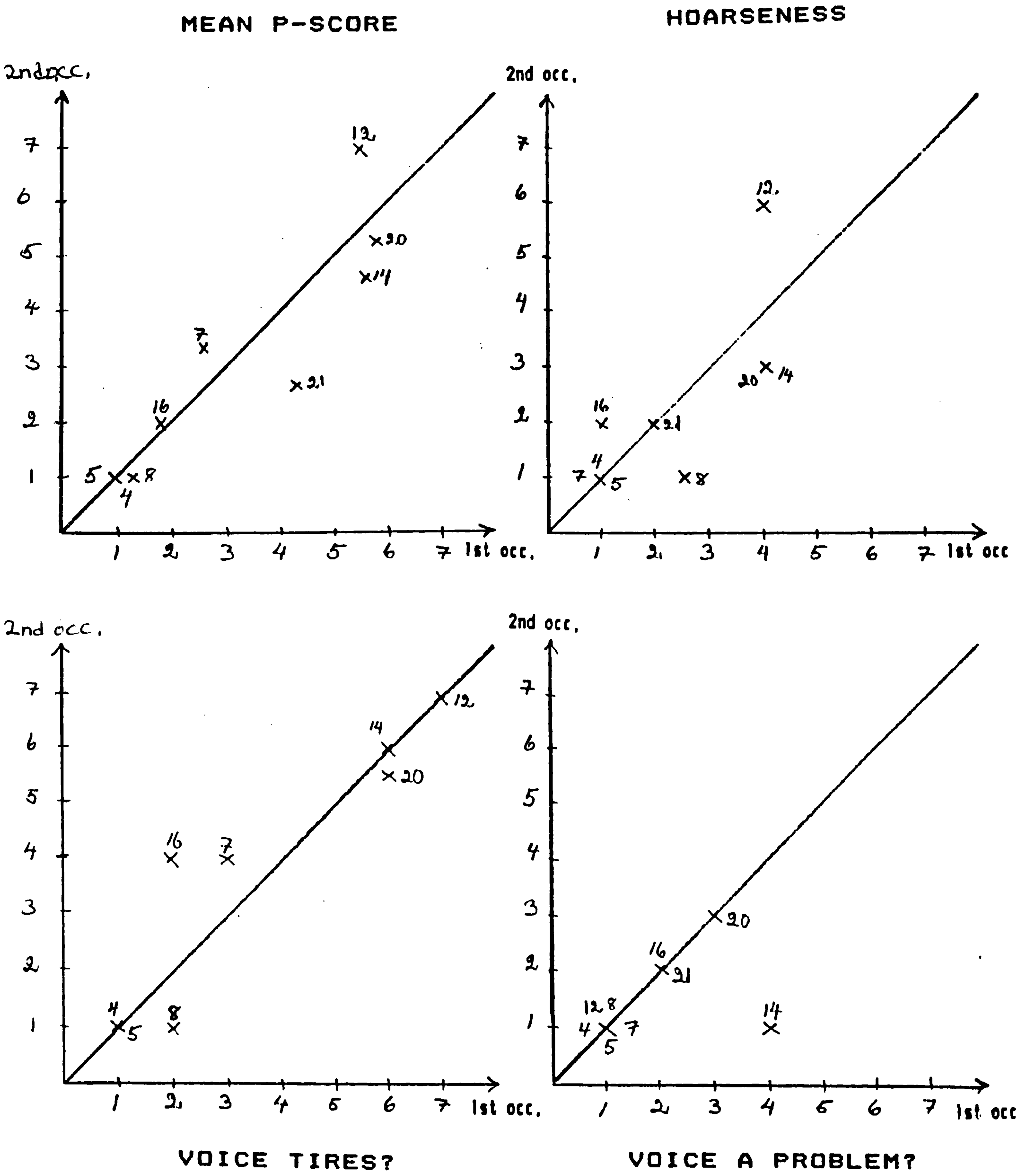
Figure 89, showing individuals' ratings on the first occasion against ratings on the second occasion, of a few parameters from the questionnaire, seems to illustrate how subjects who originally reported few if any difficulties do so also at this late stage, for instance subjects 4, 5, 7 and 8. Subjects who reported significant difficulties early on keep rating their symptoms and limitations relatively higher, for instance subjects 12, 14 and 20. Subject 12, described in some detail in Chapter VIII,ix, was thought to be using ventricular band phonation and subject 14, who will be described in more detail later, had persistent and increasing difficulties due to late radiation fibrosis.

TABLE 50

AVERAGE SELFRATINGS ON THE QUESTIONNAIRE
'PILOT SUBJECTS REVISITED'

	First occasion Pilot Study (2MPRx+)	First occasion Main Study
AGE		
Range (yrs)	39-66	43-88
Median (yrs)	63	68
MPRX		
Range	2-53	55-149
Median	5	92
N_{subj}.	11	9
Working	6	2
Hoarseness?	2,4	2,2
Voice use now?	4,1	3,6
Voice a problem?	1,9	1,4
*Voice tires?	3,6	3,7
*Difficulty singing?	4,1 (N=10)	3,5 (N=8)
*Difficulty shouting?	3,9 (N=10)	2,8 (N=7)
*Difficulty speaking over noise?	3,7 (N=9)	2,3 (N=8)
*Difficulty using the telephone?	2,4 (N=10)	2,1
Mean P-score	3,5	3,1

* Item included in Mean P-score. Where the averages are based on fewer than the original number of subjects, the number is indicated by (N=)


MEAN P-SCORE

HOARSENESS

VOICE TIRES?

VOICE A PROBLEM?

PILOT SUBJECTS 'REVISITED'
Self ratings on the Questionnaire.

Figure 89

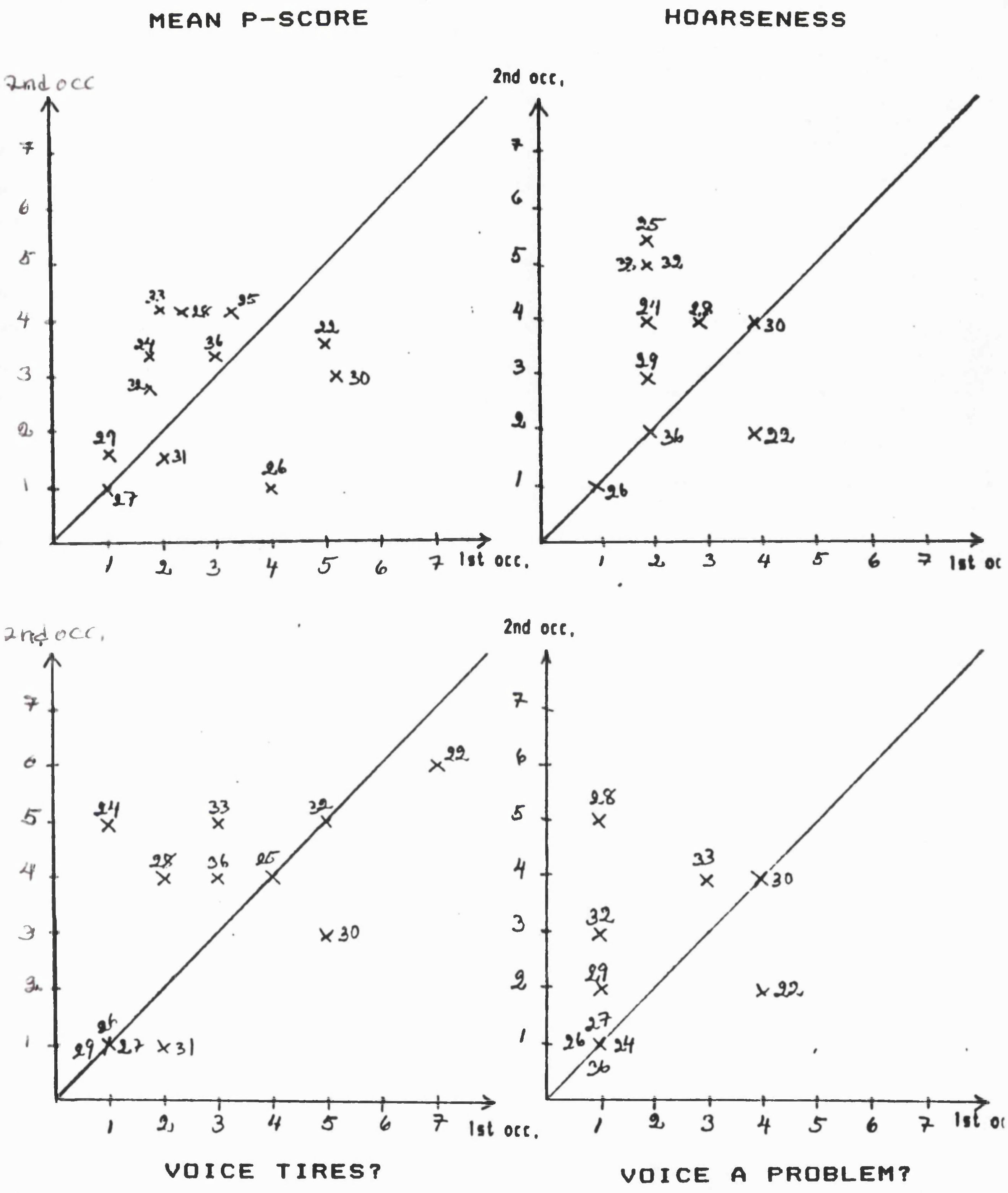
Among Main study subjects (Table 51) reassessment between two and five years later reveals an increase in their ratings of 'Hoarseness', the voice posing a 'Problem' and 'The voice tiring' after use. Shouting and speaking over noise are also rated slightly more difficult. Four of the twelve subjects are still working (subjects 22, 25, 31 and 32) and two T2 subjects, 24 and 29 have both recurred.

TABLE 51

AVERAGE SELFRATINGS ON THE QUESTIONNAIRE
MAIN STUDY SUBJECTS - 24 month interval +.

	First occasion Main Study (2MPRx+)	Last occasion Main Study
AGE		
Range (yrs)	49-70	53-73
Median (yrs)	61,5	65,0
MPRX		
Range	2-37	31-63
Median	5	49
$N_{subj.}$	12	12
Working	6	4
Hoarseness?	2,4 (N=10)	3,3 (N=11)
Voice use now?	4,6 (N=11)	4,3 (N=11)
Voice a problem?	1,8 (N=11)	2,3
*Voice tires?	2,9	3,3
*Difficulty singing?	3,2 (N=11)	2,9 (N=11)
*Difficulty shouting?	2,7	3,3
*Difficulty speaking over noise?	3,1 (N=9)	3,3
*Difficulty using the telephone?	1,8	1,5
Mean P-score	2,7	2,8

* Item included in Mean P-score. Where the averages are based on fewer than the original number of subjects, the number is indicated by (N=).



MAIN STUDY SUBJECTS
Self ratings on the Questionnaire.

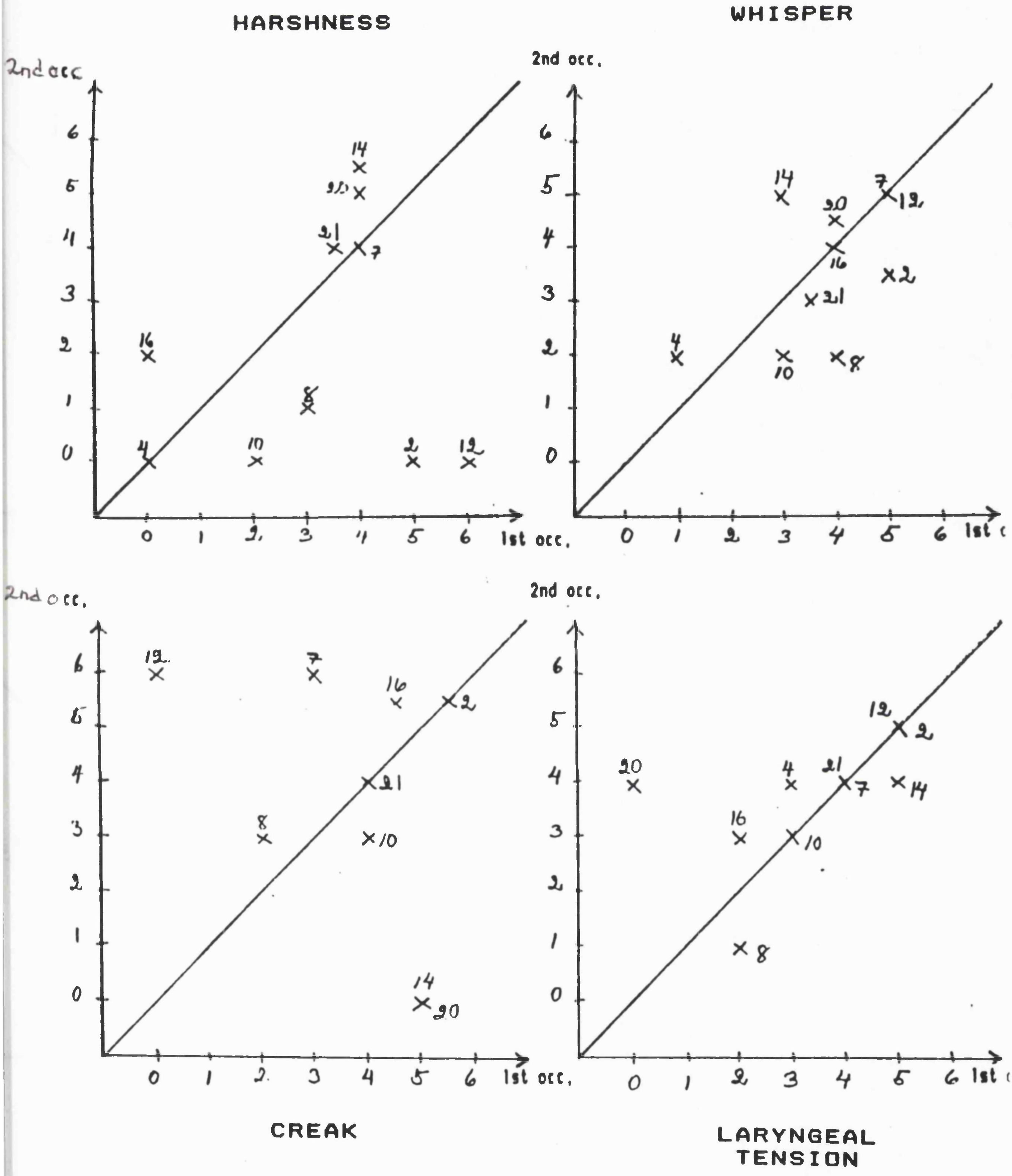
Figure 90

Inspection of their plots (Fig. 90), confirms the tendency in this group to rate symptoms and vocal limitations (Mean P-score) worse on the second occasion; there are more ratings above the regression line than below it. It can be seen, however, that subjects 24 and 29, whose tumours recurred, although they rate their symptoms as more severe on the second occasion, are not the only ones, nor do they give the highest ratings. It is notable how neither of them seems to consider the voice as a problem, despite the deterioration in voice quality.

v **VPA ratings.**

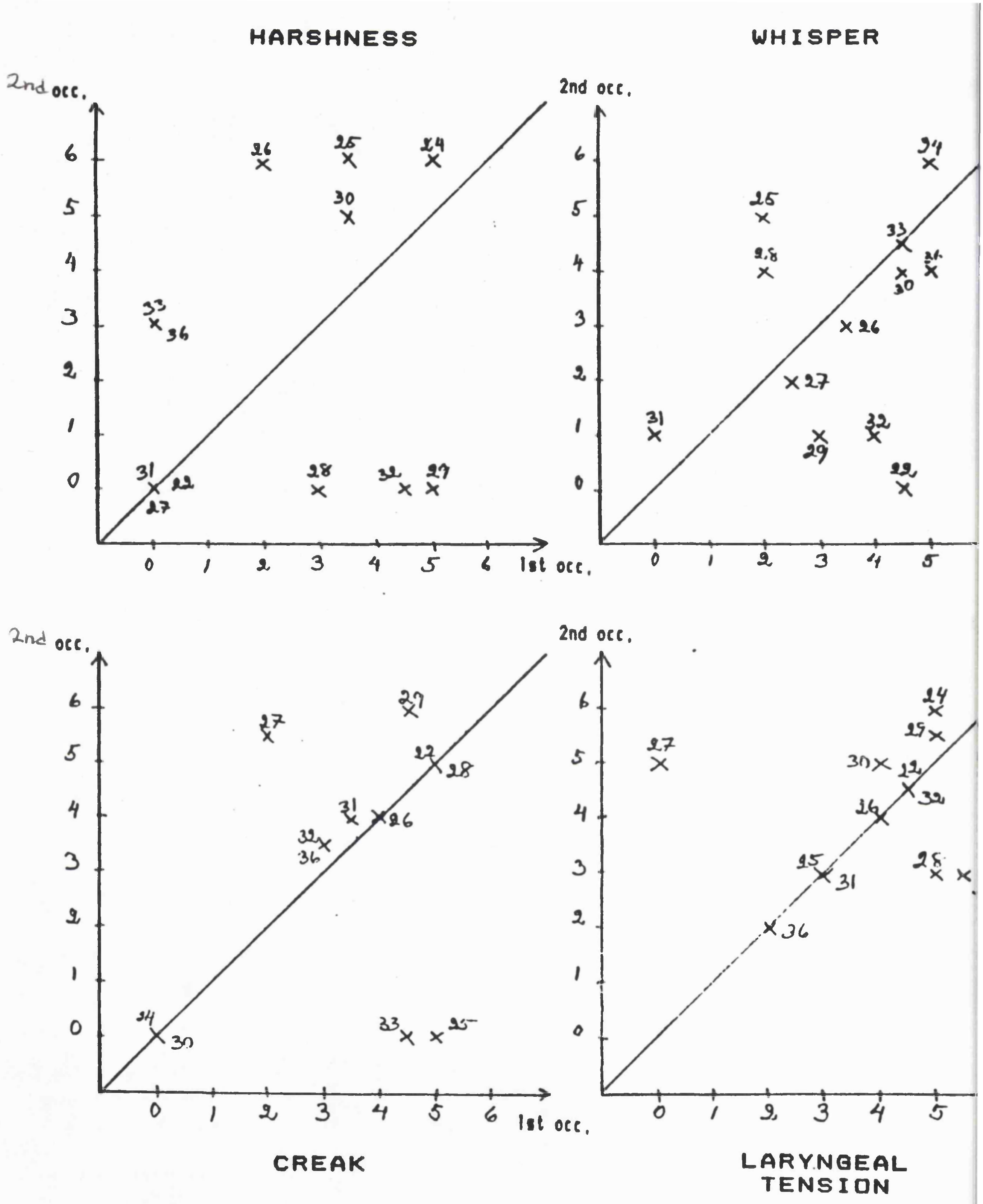
Finally, the experimenters', EC's perceptual ratings of the subjects voices on the VPA (Laver, et al, 1986) on the first and second occasions are shown in Fig. 91 and 92. The rating scale and parameters used have been described in an earlier chapter but briefly: scalar degrees 1-3 are considered to show features within 'Normal' limits. Ratings 4-6 reflect 'Abnormal' degrees of a feature. Despite variable and sometimes poor inter-rater agreement between EC and two independent judges discussed earlier (Chapter Xi, iv, v), her intra-rater agreement was good. All VPA ratings for these subjects have been carried out by EC on every recording occasion. Age ranges and MPRx are the same as in Tables 50 and 51.

Several Pilot subjects seem to have reduced degrees of Harshness and Whisper on the second occasion (Fig. 91). This may reflect the recovery of the mucosa and condition of laryngeal tissues such a long time after the end of radiotherapy. Creak shows a very different pattern in what seems to be a tendency for subjects who were perceived to have low degrees of Creak on the first occasion to have higher degrees on the second (7, 8 and 12), others have high degrees of perceived 'Creak' (2, 16 and 21). Laryngeal Tension also seems perceived to be very similar on both occasions for most Pilot subjects except subject 20. The latter two parameters may be more likely to reflect habitual use of the vocal mechanism and be less subject to change over time than Harshness and Whisper.



PILOT SUBJECTS 'REVISITED'
Vocal Profile Analysis ratings.

Figure 91



MAIN STUDY SUBJECTS
Vocal Profile Analysis ratings.

Figure 92

Among Main study subjects there seems also to be a strong tendency to perceive the same individual subjects' voices as Creaky or produced with high degrees of Laryngeal Tension on both occasions (Fig. 92). Harshness shows an increase in six Main study subjects on reassessment whereas the degree of Whisper has decreased in the majority of subjects (Fig. 92).

Subject 24 whose tumour had recurred on the last occasion shows increased Harshness, Whisper and Laryngeal tension. His voice, however, as can be seen from the plots (Fig. 92) had been perceived as very abnormal already on the first occasion at 11 MPRx.

Subject 29, whose tumour also had recurred, had an extremely low pitched voice, which may explain the very high degree of perceived Creak and Laryngeal tension (Fig. 92). He was the only subject in the study whose original diagnosis was of a stage T2 N1 tumour.

Subjects 25, with increased degree of perceived Harshness and Whisper has started smoking again; so has subject 30.

vi Summary of findings - Total Data.

The descriptive analysis suggests answers to the first two questions posed in the Main Study, Chapter VIII (p. 249):

Patients irradiated for T1 tumours of the vocal folds seem to have better voice quality than those with T2 tumours as judged by objective (Table 41) and perceptual means (Tables 42 and 43). This is also reflected in their self ratings of hoarseness and vocal limitations which are lower than T2 subjects' (Table 44).

The second question asked whether subjects treated in 5F/week have better voice quality than those treated in 3F/week. Indications are negative, as our findings instead suggest slightly more regular vocal fold vibration (%TS), and higher fundamental frequency (Fx) in subjects treated in 3F/week (Table 41). These subjects also show, less severe perceived degrees of 'Harshness', 'Whisper', 'Creak' and 'Laryngeal Tension' (Table 42 and 43, Appendix 5, Tables 1-2). This is supported by their self ratings of lower degrees of 'Hoarseness' and vocal limitations (Table 44).

The differences in Speech fundamental frequency and %TS, seem to indicate slightly more regular voices among workers than non-workers (Table 45) and lower pitch. Both these may, however, also be influenced by the age differences between the groups. Retired and unemployed T2 subjects are closer in both age and Fx to the 'Workers' than to 'retired' T1 subjects. The workers do not seem to consider their voices much more of a Problem than the non-workers, despite slightly higher self ratings of hoarseness and limitations in function (Table 46 and 47)

Although the number of subjects who are smoking is very small in each group (Table 45), it seems as if subjects treated in 3F/week have heeded advice to stop smoking more than other subjects. Only three subjects treated in 3F/week are smoking (including T2 subjects 12 and 13) but five subjects treated in 5F/week (including subject 29, T2N1, unemployed). (Table 41). This could possibly contribute to the better voice quality measures, group Fx and %TS values and VPA ratings in T1 3F/week subjects (Tables 41-43) and the worse voice quality measures among T1 subjects treated in 5F/week and among retired and unemployed subjects than among workers (Table 45).

Fewer of those who are working are smoking again, compared to those who are not (Table 45), possibly reflecting a greater concern about the effect smoking may have on their future ability to work and a concern to avoid recurrence; on the other hand, unemployment or retirement may lead to more smoking. Of the three subjects who have recurred in the T2 group, one, subject 29, is a very heavy smoker and drinker. The other two did not smoke during or after radiotherapy.

The only way to determine the effect of smoking on our subjects' voice quality would have been a randomised controlled trial, for which our number of subjects were too small.

The Pilot subjects seem to report fewer problems and vocal limitations. Due to their almost double follow-up time, part of this may be related to fewer of them still working. It is however, possible that the differences between the groups both in their self-reported voice characteristics and in VPA ratings reflect different long term effects of 3 versus 5 fractions per week treatment regimes. All Pilot subjects were treated in 3F/week, but only 2/12 Main study subjects. Only a randomised controlled study would have been able to confirm or refute this impression, however.

CHAPTER XIII - THE EFFECT OF VOICE THERAPY ON VOICE QUALITY MEASURES AFTER RADIOTHERAPY.

I Voice therapy rationale.

Hoarseness may give rise to limitations in vocal function and to laryngeal discomfort which interferes with a person's daily life and ability to interact with others at home and at work. Voice therapy involves teaching the dysphonic patient to modify the manner in which voice is produced to achieve optimum function and quality for that person's daily vocal needs. These will be slightly different for people in different professions and at different stages in their lives, but also vary according to individual inclination to socialize, to use the voice in different situations and conditions and for different purposes in and out of work. This also means that hoarseness is experienced and tolerated by different people in different ways. This may determine how much time and effort individuals are prepared to put into voice rehabilitation.

There are some reports of the beneficial effect of voice therapy on vocal function in the literature (Gramming, 1988, Motta et al, 1990, Kotby, El-Sady, Basiouny, Abou-Rass and Hegazi, 1991, Kitzing and Åkerlund, 1992, Carding and Horsley, 1992, Södersten and Hammarberg, 1993, Carlson, 1986, 1988 a, 1993 a, b). The only one pertaining to voice rehabilitation of irradiated subjects however, is the one earlier referred to by Fex and Henriksson (1969), who found that the effects of radiation damage could be reduced by offering voice therapy, in their case, in parallel with radiotherapy. Both Stoicheff (1975) and Lehman et al (1988) suggest that such patients might benefit from therapeutic intervention.

Methods used to effect change in vocal behaviour vary greatly and a summary of direct and indirect approaches which may be used by therapists on their own or in combination is given by Carding and Horsley (1992). Single case studies and therapeutic approaches are reported in Stemple (1993), many of which employ objective measurements in evaluating the effectiveness of treatment.

ii The Accent Method.

The method used in treatment of the irradiated subjects in this study is based on the Accent method (Smith and Thyme, 1976, 1978, Kotby et al 1991). The approach emphasises abdominal control of expiratory airflow to regulate subglottal pressure. This achieves increased speed and duration of vocal fold closure through enhancement of the Bernoulli effect resulting in reduced glottal air wastage (Frøkjær-Jensen, 1983, Kotby et al, 1991) and enhancement of the number and intensity of overtones (Smith and Thyme, 1978).

Kotby et al (1991) suggest that the expiratory breath pulses used in phonation exe rcises in the Accent method result in:

a) restoration of the symmetry of the vibrator (glottis).
b) reduction in air wastage through the glottis.
c) reduction of excessive glottic muscular effort.

They compared patients' ratings of their voices on a five point scale; therapists' voice quality ratings using a modified GRBAS scale (Hirano, 1981) and objective voice quality measures pre- and post voice therapy using the Accent method. Therapy was given to 28 subjects in 20 minute sessions three times per week for 20 to 25 sessions.

Statistically significant differences were found in perceptual ratings of degree of dysphonia (G), Strain (S), and 'Leaky' voice (L): all were reduced; Maximum phonation time (MPT), Maximum flow rate (MFR), subglottal pressure and glottal efficiency (GE) had also improved significantly. Unfortunately they do not describe how these parameters are measured. They conclude, however, that the measures appear helpful in evaluation of the effectiveness of voice therapy.

Using videostroboscopy they demonstrated reduction in size of vocal nodules after voice therapy in 6/6 subjects and also reduction in 'phonatory gap' in 4/6 cases.

In the present study, an initial detailed explanation of laryngeal structure and function using a model of the larynx and Xeroradiographs of the vocal tract at rest and during phonation, was followed by teaching voice production and control according to the Accent method. The technique for breath control is taught in graded exercises where the patient's attention is moved from the larynx to the abdominal musculature. The emphasis is on reduction of laryngeal effort by using an open vocal tract, a lowered larynx and phonation at low pitch and intensity, to get the patient to experience the driving force of the voice source as breath support not laryngeal muscle power.

iii Biofeedback as an aid to voice therapy.

Biofeedback offers a means of monitoring aspects of an individual's internal physiological environment to increase his awareness and perception of bodily processes that are not ordinarily consciously perceived (Stemple, Weiler, Whitehead and Komray, 1980). It has been used as a tool in behaviour modification (Brown, 1975) and with particular reference to voice disorders by Holbrook, Rolnik and Bailey (1974), Sturlaugsson (1975), Lyndes (1975), Wechsler (1977), Stemple et al (1980) and Andrews, Warner and Stewart (1986).

Its effectiveness is explained by theories of operant conditioning where a target behaviour is 'shaped' by making desirable consequences contingent on gradual approximations of the behaviour.

Stemple et al (1980) report an experiment where electromyographic (EMG) biofeedback was used in training seven dysphonic subjects with hyperfunctional voice disorders (vocal nodules) to reduce general laryngeal area muscle tension. Treatments using EMG were given in twice weekly sessions for four weeks, including one session where no feedback was offered. Significant correlations were found between improved perceptual voice quality ratings and EMG measures after therapy.

Another study using EMG biofeedback from the cricothyroid muscle for tension reduction in hyperfunctional dysphonias is reported by

Andrews et al (1986). They compared the effectiveness of EMG feedback for inducing reduction in laryngeal tension with a program of 'Progressive relaxation'. Both approaches were effective but there was no difference between them. Both methods were used as part of a 'graded voice therapy program' including the teaching of *'expiratory control, humming on a monotone, pitch and range exercises and extending phrase length'*.

Significant improvement in voice quality was achieved with both approaches in 'control of vocal fold vibration' as judged from the smoothness and regularity of Fx contours obtained by ELG. The speech tasks recorded were sustained phonation and glides on [ɑ] and repetition of six standard sentences. Significant improvement was also found in subjects' self rated voice quality measures.

The authors concluded regarding EMG feedback that: *'Although the EMG was useful in detecting global increases in laryngeal tension, it did not detect increases in adductive tension or medial compression in glottal attack. Hard attack in 3/5 EMG subjects had to be eliminated through auditory monitoring'*.

iv The Lx waveform as visual feedback in voice therapy.

The Lx waveform reflects the speed, extent and duration of vocal fold contact (Fig. 23 p. 135 and Fig. 24, p. 137). Orlikoff (1991) derives from the Lx waveform a Contact Quotient (CQ) calculated as the ratio of the duration of the Contact Phase (CP) and the total period (Fig. 35, p. 154) and suggests that this ratio is related to the degree of approximation and relative compression of the vocal folds in the horizontal plane.

The difference between the duration of the 'contact closing phase (ccp)' of the waveform and the duration of the opening phase (cop) divided by the duration of the Contact phase are used in calculation of what Orlikoff calls a 'Contact Index'. He suggests this reflects vocal fold tonus and vertical fold dynamics. Both CQ and CI have been described in detail in Chapter V, vii, with relevant experimental evidence.

Facilities for measuring Lx waveform features were not available until late in this study. The relative duration of the 'Contact closing phase', as reflected in the steepness of the closing slope (Fig. 43, p. 202), duration of the Contact phase relative to total period and the relative amplitude of the Lx waveform have, however, been found to offer extremely powerful means of providing real time visual feedback of changing and improved glottal aerodynamics in voice therapy based on the Accent method (Carlson, 1986, 1988 a,c, 1992, 1993 a, b).

On the first assessment occasion the shape of the Lx waveform is explained to the patient and the effect on Lx of varying degrees of breath support, laryngeal tension and vocal intensity are demonstrated by the therapist. As all the subjects in this study were male the difference in waveform features between males and females was also explained.

Lx waveforms produced by excessive laryngeal effort as in hyperfunctional voice disorders (Fig. 93 a), were quite prevalent among the irradiated subjects in this study. Lx has been found to be very effective in helping such subjects to reduce laryngeal effort and use abdominal breath support to sustain phonation with much reduced harshness and irregularity of vocal fold vibration (Fig. 93 b) (Carlson, 1986, 1988 c, 1993 a). In such cases it was particularly the excessive duration of the contact phase and limited amplitude of the waveform that were the target for modification.

The target waveform in rehabilitation of puberphonic male voices provides another example. The aim is increased duration of the Contact Phase, particularly the 'Contact opening phase (cop)' (Orlikoff, 1991) and increased amplitude of the waveform (Fig. 8 b, p. 72). Waveform modifications reflect changed laryngeal aerodynamics from falsetto to modal voice production (Fig. 8 a and b, p. 72) (Carlson 1993 b).

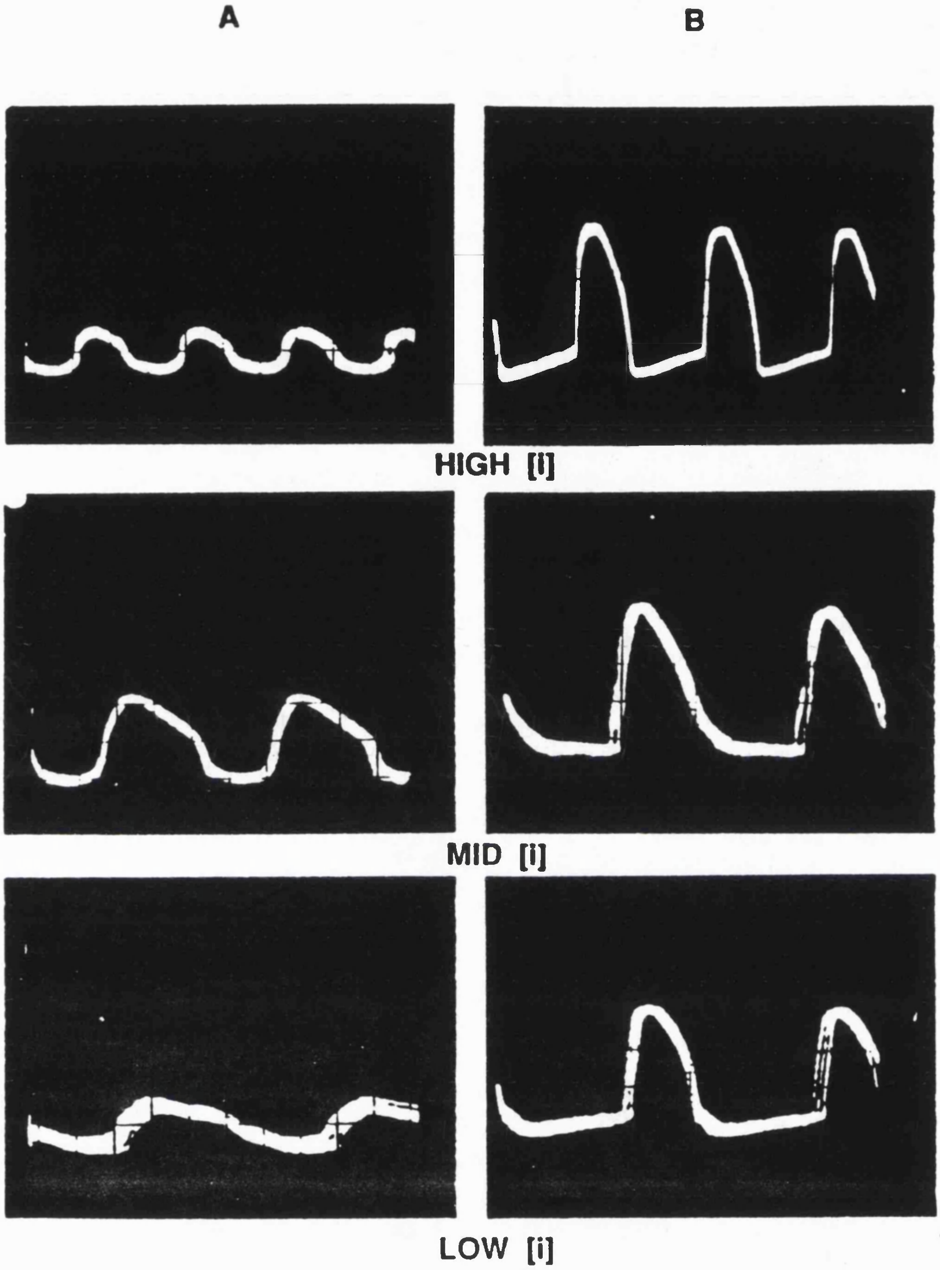


Figure 93 Subject 22 - Lx waveforms
a) before b) after voice therapy

v Approach to treatmnent of voice problems after radiotherapy.

The only difference in approach compared to the therapy offered any other voice patient, was an explanation of the irreversible effects of irradiation on laryngeal irrigation and mucosa and the possible limits imposed on recovery of a completely normal voice as a result of this. The need to heed and not exacerbate the symptoms of dryness, irritation and vocal fatigue that may be experienced as a result were emphasised. The features and order of voice assessment and therapy stages can be described as follows:

a) Explanation of normal laryngeal structure and function and the expected effects of irradiation, smoking, alcohol intake, and vocal abuse as relevant to the individual.

b) Explanation of the Lx waveform as described above and rationale for its use in recording voice samples at different times post radiotherapy and voice therapy.

c) Recording of conversational speech and reading aloud

d) Recording and storing of Lx waveforms of sustained [i] at habitual mid, high and low pitch and at comfortable volume.

e) Auditory analysis of the individual subject's perceived features of habitual voice production, and an outline of the main aims of therapy in his particular case.

f) Explanation of the rationale behind the teaching of abdominal breath control and increasing awareness of the difference between habitual and new breathing patterns for voice production.

g) Demonstration of the differences in effect on the Lx waveform between control of phonation by laryngeal adduction forces as opposed to abdominal breath support. Auditory feedback is offered in conjuction with this.

h) Patients are given tapes of recorded exercises of general body relaxation, graded breathing exercises according the the Accent method, initially using voiceless fricatives, gradually introducing voiced fricatives, and vowels inititated on [h] to avoid the use of hard glottal attack. In the clinic situation, once the breath support technique is beginning to be mastered, some of these are practised using the visual feedback offered by Lx, but patients are expected to practise at home between sessions without such feedback.

Mastery of the technique of abdominal breath control and support for phonation with reduction in laryngeal effort and tension is basic to all and any further voice production exercises and applications used in the clinic. These may vary according to the degree of dysphonia, symptoms and expressed needs of the individual, but usually consist of gradual introduction of more speech-like material first in the form of adapted sentences that facilitate the maintenance of well sustained, effortless modal voice and a balanace between breath support and laryngeal adduction forces.

Emphasis is put on maintenance of phonation which reflects regular, periodic vocal fold vibration. To achieve this, a departure is made from strict Accent method technique and chanting on a monotone at a comfortable pitch in the middle or low range of the individual's speaking voice is employed, often using fundamental frequency (Fx) contours as feedback of smoothness and regularity of vocal fold vibration (Fig. 94 a and b).

Once this is achieved, the same phrases are spoken immediately after chanting. This allows the patient to experience the difference in phonatory control between the chanting mode of vocal fold vibration, sustained with effective breath support, and habitual voice production in speech with varying intonation and stress patterns where often interference of excessive laryngeal tension is introduced as revealed in broken or irregular Fx contours (Fig. 94 a). The experience reminds the individual to maintain a more comfortable speaking voice by maintaining breath support. At this stage, several weeks into the period of attending for voice therapy, patients

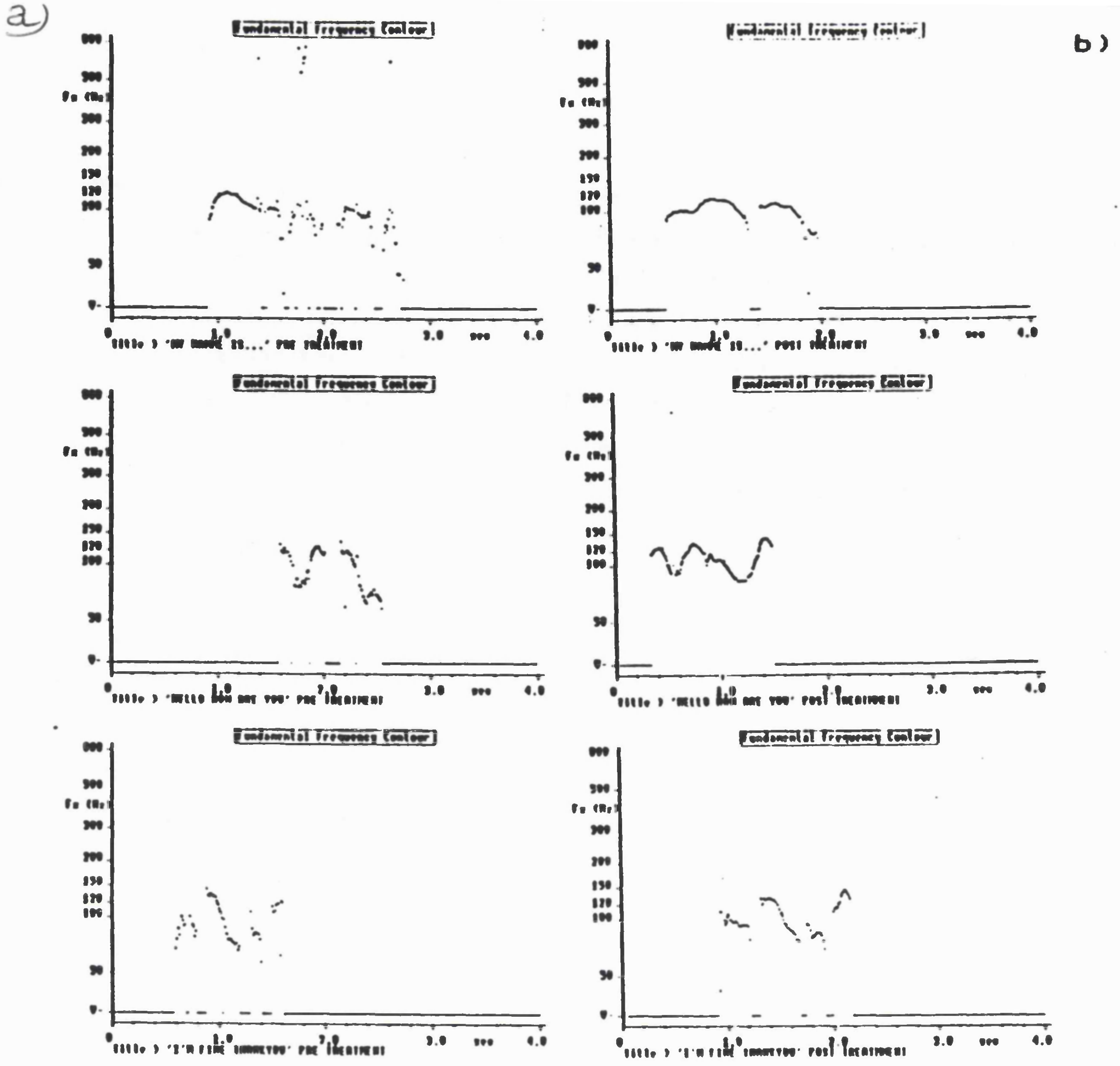


Figure 94　Subject 22 Fx contours
a) before voice therrapy, 5 MPRx
b) after voice therapy, 12 MPRx

usually report increased awareness of the need for breath support in everyday speaking situations and are beginning to monitor themselves.

Carryover into natural speech is encouraged in the clinical situation by offering the patient increasingly demanding reading material including poetry and prose passages that are not adapted to facilitate but to demand greater control of prosody, vocal pitch and intensity.

Successful therapy outcomes aimed at improving vocal function should be possible to demonstrate by subjective and/or objective voice quality measures as employed in this study and as demonstrated in the aforementioned case studies reported by the experimenter. To what extent such evidence is demonstrable, depends to a great extent on the severity of the voice problem, the potential for improved function as evidenced by laryngeal tissue status, and the motivation and need for the patient to improve and modify his habitual pattern of voice production.

This clinical study of voice quality features after radiotherapy for laryngeal carcinoma was not designed to conclusively prove the effect of voice therapy for such patients. The subjects who underwent courses of voice therapy were self selected and varied greatly in the severity of their voice problem and the effect they perceived it had on their quality of life. What follows is a descriptive analysis of the trends in objective and perceptual voice quality measures used in this study before and after a course of voice therapy. A comparison is made with the same measures from a group of subjects who did not need or request therapy, or who declined it when offered.

vi Subjects.

Voice therapy was offered to subjects who were complaining of vocal limitations and/or laryngeal discomfort after irradiation, irrespective of whether laryngeal examination revealed any abnormalities apart from the expected slight erythema of laryngeal tissues (Fig. 95) (Appendix 16, Table 3).

Subjects 14 and 16 had received voice therapy before and are also described in the Pilot study (Table 11, p. 241).

Subject 14 was, however, experiencing severe late radiation fibrosis with swallowing difficulties and severe deterioration in voice quality and function. He had now retired from work.

Subject 16 requested help as he was finding increasing difficulty in maintaining good voice quality. There was no evidence of laryngeal tissue deterioration or recurrence. His right vocal fold had always been described as 'red' and his laryngeal mucosa 'dry' throughout the five year follow up period, before he was referred again. He was troubled by a lot of mucus, suffered from asthma and was on medication for this. He was in full time employment as catering manager for a large company.

Subject 21 had recently, almost five years after the end of radiotherapy, had a biopsy of an extensive area of leukoplakia on his irradiated right vocal fold. There was no evidence of malignant changes, however. He was working as a furniture restorer.

Subject 22 was referred by the ENT surgeon as his voice was found to take a long time to recover and there was evidence of persistent oedema. He was managing director of his own company.

Subject 25 was referred as he found his voice did not stand up to all the talking he had to do as a lecturer. Throughout the period before referral there was evidence of persistent oedema.

Subject 33 was dissatisfied with his voice as he had been a keen amateur singer and could no longer take part in singing in church. He also complained of irritating dryness of his throat. He was a retired diplomat and a great 'communicator'.

Subject 39 had received radiotherapy for a T1 tumour of his right vocal cord in December 1992. He had not been referred for assessment or advice and when contacted for assessment in February, he was virtually aphonic, producing extremely effortful voice using his ventricular bands as confirmed by videostroboscopy. He was very keen to try to improve his voice quality as he was embarrassed to talk to people.

Finally, Subject 40 was referred nine months after the end of radiotherapy as his voice was not improving as expected. He had been found to suffer from myxoedema three months prior to referral and was treated with thyroxine. He used his voice a lot in his work as an editor.

Two more subjects received voice therapy after radiotherapy but are not included in the descriptive analysis: Subject 24 suffered a recurrence of his tumour and his voice quality showed a steady decline (Appendix 9 A and B). In the case of Subject 28 there were many problems with data collection. He did not return his questionnaire on the first occasion. He was using his ventricular bands for phonation initially, resulting in extremely poor Lx signals and erratic objective measures and waveforms on the first few recording occasions at two and three MFRx (Appendix 10 A).

The average age of the voice therapy subjects was 61.6 years ranging from 44 to 71 (Appendix 16, Table 3). Seven subjects had been treated for T1 tumours, two for T2 tumours (Appendix 2 A and B). The three Pilot subjects who were re-referred and subject 22 had been treated with 3F/week, all the others in 5F/week.

Voice Therapy Subjects

Subject Number	Laryngeal examination findings.	
	On referral	At the end of therapy/Last assessment occasion.
14	IDL - Poor adduction and abduction	Stroboscopy - Stiff supralarynx Good movement and mucosal wave on right. Left v.c. thin, no mucosal wave. Bilat. bulky false vocal folds.
16	IDL - Both cords smooth and mobile. Normal looking mucosa.	Stroboscopy - Irradiated right v. c. slightly thinner than the left. Reduced mucosal wave.
21	IDL - Irregular right v.c. on site of biopsy.	
22	IDL - Bilaterally 'red' and inflamed v. c. and arytenoids.	IDL - A little redness on the the right v. c. No oedema.
25	IDL - Cords look normal	Stroboscopy - Irradiated left v. c. looks thin with a reduced mucosal wave. Right cord moving well with good mucosal wave.
30	IDL - Splinter haemorrhage on irradiated left v.c. slight chronic laryngitis.	IDL - Telangiectasia left v.c. Stroboscopy - Irregular anter. right v.c. Very bulky false c.
33	IDL - No sign of recurrence.	IDL - No sign of recurrence. Stroboscopy - Thin left v.c. bulky false cords.
39	Stroboscopy - Bulky false cords, ventricular band phonation, oedema of posterior left v.c.	
40		Stroboscopy - slightly atrophied irradiated right v.c. moving well. Bulky false cords. Stroboscopy - Cords move well, Good mucosal wave.

FIGURE 95

Table 52

Time elapsed since the end of radiotherapy before and after voice therapy.

Subject number:	**14**	**16**	**21**	**22**	**25**	**30**	**33**	**39**	**40**
Assessment pre voice therapy (MPRx)	149	61	55	5	12	9	3	3	9
Assessment post voice therapy (MPRx)	156	63	63	9	14	12	7	-	10
Number of sessions	*8*	*6*	*8*	*11*	*15*	*5*	*3*	*9*	*6*
Last Assessment (MPRx)	165	94	101	57	50	40	35	12	16

The time that had elapsed since the end of radiotherapy in the Voice therapy group was less than one year (except for the three Pilot subjects) (Table 52). A comparison will be made between their voice quality measurements before (Pre) and after (Post) voice therapy and on the 'Last' assessment occasion, with those of a group of subjects referred to the Main study, who did not request or need voice therapy treatment, the 'No therapy' group.

As shown in Table 52, the number of sessions offered to subjects varied between 3 and 15. The interval between pre and post voice therapy evaluation varied between 1 and 8 months.

Subject 39 had great difficulty producing modal voice in conversation during his period of voice therapy. There was some evidence of modal phonation in exercises using Lx for visual feedback but stroboscopy examination shortly after initiation of therapy confirmed consistent use of his ventricular bands for phonation and some oedema of the posterior end of his left vocal fold (Fig. 95). Concern about possible recurrence of the tumour in view of his continued smoking resulted in discontinuation of his therapy and referral back to ENT for further direct laryngoscopy. There was however no sign of recurrence. His Lx signal was extremely poor as a result of his ventricular band phonation. No Post- voice therapy assessment was therefore carried out (Table 52). When reassessed nine months later, on the last occasion (Fig. 96), his voice was considerably improved, however.

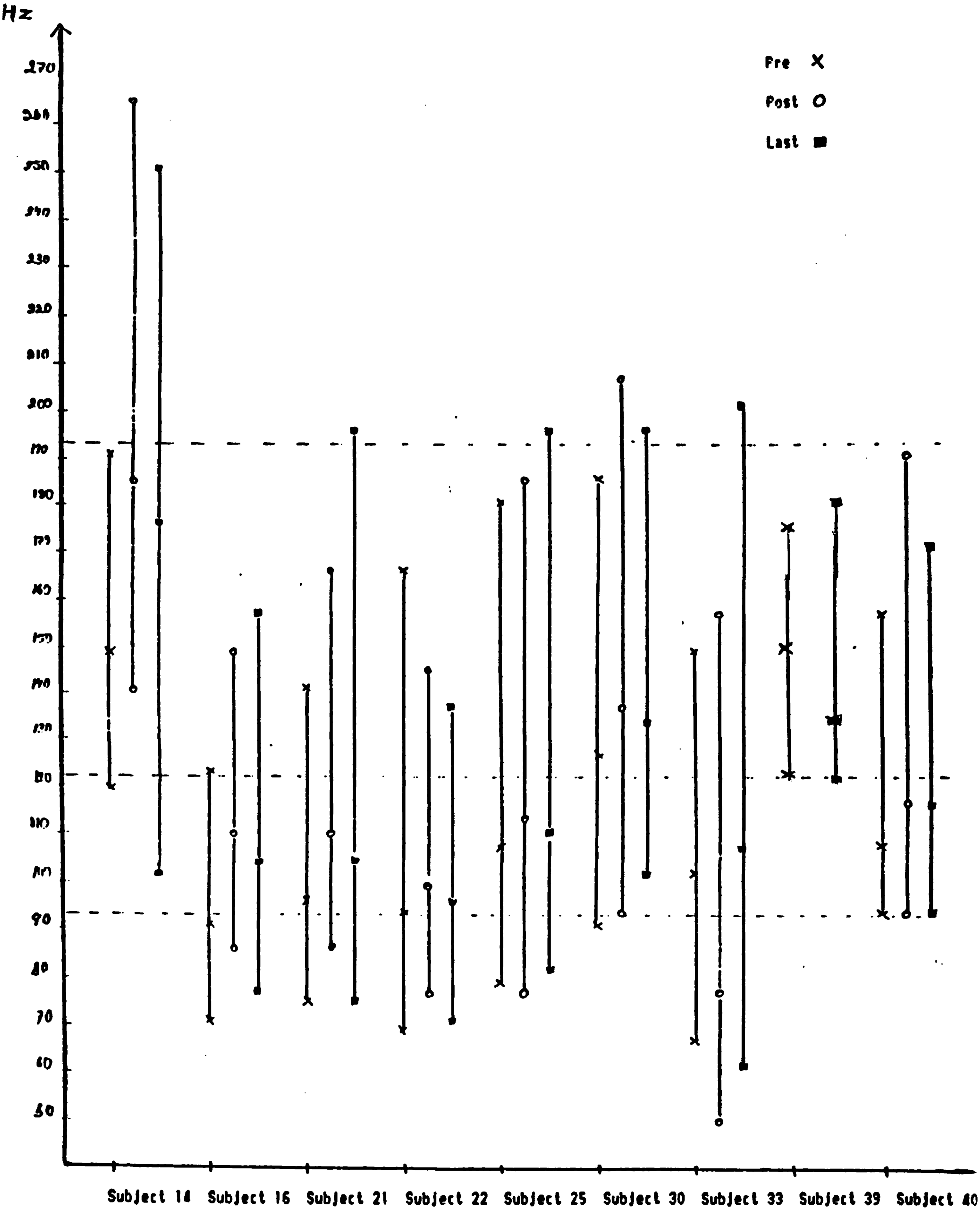
vii Description of trends in the objective measurements.

In an earlier part of this study irradiated subjects were found to have significantly lower Maximum range measures for Speech (Table 29 and 30, pp. 307-8) and less regularity of vocal fold vibration, %TS (Table 28 a and b, p. 306), than normal speakers.

Figure 96 illustrates fundamental frequency values derived from conversational speech samples for the subjects before and after the end of a course of voice therapy. Second order Mean, 90 % Maximum and Minimum range measures are given within each bar for each occasion. Crosses refer to the Pre- therapy assessment, circles to the post therapy occasion. Squares indicate measures derived from the last occasion for this study and will be discussed later. Dashed lines indicate 'Normal Average' 2nd order Mean, Maximum and Minimum ranges, arrived at in the study of the significance of differences between a random group of Normal speakers and a group of T1 subjects described in Chapter X.

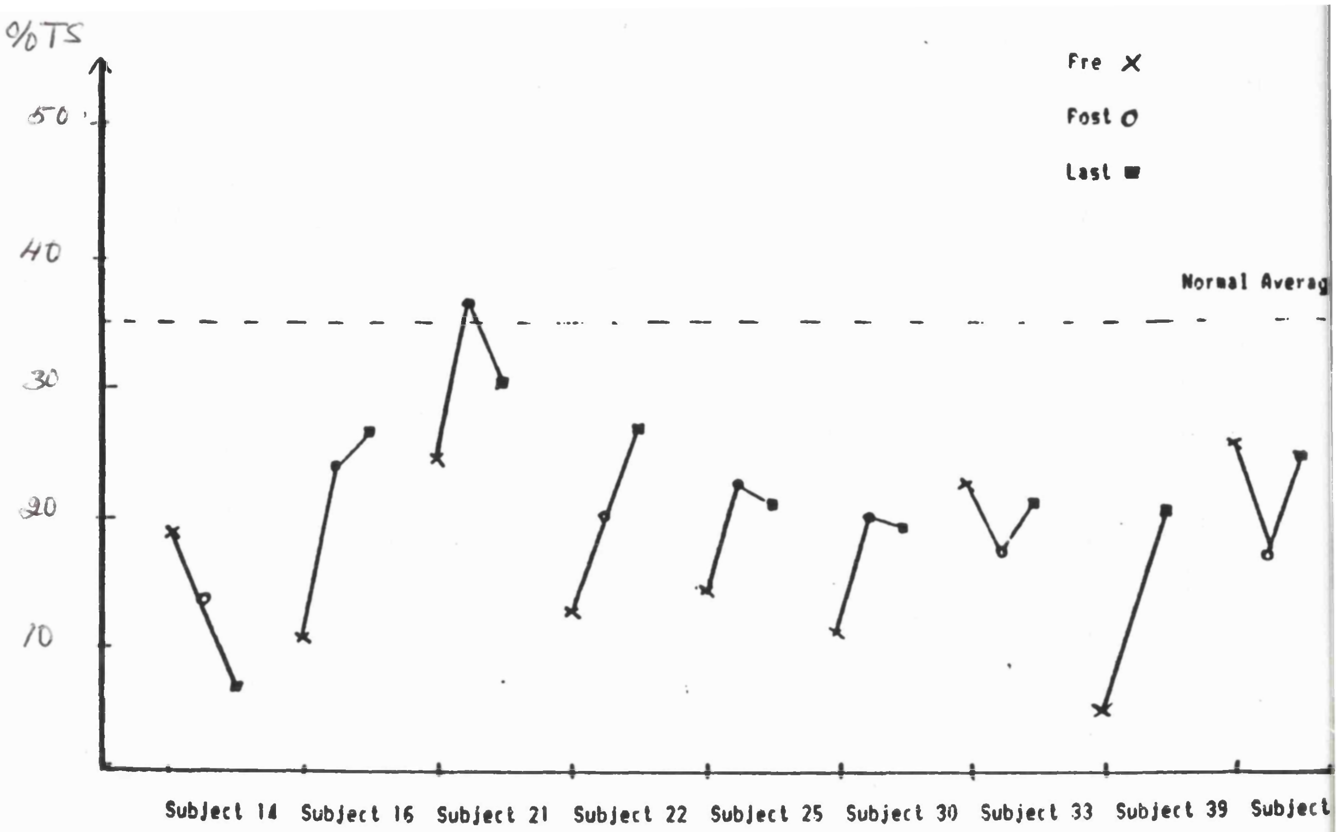
The following observations of fundamental frequency measures before and after voice therapy in Figure 96 are made:

a) Seven of the eight voice therapy subjects who were recorded both before and after, except number 33, increase their Mean from the Pre- to the Post therapy assessment. Most subjects use a lower Mean than the Average Normal speaker, particularly subjects 22 and 33. The latter two were treated quite soon after the end of radiotherapy (Table 52) and may still be in the process of recovering. Subject 33, treated for a T2 tumour, received only three voice therapy sessions before reassessment as his main concern was his singing voice and he did not seem inclined to modify the manner in which he habitually used his speaking voice. A month later he was found to have a



SUBJECTS WHO RECEIVED POST RADIOTHERAPY VOICE THERAPY
Speaking fundamental frequency
2nd order Fx Mean and Range

Figure 96



SUBJECTS WHO RECEIVED POST RADIOTHERAPY VOICE THERAPY

Speech %TS

Figure 97

recurrence in some neck nodes and had a radical neck dissection. His larynx was not involved.

b) The higher Mean Fx is the result in most subjects (6/8) of an extended and higher range of fundamental frequencies in their conversational speech after the end of therapy. Few, however, approach the level indicated of 'Normal Average' maximum range. Only subjects 25 and 40 seem to achieve this on this occasion. Five of eight subjects have Minimum range levels well below Normal average levels. Barry et al's (1990 b) observation of less variability in Minimum than in Maximum range measures seems to be confirmed.

c) Subject 14 with severe late radiation fibrosis shows abnormally high fundamental frequency values both before and after therapy. It may be due to stiffness in the vocal cord 'cover' and the lack of a mucosal wave on the irradiated side (Hirano et al, 1991). This was observed on Stroboscopy examination on the last assessment occasion (Fig. 95). Subject 39 also shows abnormally high Fx Mean before therapy. It is important to note that his data is only based on 5.6% of the total sample carried into second order (%TS) (Fig. 97) as a result of his ventricular band phonation. He was recorded only three months after the end of radiotherapy and had continued smoking. As explained above no Post therapy assessment is available for subject 39 as therapy was discontinued.

Figure 97 illustrates the regularity measure %TS, corresponding to the conversational speech samples represented in Figure 96. The same symbols are used for each occasion and we are first concerned with the Pre and Post voice therapy symbols, crosses and circles repectively.

It is noted that five of eight subjects recorded on both occasions, 16, 21, 22, 25 and 30, have increased regularity of vocal fold vibration quite considerably. Subjects 14, 33 and 40 show a decrease. In subjects 14 and 33 this reduction in vocal fold regularity is also reflected in their responses on the questionnaire as increased ratings of Hoarseness and Mean P-score, parameters which were earlier found to correlate negatively with %TS (Table 39, Chapter XI, ix). In all three subjects the decrease in %TS is accompanied by an increased range of Fx (Fig. 96). No subject except number 21, the youngest subject in the Voice therapy group, irradiated for a T2 tumour, approaches the level indicated by the 'Normal Average' of 35 %.

The group of subjects who received voice therapy was self selected or seen on request of an ENT surgeon, who was concerned about lack of expected improvement in voice quality after radiotherapy. As mentioned in the introduction to this chapter, individuals will vary in how much time and effort they will put into changing habitual voice production and to what extent they experience the effects of irradiation on laryngeal tissues.

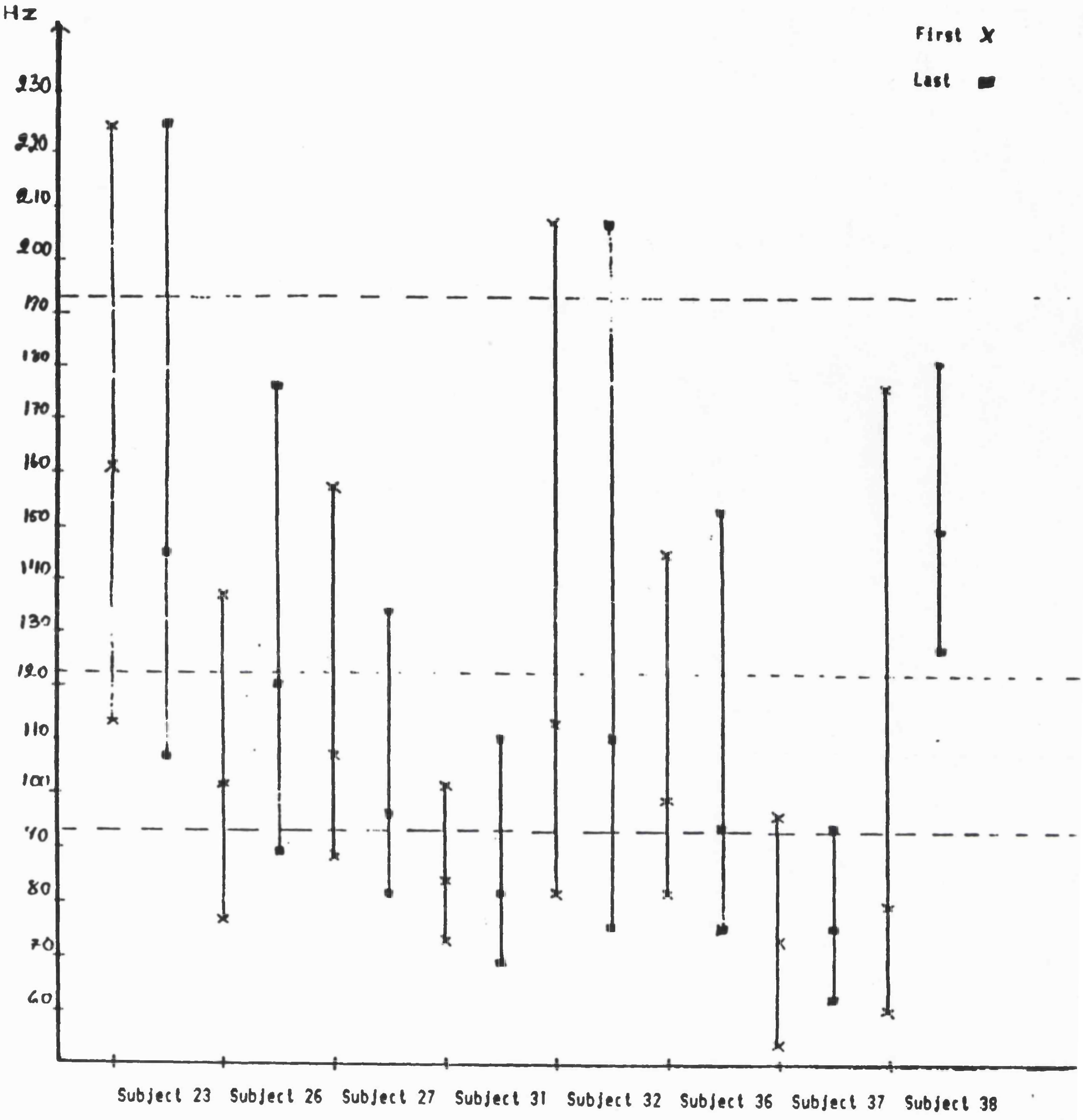
For comparison purposes, voice quality measures on the First and Last recording occasion for the Main study, of a group of subjects who did not request or were not considered to need voice therapy are shown in Fig. 98 and Fig. 99. Crosses indicate the first assessment and squares the last assessment for this study. Details of the subjects' ages, occupations and smoking habits are shown in Appendix 16, Table 4. Their assessment occasions are shown in Table 53. The Voice therapy subjects' Last assessments are also represented by square symbols in Figures 96 and 97.

Table 53.

No Voice therapy

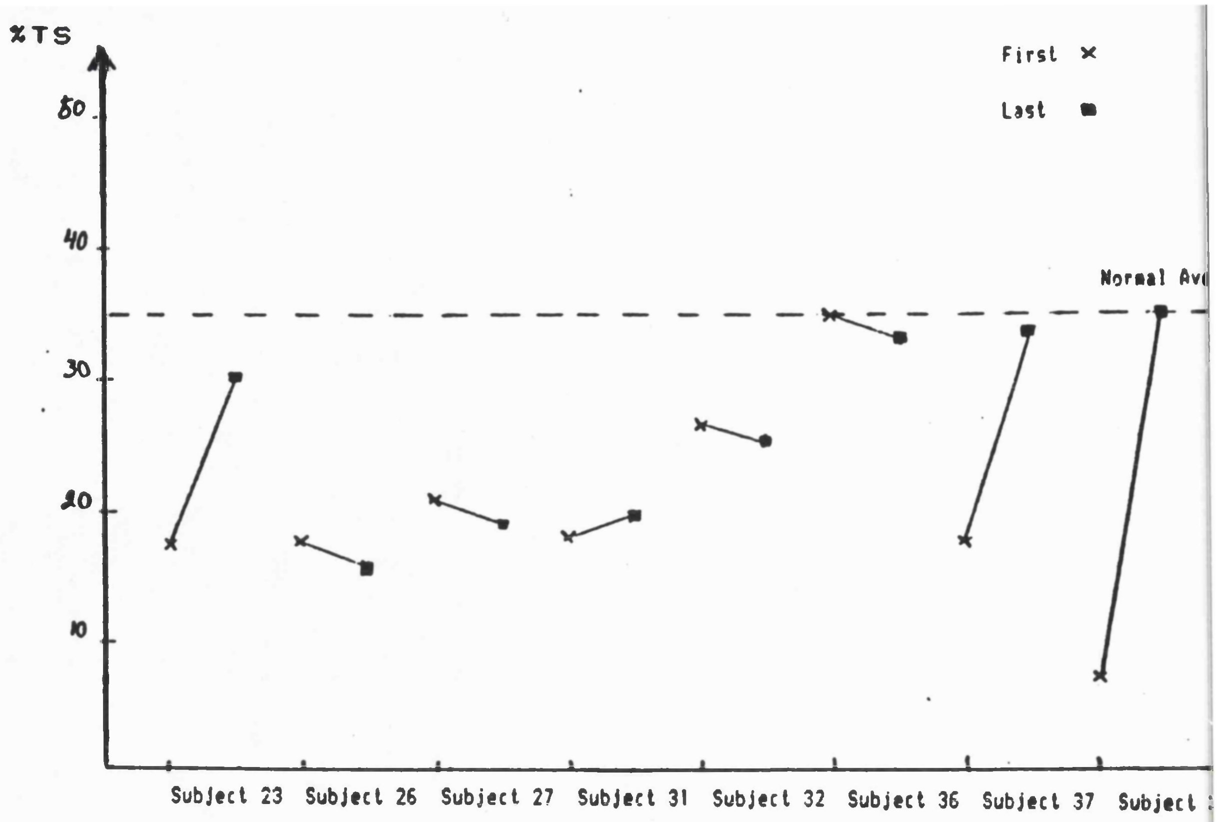
Subject number:	23	26	27	31	32	36	37	38
First Assessment (MPRx)	3	4	18	6	5	37	5	168
Last Assessment (MPRx)	5	61	63	37	48	63	11	178

Three subjects who did not receive therapy are excluded from the analysis; Subject 29 as his tumour recurred; Subject 34 as he was only recorded once and subject 35 as he was the only non-English



SUBJECTS WHO DID NOT RECEIVE POST RADIOTHERAPY VOICE THERAPY
Speaking fundamental frequency
2nd order Fx Mean and Range

Figure 98



SUBJECTS WHO DID NOT RECEIVE POST RADIOTHERAPY VOICE THERAPY

Speech %TS

Figure 99

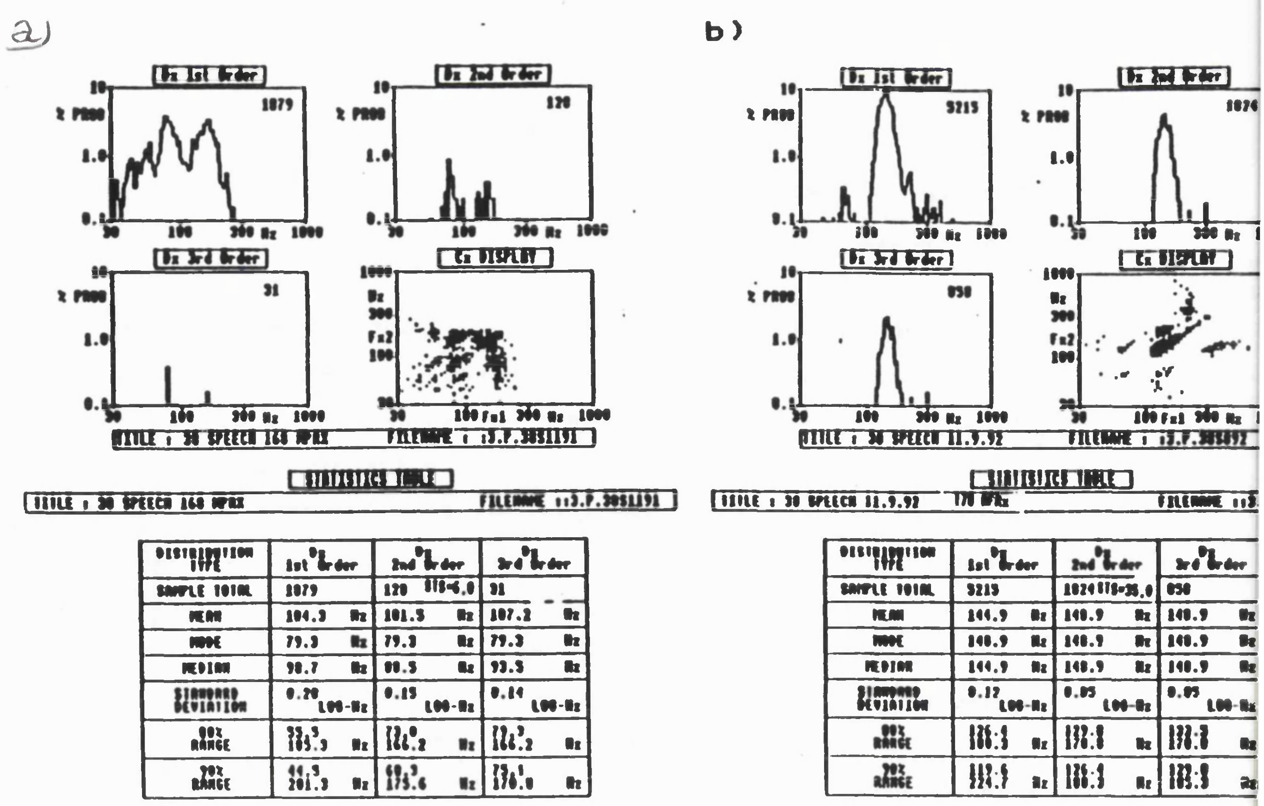
speaking subject, which might bias both objective and VPA measurements.

Inspection of Figure 98 showing fundamental frequency measures for conversational speech in the untreated group shows a decrease in Fx Means from the first to the last occasion in five of eight subjects (23, 27, 31, 32, and 36). In the Voice therapy group (Fig. 96) all except one subject increased their Fx Mean from the Pre- to the Post-therapy occasion. Their Mean Fx on the last occasion has decreased compared to the immediate post voice therapy occasion but not to the level before therapy (14, 16, 21, 22, 25, 30). As in the untreated subjects (23, 31, 32, 36) a slight decrease in Fx Mean on the last occasion is often accompanied by an increased overall range in the treated group (14, 16, 21, 25) (Fig. 96). Subject 39 had a much improved voice quality compared to his Pre voice therapy assessment, and an Fx Mean and range within more normal limits.

Figure 98 shows a very slight increase in Mean Fx in untreated subject 37 and an extreme increase in subject 38's modal value on the last occasion. Modal values are used as the first distribution is multimodal and the Mean would not be the appropriate central measurement to describe such a distribution. On the last occasion the Mean and the Mode are the same as the distribution has now a single Mode as the voice quality has improved as a result of treatment for severe oedema (Fig. 100 a and b).

Subject 38 was another subject who was suffering from severe late radiation fibrotic changes in the neck and larynx exacerbated by heavy smoking throughout the follow up period. Direct laryngoscopy had shown a fixed left vocal cord but biopsies did not show recurrence of the tumour just previous to the first assessment, when his voice was extremely effortful and intermittently aphonic. Fx values are based on only 7 % of the total sample (%TS) (Fig. 99).

The high Fx values of subject 38 on the last occasion (Fig. 98) may be due to stiffness in the vocal fold cover as suggested by Hirano et al (1991). The data is now based on 35 % of the recorded sample



a)

DISTRIBUTION TYPE	Dx 1st Order	Dx 2nd Order	Dx 3rd Order
SAMPLE TOTAL	1079	120 ITS=6.0	31
MEAN	104.3 Hz	101.5 Hz	107.2 Hz
MODE	79.3 Hz	79.3 Hz	79.3 Hz
MEDIAN	98.7 Hz	98.5 Hz	93.3 Hz
STANDARD DEVIATION	0.20 LOG-Hz	0.15 LOG-Hz	0.14 LOG-Hz
80% RANGE	55.5 185.3 Hz	73.0 166.2 Hz	73.3 166.2 Hz
90% RANGE	44.5 201.3 Hz	68.3 175.6 Hz	73.1 176.0 Hz

b)

DISTRIBUTION TYPE	Dx 1st Order	Dx 2nd Order	Dx 3rd Order
SAMPLE TOTAL	3215	1024 ITS=35.0	850
MEAN	144.9 Hz	140.9 Hz	140.9 Hz
MODE	140.9 Hz	140.9 Hz	140.9 Hz
MEDIAN	144.9 Hz	140.9 Hz	140.9 Hz
STANDARD DEVIATION	0.12 LOG-Hz	0.05 LOG-Hz	0.05 LOG-Hz
80% RANGE	126.4 168.3 Hz	132.8 170.8 Hz	132.3 170.8 Hz
90% RANGE	112.6 224.7 Hz	126.4 180.3 Hz	123.8 185.3 Hz

Subject 38, Dx and Cx plots for Speech
a) first recording b) last recording

Figure 100

and his voice was modal although still extremely harsh and whispery in quality and produced with massive laryngeal tension (Appendix 5, Table 1).

Subject 37 had extremely low Fx values on both occasions and used a very creaky voice quality (Fig. 98 and 99). He was also smoking heavily.

Increased ranges are noted on the last assessment in untreated subjects 23, 31, 32 and 36, all had good voice qualitites and none smoked (Fig. 98).

Figure 99 shows regularity measures corresponding to the fundamental frequency values in Fig. 98 and reveal marked improvement in vocal fold regularity in subjects 23, 37 and 38. In subject 38 and to some extent in 37, this is related to a marked decrease in Fx range used. This indicates that voices with an abnormally narrow range may show a very high regularity of vocal fold vibration (%TS) within that narrow range of frequencies. The effect on %TS of increased range is seen in subjects, who show a very slight decrease in %TS between occasions 26, 27, 32 and 36, in all except subject 27 it is accompanied by a slight increase in range. Such slight reduction in %TS may therefore not neccessarily mean a deterioration in voice quality but be an artefact of the way in which %TS is calculated.

Table 54 shows the average of all the parameters illustrated in Fig. 95-99 for both groups of subjects on each assessment occasion.

The Median MPRx for the groups are used in Table 54 as the range of times after radiotherapy is very wide (Table 52 and 53). There is no major difference between the groups' MPRx Pre voice therapy or between the first and last occasion.

Comparing the Pre voice therapy values with the First assessment measures of the untreated group there seem to be slightly higher Mean and Range Fx values in the therapy group but lower regularity, %TS, than in the untreated group. After completed therapy there is an

increase in the Mean and Maximum range Fx measures and also in %TS as observed in the graphs in Fig. 96 and 97.

Table 54

GROUP AVERAGES

OBJECTIVE VOICE MEASURES - SPEECH SAMPLES

GROUP	**MEDIAN MPRx**	**2nd order Mean (Hz)**	**90% Minimum Range (Hz)**	**90% Maximum Range (Hz)**	**%TS**
With Voice Therapy					
Pre (N=9)	9.0	113.4	87.7	163.6	16.2
Post (N=8)	12.5	118.6	88.0	183.1	21.5
Last (N=9)	50.0	120.5	87.0	187.8	21.9
No Voice Therapy (N=8)					
First	5.5	102.3	78.6	155.5	19.9
Last	54.5	108.6	85.6	159.7	25.7

Comparing the Pre voice therapy values with the First assessment measures of the untreated group there seem to be slightly higher Mean and Range Fx values in the therapy group but lower regularity, %TS, than in the untreated group. After completed therapy there is an increase in the Mean and Maximum range Fx measures and also in %TS as observed in the graphs in Fig. 96 and 97.

On the 'Last' occasion, for the Voice therapy group the average Fx Mean is marginally higher still, the average Minimum range Fx remaining very similar on all occasions for this group. The Maximum range has increased.

Average %TS remains roughly the same on the Last occasion as Post voice therapy (Fig. 97).

The untreated group seem to have increased their Fx Mean and Range measures slightly (Fig. 98), mostly influenced by Subject 26, whose range has been raised overall and subject 38, whose extremely deviant values on the First occasion have been replaced by extremely high Mean and Range on the Last one. There are, however, some very deviant voices represented in the Voice Therapy group also, e.g. subjects 14 and 39.

The untreated group also show an improvement in their average %TS from the first to the last occasion, the likely result of three subjects' dramatic increase in %TS, subjects 23, 37 and 38. The other subjects in this group do not show marked change in %TS in any direction (Fig. 99), whereas the Voice therapy group displays a variety of patterns including a steady increase from Pre to post therapy (Fig. 97) in subjects 16, 21, 22, 25, and 30 and a decline in subjects 14,33, and 40. On the Last occasion subject 14 shows a further decline, subjects 16 and 22 a further improvement. In the subjects whose %TS decrease from the post therapy to the Last occasion none decrease to the Pre therapy levels.

In summary the data seems to indicate a positive effect of voice therapy on some irradiated subjects' ability to achieve and maintain a slightly higher Fx Mean and wider range and also to improve the regularity of their voices compared to a group of subjects who did not receive any instruction or training in voice production. It may contribute evidence of Lehman's et al's suggestion that therapy may help some irradiated speakers to make better use of the damaged vocal instrument.

viii Trends in perceptual and self rated voice quality measures after voice therapy.

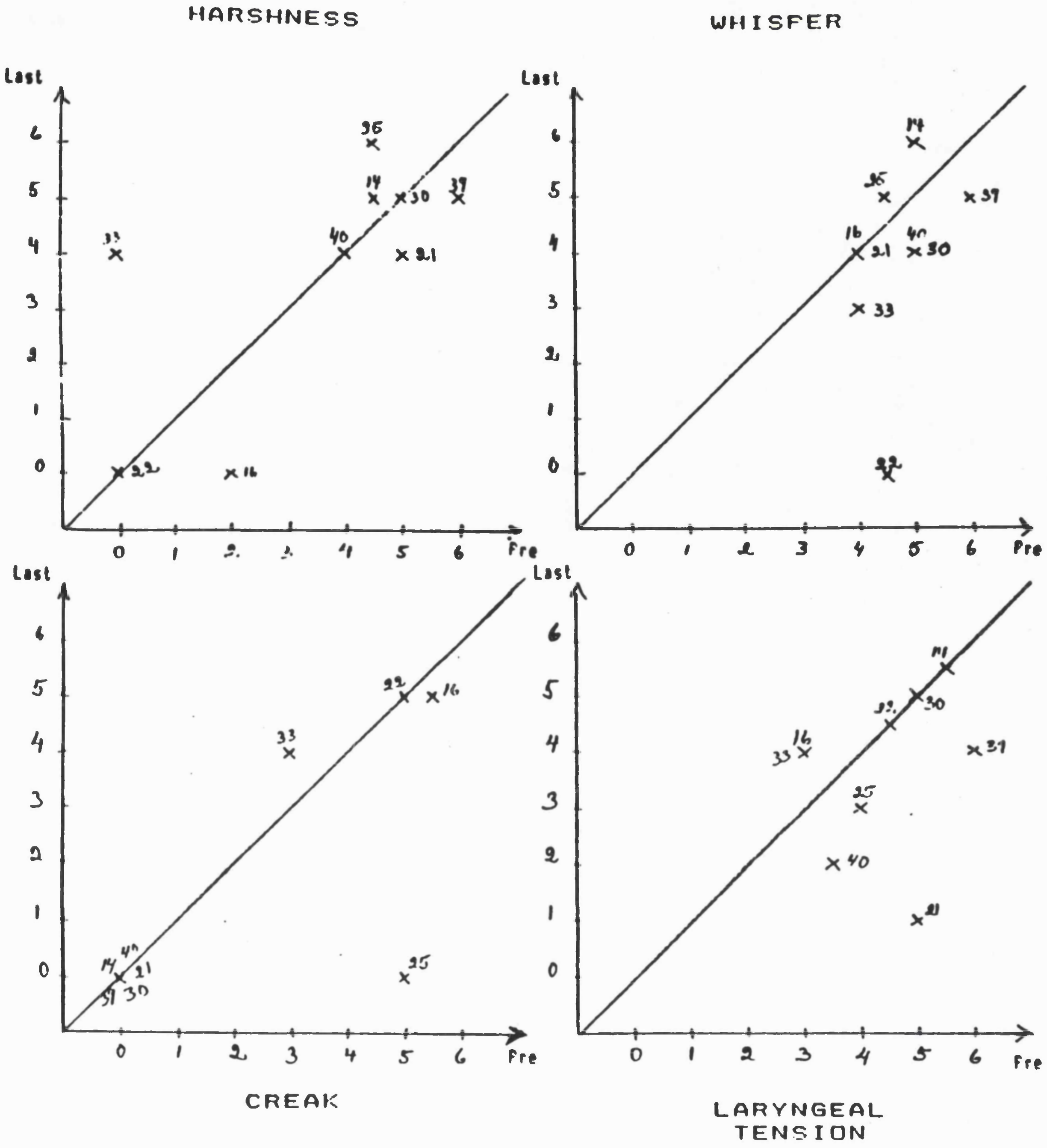
Having examined the objective voice measurements of treated and untreated subjects, perceptual ratings on the VPA, and the subjects' self-ratings on the questionnaire will now be examined for similar trends. Table 55 shows group averages of VPA ratings on the three occasions. As not all the subjects show all the voice quality features the numbers that the averages are based on are given in the table (N=). Figures 101 and 102 show the Pre-voice therapy versus the 'Last', and 'First' and 'Last' assessments for individual subjects in the two groups respectively.

Table 55

VPA GROUP AVERAGES - SPEECH SAMPLES

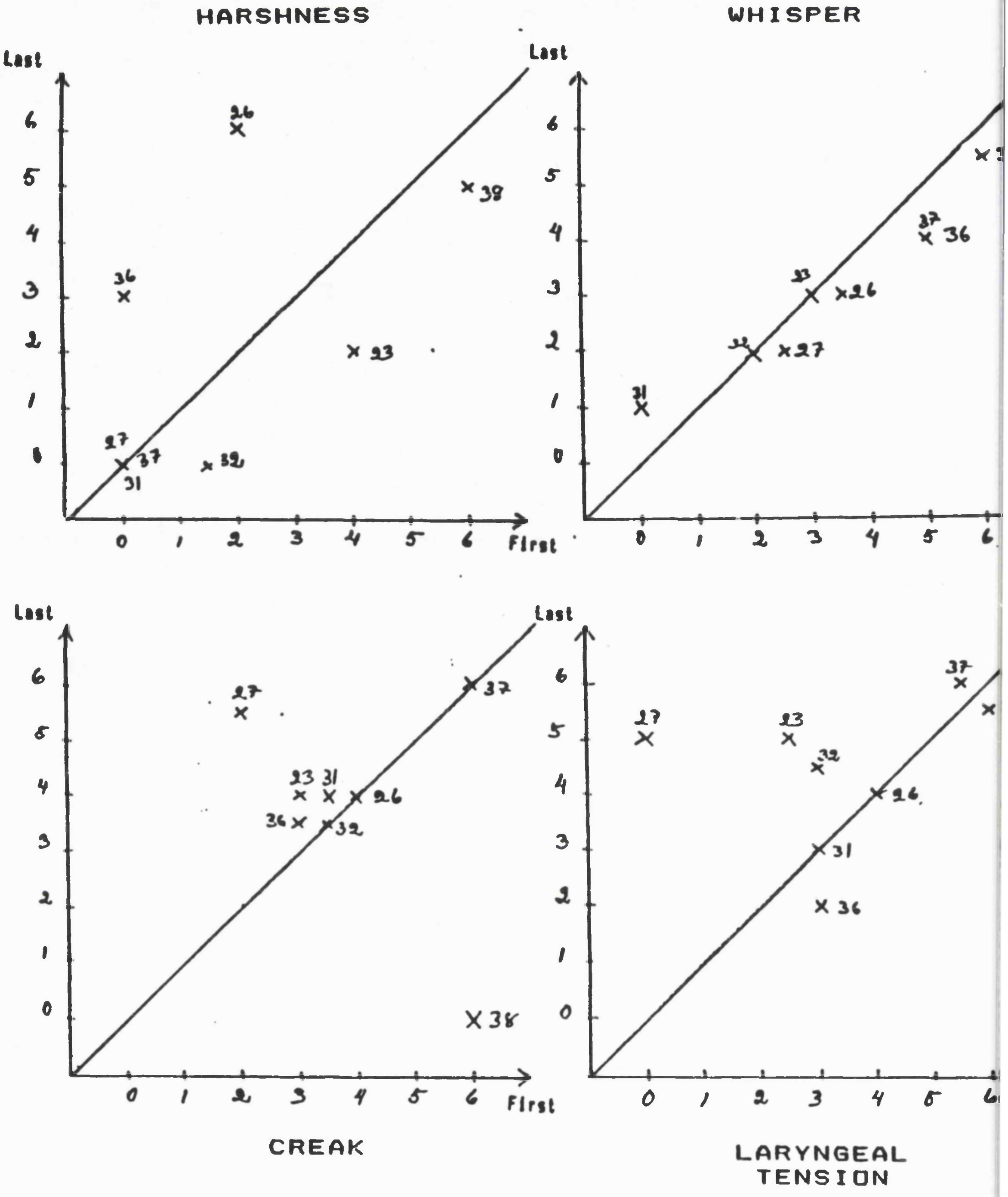
GROUP	MEDIAN MPRx	HARSHNESS	WHISPER	CREAK	LARYNGEAL TENSION
With Voice Therapy					
Pre (N=9)	9.0	4.4 (N=7)	4.7 (N=9)	4.6 (N=4)	4.4 (N=9)
Post (N=8)	12.5	3.4 (N=7)	3.8 (N=8)	3.4 (N=5)	3.2 (N=8)
Last (N=9)	50.0	4.7 (N=6)	4.1 (N=7)	4.7 (N=3)	3.4 (N=8)
No Voice Therapy (N=8)					
First	5.5	3.4 (N=4)	3.8 (N=7)	3.8 (N=8)	3.9 (N=7)
Last	54.5	4.3 (N=4)	3.1 (N=8)	4.2 (N=7)	4.4 (N=8)

Comparison of average Pre-therapy and 'First' group ratings (Table 55) seems to indicate that perceptually the subjects referred for and/or requesting voice therapy had poorer voice quality than the untreated group. This is despite these subjects being assessed slightly later than the untreated group (Median MPRx). Treatment seems to have had an immediate beneficial effect on all the perceptual ratings (Post) but on the 'Last' assessment occasion the only feature that stays within 'Normal' limits, i.e. below a rating of '4', is Laryngeal Tension. This is interesting, as it is a feature that would have been particularly targeted for reduction in therapy. The untreated group also show some but less marked deterioration in ratings between the First and the Last occasion and Laryngeal Tension retains a rating within 'Abnormal' limits, if we use the VPA form terminology (Fig. 10, p.83).



SUBJECTS WHO RECEIVED POST RADIOTHERAPY VOICE THERAPY
Vocal Profile Analysis ratings

Figure 101



SUBJECTS WHO DID NOT RECEIVE POST RADIOTHERAPY VOICE THERAPY
Vocal Profile Analysis ratings

Figure 102

These differences in voice quality between the groups are illustrated in that more Voice therapy subjects (Fig. 102) show higher ratings of 'Whisper' and 'Harshness' than the untreated group on both occasions (Fig. 102). 'Creak', however, seems to be a strong feature of the untreated group's voice quality also reflected in the overall lower fundamental frequency Mean and Range found in this group (Table 54, Fig. 98) compared to the Voice therapy group, where six of nine subjects get a rating of '0 Creak' or 'None perceived' on the 'Last' occasion (Fig. 101). The average rating of 'Creak' in Table 55 looks high for the Voice therapy group only on account of three subjects who still have this as a prominent feature.

High and increased 'Laryngeal Tension' ratings on the 'Last' occasion also seem a feature of the untreated group, whereas several of the Voice therapy subjects show a reduction (Fig. 101 and 102) although some have increased tension ratings compared to the immediate Post therapy occasion (Subjects 16, 22 and 25).

An obvious criticism and limitation of these VPA ratings is that the therapist who treated the subjects has also carried out the ratings before and after therapy and on the 'Last' occasion. She may be biased in favour of the treated subjects. However, the deterioration in their ratings on the 'Last' occasion may be some indication that this knowledge is not unduly influencing ratings.

Average group self ratings on the Questionnaire (Appendix 1b) on the three occasions are shown in Table 56. Individual Ratings Pre-therapy and 'Last', and 'First' and 'Last' are illustrated in Fig. 103 and 104.

Table 56 illustrates some overall reduction in Voice therapy subjects' ratings immediately after therapy with a slight increase in ratings, notably in 'Voice tiring after a lot of use' to almost Pre therapy levels on the 'Last' occasion.

Table 56

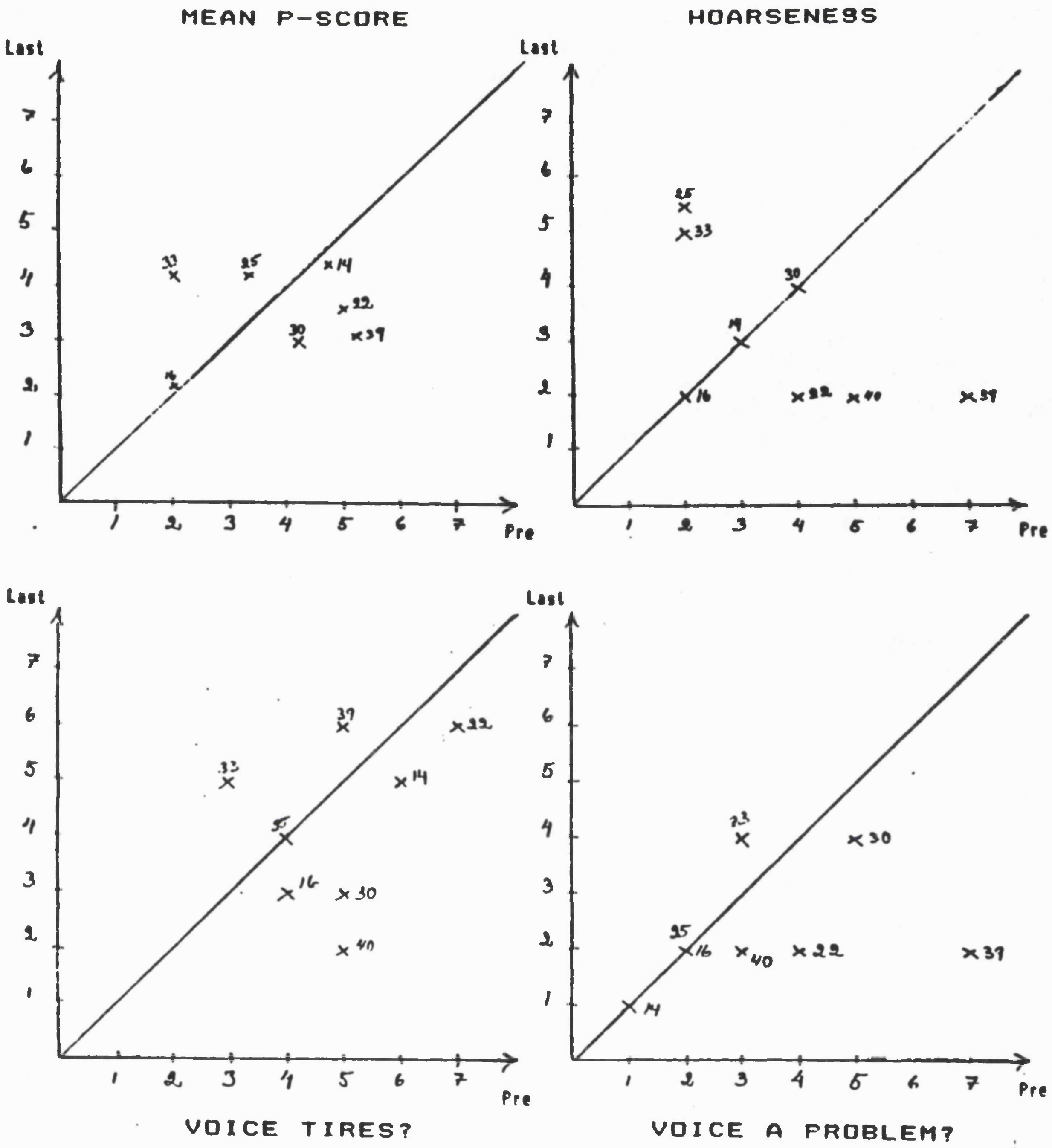
SELF RATINGS ON THE QUESTIONNAIRE, GROUP AVERAGES

GROUP	MEDIAN NPRx	MEAN P-score	DEGREE OF HOARSENESS	VOICE TIRES?	VOICE A PROBLEM?
With Voice Therapy					
Pre (N=9)	9.0	3.7	3.4	4.7	3.2
Post (N=8)	12.5	3.3	3.1	3.8	2.6
Last (N=8)	50.0	3.5	3.2	4.3	2.4
No Voice Therapy (N=7)					
First	5.5	2.4	2.4	2.4	1.1
Last	54.5	2.4	3.1	2.4	1.4

Just as the VPA ratings seemed to indicate poorer voice quality in the Voice therapy subjects, their Self ratings seem to show higher i.e. worse, ratings of 'Hoarseness', 'Voice tiring' and Mean P-score and consequently greater experience of the voice as a Problem compared to the untreated group (Fig. 103) where subjects return remarkably low ratings on both the 'First' and 'Last' occasion (Fig. 104).

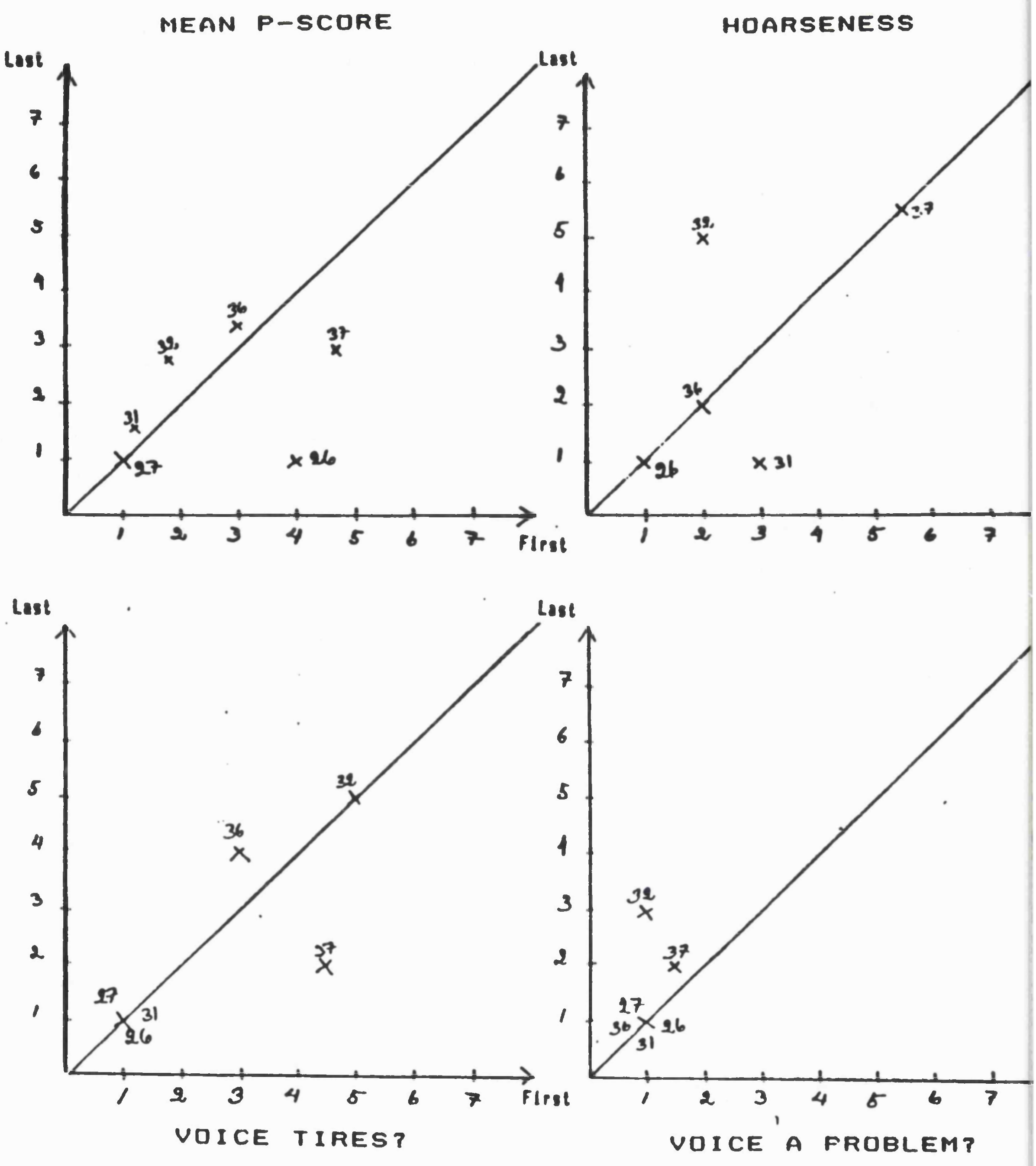
Voice therapy subjects seem to have accepted the limitations in voice function by the 'Last' occasion and return even lower 'Problem' ratings than on the Post therapy occasion. Their higher pitch and more regular voices produced with less Creak and Laryngeal tension may have contributed to this.

Four subjects' paired ratings are missing due to non-return of questionnaires on one occasion; Subject 21 on the 'Last' occasion, Subject 39 - Post therapy, Subject 23 on the 'First' occasion and subject 38 on the 'Last' occasion.



SUBJECTS WHO RECEIVED POST RADIOTHERAPY VOICE THERAPY
Self ratings on the Questionnaire

Figure 103



SUBJECTS WHO DID NOT RECEIVE POST RADIOTHERAPY VOICE THERAPY
Self ratings on the Questionnaire

Figure 104

SUMMARY

Table 57

		VOICE THERAPY	NO VOICE THERAPY
N		9	8
AGE ON LAST OCCASION	Mean	61.6	63.3
	Range	44 - 73	51 - 73
SMOKERS (Subj. No)		30 and 39	23. 37 and 38
WORKERS (Subj. No)		16, 21, 22, 25 and 40	31, 32 and 37

Table 57 reveals a discrepancy between the groups in the numbers of continuing smokers and, more strikingly, the greater number of workers in the group that received voice therapy compared to the one that did not.

An important factor influencing the worse Self ratings in the voice therapy group is likely to be the fact that this group contains five subjects who are still working and using their voices more than the higher proportion of non-working subjects in the untreated group. This may also explain the worse voice qualities as rated on the VPA in this group, possibly as a result of their greater voice use (cf Table 46 and 47, pp. 354-355) and vocal abuse post radiotherapy.

Subject 14 in the Voice therapy group has a very abnormal voice as a result of late radiation fibrosis; he is, however, retired and does not experience his voice limitations as a problem on either occasion (Fig. 103). However, subject 32 in the untreated group retired in the period between the assessment occasions but had started working part time again just before the 'Last' occasion. This may account for his increased self perception of Hoarseness and of the voice posing more of a Problem (Fig. 104).

Subject 37 in the untreated group was perceived as having very high degrees of Whisper, Creak and Laryngeal tension (Fig. 102). He perceived himself as very Hoarse but with surprisingly few vocal limitations and did not consider his voice a Problem (Fig. 104). His objective measures were also extremely deviant from the norm (Fig. 98 and 99). He was working throughout the period of radiotherapy and was offered Voice therapy, which he declined. He was smoking and drinking heavily.

In summary, the impression is of voice therapy being effective in improving voice quality in some irradiated subjects. There is some evidence for this in objective, perceptual and Self rated measures.

Regarding who will need or benefit from therapy, fundamental frequency (Fx) measures on their own cannot give an indication of this as both the treated and the untreated subjects show very low or deviant Fx measures compared to normals on the whole. The measures do, however, seem to show a possible effect of therapy in increased Fx Mean and Range and in greater regularity %TS (Table 54, Fig. 96 and 97) Post therapy.

What ultimately will determine whether a patient will co-operate and benefit from voice therapy is his own perception of whether he has a problem or not and to what extent he experiences limitations in vocal function. The people most likely to do so are those who go back to work after radiotherapy and who need to use their voices for work purposes, or who are self confessed keen 'communicators' in a general sense e.g subjects 30 and 33 in the therapy group.

We observe, however, the still quite high ratings of Harshness, Whisper and Laryngeal tension in some subjects and conclude, supported by Fx and %TS measures of conversational speech, that radiotherapy for early vocal cord tumours results in voice qualities that are not normal.

Voice therapy may help some patients to make better use of a damaged mechanism, particularly those who go back to work, but many will experience some permanent limitations in quality and function. The fact that many do not consider this a Problem is no doubt due to having been cured of a potentially fatal disease and the post radiotherapy voice quality being considerably better than the pre-treatment one.

CHAPTER XIV - Lx WAVEFORM ANALYSIS AFTER RADIOTHERAPY.

i Introduction

In this study Lx waveforms have been used for three main purposes:

Firstly, to monitor the strength and stability of the signal on which objective voice quality parameters, Fx and %TS calculations, are based. As described in chapter VII, i , the Laryngograph enables calculation of the interval (Tx) between successive vocal fold contacts (Fig. 43, p. 202), on which the TPS or PCLx program calculates and plots fundamental frequency Fx, parameters pertaining to a particular sample of speech or reading.

Secondly, Lx of sustained phonation has been used to keep a record of vocal fold vibratory characteristics at different times in the recovery process after radiotherapy. Irradiated subjects sustained the vowel [i] at habitual mid, high and low pitch at comfortable loudness (cf. Motta et al, 1990). This chapter will give examples of intra-subject comparisons of waveforms produced:

a) at different times after radiotherapy
b) after recurrence and before and after biopsy
c) after voice therapy
d) by subjects with persistent voice problems after radiotherapy.

A third use of Lx waveforms, as described in the previous chapter, was for visual feedback in voice therapy with irradiated subjects. Lx provides immediate feedback about the effect of changing laryngeal aerodynamics as a result of subjects learning to control and maintain subglottic pressure using breath support during phonation.

The study was designed to be part of the routine review procedure for patients irradiated at the hospital. The ENT surgeons' main concern was signs of recurrence of the tumour. Review of the subjects' notes, if the therapist had not been present at the examination, would often only state 'NSR', 'No sign of recurrence', without any further description of vocal fold status.

Several different ENT surgeons performed the numerous examinations of subjects described in this longitudinal study, mainly by indirect mirror laryngoscopy (IDL). In recent years more examinations have been performed using nasendoscopy and stroboscopy, which has given more detailed information about the effect of radiotherapy on laryngeal tissues of our subjects (Fig. 95, p. 384). As we have not had access to synchronised stroboscopy with Lx any attempt to relate Lx features to observed laryngeal abnormalities will be tentative.

ii Method.

To obtain the best possible Lx signal for individual subjects, the subject was seated in a swivel chair, half facing the examiner and was asked not to turn his head during the recording to avoid gross movement of the electrodes and interference with the signal. Subjects were encouraged to keep both feet on the floor i.e. not to sit crosslegged as this often leads to a slightly reclining posture and a lowering of the chin which changes the position of the electrodes relative to the vocal folds.

The examiner identified the notch in the thyroid cartilage, and once located, the vertical extent and level of the subject's larynx in the neck was examined. The round electrodes were then applied approximately in the middle of the thyroid alae and adjusted to where the strongest Lx signal was elicited with the subject phonating on [i]. This was monitored on an oscilloscope screen.

The electrodes were held in place by velcro fasteneres on the inside of an elastic neck band. The distance between them was adjusted according to the size and shape of the individual's larynx. As shown by Titze (1990) the angle between the electrodes and their distance from the glottis affects the strength of the EGG signal (Fig. 20 a, p. 121). This may affect the signal to noise ratio and any contact area measurements.

The first part of the voice assessment was a conversation between the examiner and the subjects. During this and during recording of the subsequent reading sample the examiner was monitoring the quality and strength of the Lx signal on the oscilloscope screen. Any indication that the signal was deteriorating by a loss of contact or movement of the electrodes was easy to remedy. It was evident that many subjects produce a lot of laryngeal movement during speech. This resulted in gross baseline movement of the Lx signal but seldom in any marked deterioration in the recorded signal, which was used for calculation of Fx parameters.

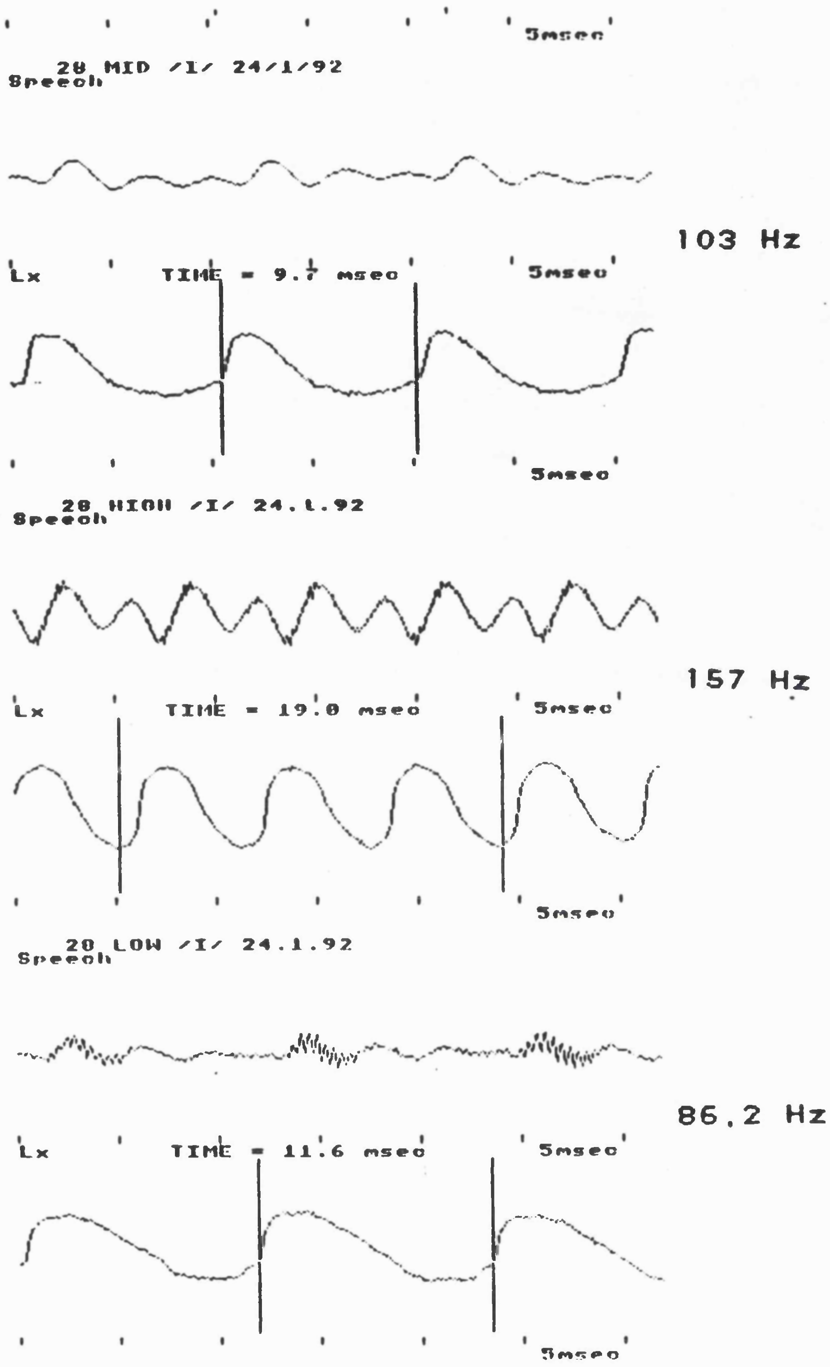
The last task the subjects were asked to perform was the sustained phonation on the vowel [i] at comfortable volume, at habitual high, middle and low pitch. The examiner demonstrated what was expected, and the subjects had an opportunity to practise and monitor their performance on the oscilloscope screen before the Lx waveforms were photographed or, more recently, recorded on DAT and stored for analysis on computer, at each pitch level.

iii Recording and evaluation of Lx waveforms.

Lx waveforms were originally photographed 'live' from the oscilloscope screen as shown in the Pilot study (Chapter VIII, Fig. 51, 54, 56, 58 and 59). Later, using the Lx ROM software, Lx and Speech waveforms could be stored and analysed e.g. by measuring the time span (Tx) between Lx cycles, on which fundamental frequency calculations could be made (Fig. 105). It also allowed a compariosn between features of the acoustic 'Speech' and Lx waveforms as described in the chapter on verification of ELG waveforms (Chapter V, v).

The most recent Laryngograph Processor and DAT recording equipment allows storage, playback, manipulation and analysis of Lx and Speech waveforms without phase distortion. Features of the waveforms can be identified and enlarged for detailed examination and measurements taken of durational aspects of the vibratory cycle (Fig. 24, p. 137).

Intensity of phonation has a direct bearing on the duration of vocal fold contact. Orlikoff (1991) found significant differences in his Contact Quotient (CQ) between soft, moderate and loud phonation. Our subjects were asked to sustain phonation at 'comfortable' loudness.



Subject 28, 34 MPRx;
Lx and Speech waveforms, analysed by Lx ROM.

Figure 105

During practice before recording, subjects were encouraged to sustain phonation at loud enough volume to produce an Lx signal, but intensity has never been measured or controlled in this study. Quantification of waveform features for intra-subject comparison between different occasions is therefore not attempted. Instead, qualitative judgements will be made of waveform features which have been observed in clinical practise. These will be related to laryngeal observations where available (Motta, Cesari et al, 1990, Colton and Conture, 1990, Carlson, 1993 a, b).

Due to the long duration of this study, during which facilities for analysing Lx waveforms have become more sophisticated as described above, polaroid photographs and Speech and Lx printouts using Lx ROM were used for the major part for evaluation of Lx waveform features. Waveforms were evaluated as proposed by Fourcin (1982), in terms of:

a) Uniformity of Lx peaks
b) Definition of Lx contact
c) Duration of closure
d) Regularity of contact periodicity

Attempts were also made to judge:

e) The rate of vocal fold closure by the steepness of the closing slope (Colton and Conture, 1990)
f) Irregularities in the 'contact-closing', maximum contact and 'Contact-opening' stages of the waveforms (Orlikoff, 1990, Motta et al, 1990)

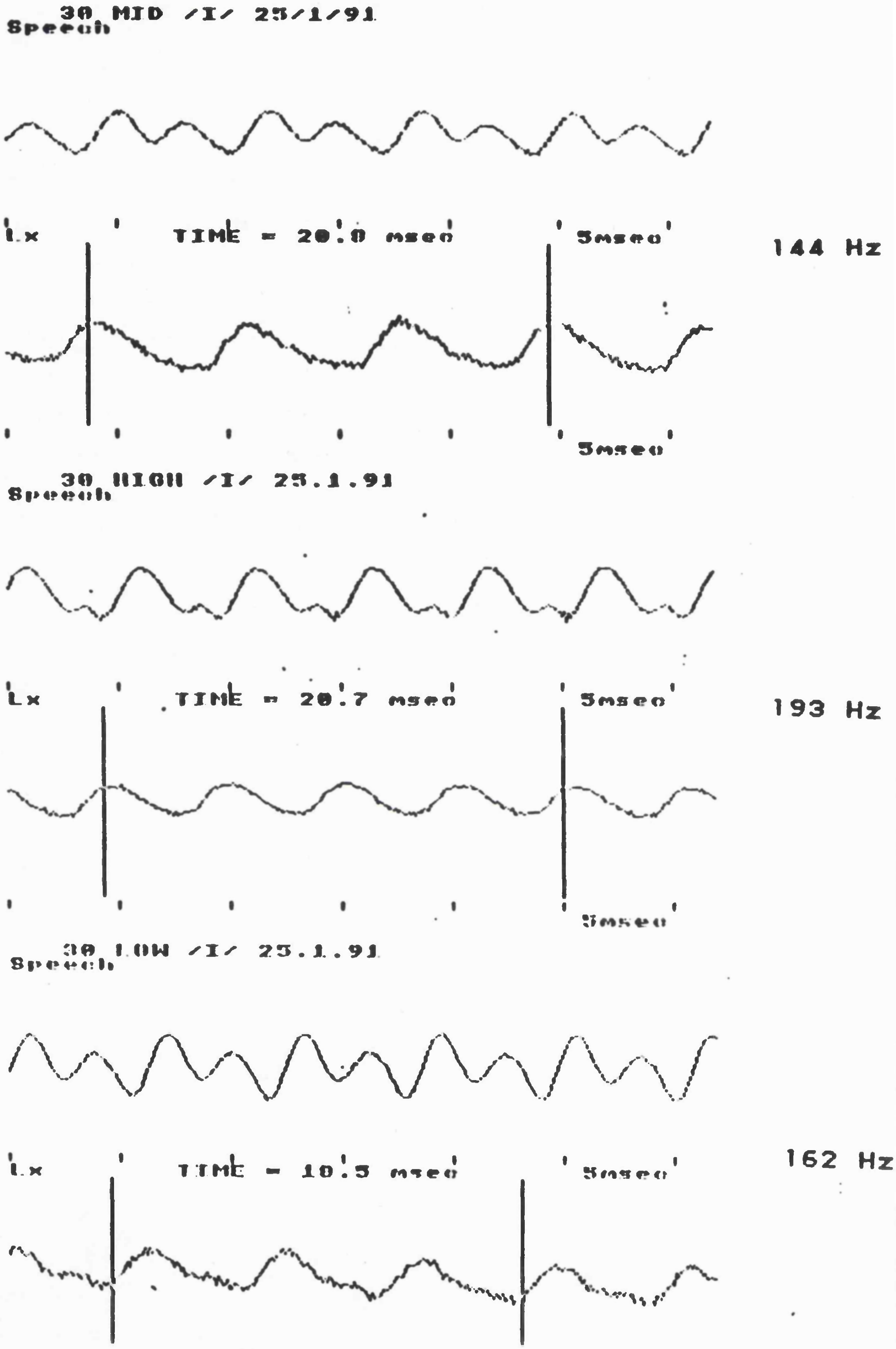
The 'idealised' Lx waveform phases in Fig. 43 (p. 202), were used to judge the relative duration of open and closed phases of the vibration cycle, keeping in mind, however, Colton and Conture's (1990) suggestion that it may not be appropriate to relate waveforms produced by pathological speakers to 'idealised' ones, and Motta, Cesari et al (1990), who found that 28 % of their patients with organic mass lesions such as vocal nodules produced near normal waveforms and so did 7 % of patients with polyps (Table 4, p. 159).

iv 'Non-analysable' waveforms.

One of Colton and Conture's (1990) issues regarding the use of EGG is described as 'subject concerns'. The most common of these is a poor Lx signal as a result of a thick layer of subcutaneous fat overlying the thyroid cartilage, more common in women and children than in men. Some male larynges are, however, positioned so low in the neck that access to the thyroid alae for securing the electrodes is difficult, resulting in a poor signal especially for sustained phonation e.g subject 30 (Fig. 106). An adequate signal without much 'noise' was however, produced on the same occasion during conversation as evidenced by good Dx and Cx plots (Fig. 107).

The Lx waveform in Fig. 106 would not be adequate for deriving quantitative information regarding phases of the vibratory cycle. However, during recording of natural speech and reading Lx looked more normal although with a very low amplitude. This is a reminder that laryngeal gesture and aerodynamics during sustained phonation on a vowel and during speech are likely to be different. This is the reason for collecting data from both, if possible. Failure to obtain acceptable Lx waveforms during sustained phonation need not preclude the use of ELG for recording and analysing running speech.

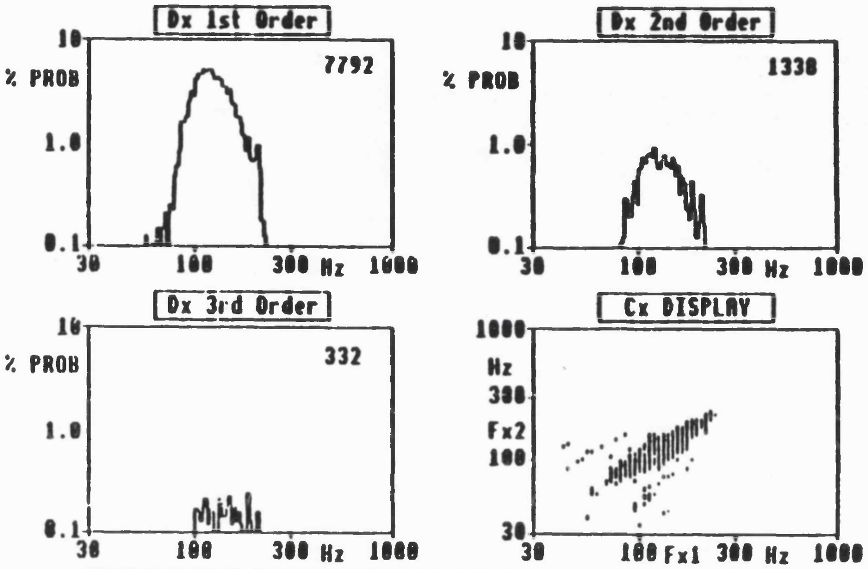
In all, 464 Lx waveforms were recorded, produced by 33 subjects with irradiated T1 (N=22) and T2 (N=11) tumours of the vocal folds on approximately 150 recording occasions before, during and after the end of radiotherapy. As described above, subjects were asked to sustain phonation on [i] at comfortable volume and at high, middle and low pitch. Occasionally subjects were unable to produce a recognizable Lx waveform at a certain pitch or all pitches before or during radiotherapy, possibly as a result of interference by the tumour with vocal fold vibration and/or the effect of acute tissue reactions to the treatment. Once treatment was completed their waveforms, as will be shown, improved with time. The proportion of Lx waveforms where it was impossible to identify phases of the 'idealized' Lx, for each pitch attempted are shown in Table 58.



Subject 30, 16 MPRx
Example of 'Non-analysable' Lx waveforms.

Figure 106

COMPACT DISPLAY Produced by the Dx Processing System (Dx)



TITLE : 30 SPEECH 25.1.91 FILENAME : :1.P.30SP191

STATISTICS TABLE

TITLE : 30 SPEECH 25.1.91 FILENAME ::1.P.30SP191

DISTRIBUTION TYPE	Dx 1st Order	Dx 2nd Order	Dx 3rd Order
SAMPLE TOTAL	7792	1338 STS=17.2	332
MEAN	122.9 Hz	129.9 Hz	137.2 Hz
MODE	113.2 Hz	122.9 Hz	190.6 Hz
MEDIAN	122.9 Hz	129.9 Hz	141 Hz
STANDARD DEVIATION	0.11 LOG-Hz	0.10 LOG-Hz	0.10 LOG-Hz
80% RANGE	96.1 170.0 Hz	98.6 185.3 Hz	104.3 195.0 Hz
90% RANGE	88.5 190.6 Hz	90.9 201.3 Hz	93.5 212.7 Hz

Subject 30, 16 MPRx
Speech, Dx, Cx and statistics.

Figure 107

Table 58

Proportion of 'Non analysable' Lx waveforms.

Phonation on [i]		**Pitch aimed for**			
		High	**Middle**	**Low**	**Total no of waveforms**
T1 (N=22)	**Total no waveforms**	104	104	104	312
	Not analysable Lx	24	16	20	60
	% of Lx	23.0	15.4	19.2	19.2
T2 (N=11)	**Total no waveforms**	50	50	50	150
	Not analysable Lx	9	6	8	23
	% of Lx	18.8	12.0	17.0	15.3

Table 58 above reveals a tendency for subjects to have more difficulty phonating at high or low pitch than at habitual 'middle' pitch, possibly because they are less confident about sustaining and varying fundamental frequency and tend to make more laryngeal effort when attempting to sustain phonation at extremes of their range. The proportion of waveforms impossible to analyse is close to or a little higher than that reported by Colton and Conture (1990) of 15 %.

There were four subjects, numbers 12, 28, 30 and 39 from whom it was extremely difficult to obtain Lx signals on sustained phonation. Subjects 28 and 30 account for 20/60 of the 'Not analysable' waveforms in the T1 group; subject 12 for 12/23 'Not analysable' waveforms in the T2 group. Subject 39 did not contribute to any of the above waveforms as he was never able to sustain an Lx trace. All other subjects produced analysable waveforms on some occasions.

This illustrates, however that the vast majority of subjects were able to produce acceptable Lx traces on sustained phonation at least at some pitches, most of the time. Among 'analysable' Lx there were, however, also 20% which were classified as 'poorly defined' in terms of not showing a well defined moment of first contact and only just perceptible left 'tilt'. The reasons for this are likely to be related to both the experimenter's inexperience in the early days of applying the electrodes and photographing waveforms and to do with subjects' phonating at low intensity. Had quantification of waveforms been our aim, many of the 'poorly defined' ones would also have had to be classified as 'non-analysable'. Here we shall use them, however, to demonstrate improvement over time, or otherwise, in vocal fold function and control.

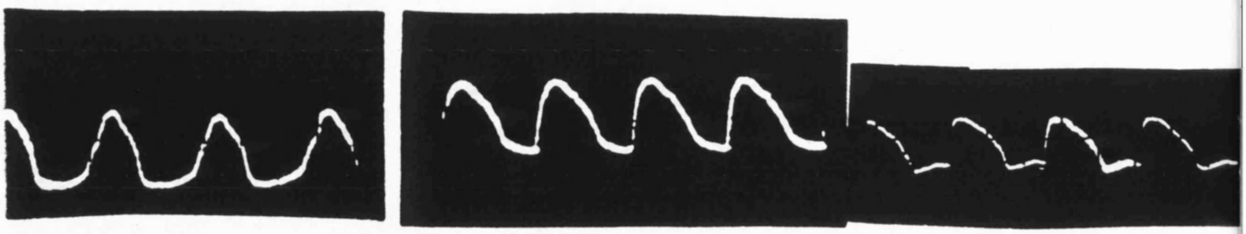
v Changes in Lx waveform shape after radiotherapy.

Fig 108 a illustrates gradual normalisation in Lx with time after the end of radiotherapy, in this case for a T2 tumour of the left vocal fold. Without detailed information about the appearance of the larynx, we can only hypothesise what may have produced the shape of Lx on the occasion, one month after the end of treatment (1 MPRx).

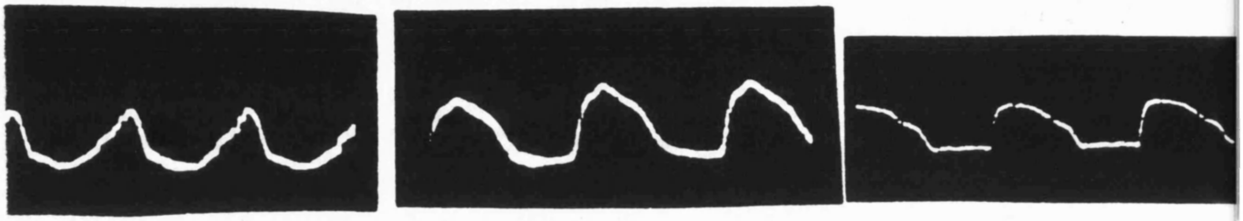
At high and mid pitch there is a delay in achieving maximum contact. At mid [i] there is evidence of gross irregularity towards the end of the contact-closing phase. A similar but less pronounced tendency is observed at high [i] on the same occasion. This may reflect tissue irregularity along the superior-anterior edge of the vocal fold, the last part of the folds to make contact on adduction, assuming it happens in an inferior-superior, posterior-anterior sequence as suggested by Childers and Krishnamurthy (1985) and Childers, Hicks, Moore and Alsaka (1986). This may be a possible explanation as the original lesion is described as being on the anterior left vocal fold with slight subglottic extension. Attempt at sustained low [i] produced a 'non-analysable' waveform with no distinctive contact features.

A month later, at 2 MPRx (Fig. 108 b) the waveforms are looking more normal. There is an increased left tilt at high and mid pitch. Long closure duration and good periodicity at high [i]. At mid pitch there is still evidence of some, but less irregularity at the end of the contact-closing phase and a delay in achieving maximum closure. This

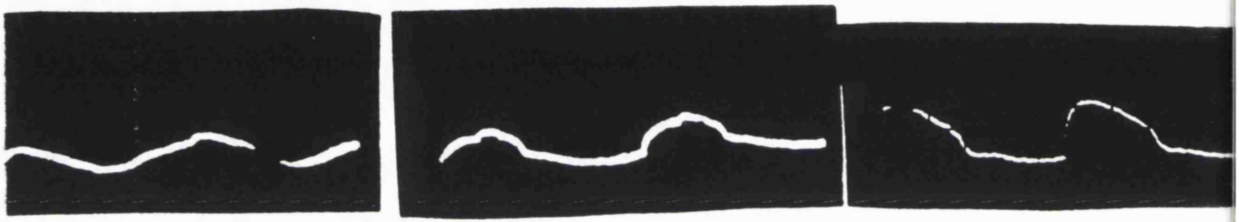
a) b) c)



High [i]



Mid [i]

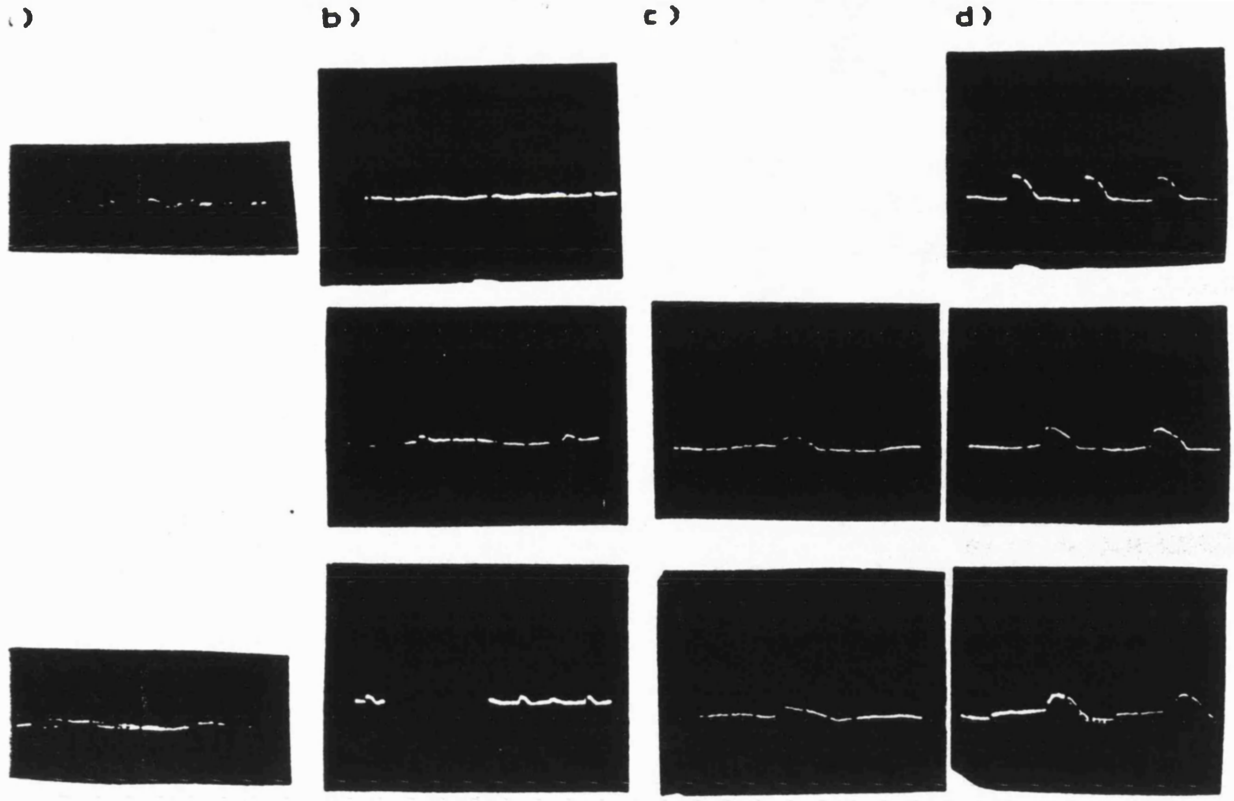


Low [i]

Subject 8, Lx

a) 1 MPRx b) 2 MPRx c) 5 MPRx

Figure 108



Subject 16, Lx
a) Pre, b) Mid, c) End d) 3 MPRx.

Figure 109

is much more pronounced at low [i] which this time shows some contact periodicity and a very long open phase.

Five months after the end of radiotherapy, Lx waveforms look entirely normal, with a short contact-closing phase and long duration contact resulting in a left tilt of the waveform, and good contact periodicity . Subject 8 had slightly 'Whispery' voice quality on this occasion as he was a hay fever sufferer. Indirect laryngoscopy showed 'good vocal fold movement'. Voice regularity, %TS, during reading and conversation illustrated in Appendix 9 A and B is almost 30 %.

Fig 109 illustrates the Lx waveforms of subject 16 who had a T1 tumour on his anterior right vocal fold, which was stripped for biopsy. Before radiotherapy (Fig. 109 a) there is little evidence of vocal fold contact periodicity, except possibly at attempted high pitch. Halfway through radiotherapy (Fig. 109 b), there are some brief moments of contact at mid and low pitch, slightly increased duration at the end of radiotherapy (Fig. 109 c) and three months later (Fig. 109 d) subject 16 produced recognizable Lx but with extremely long open phase at all pitches. Indirect laryngoscopy only commented on the right vocal fold looking 'red' post therapy. Unfortunately speech and reading recordings at 3 MPRx were not available so we have no corresponding Fx or %TS information. The recording two weeks post radiotherapy, however, shows an increased regularity of vibration, %TS, and a lowering of his speaking pitch (Appendix 8 A) His voice was very 'Harsh and 'Whispery' on this occasion (Appendix 5, Table 1).

vi <u>Lx waveforms after recurrence of vocal fold tumour and after biopsy.</u>

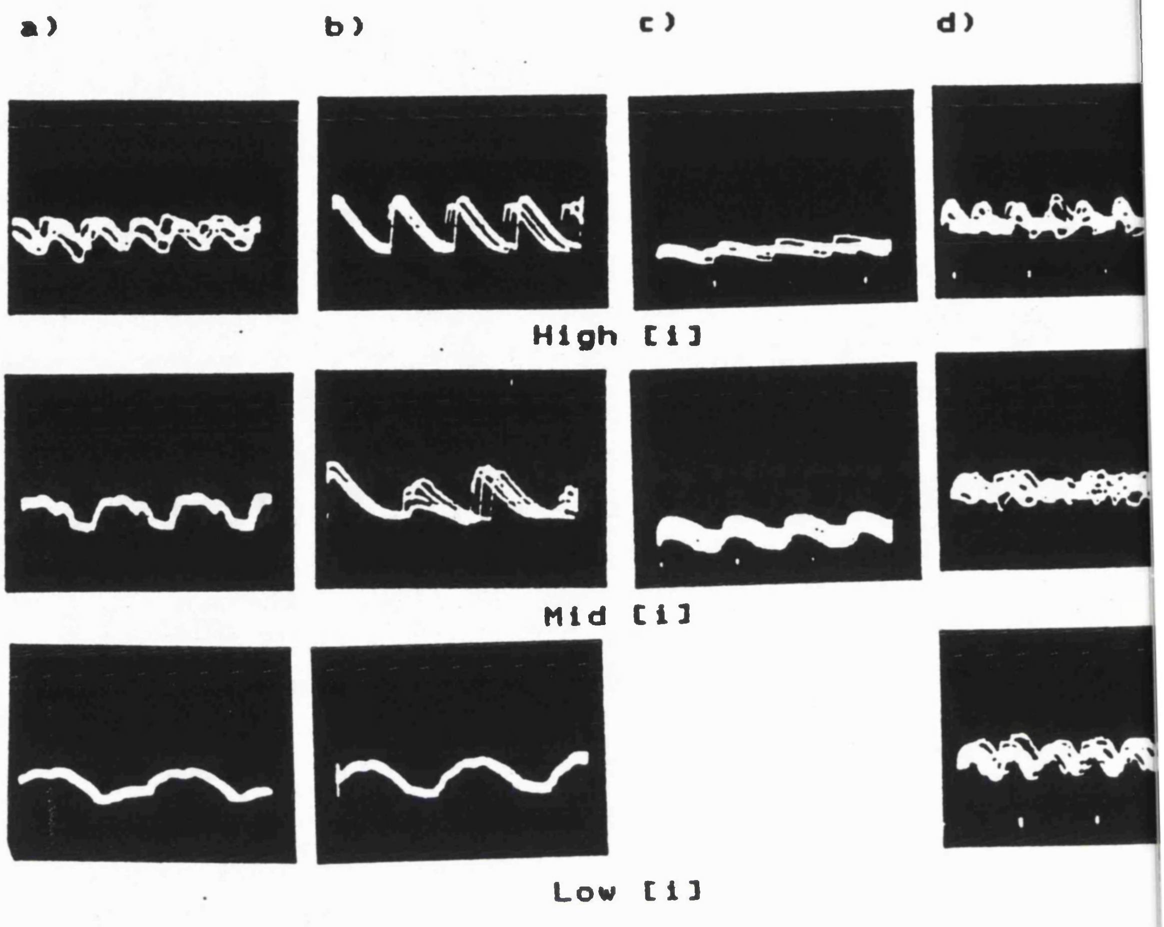
Figure 56 (p. 231), illustrated subject 3's waveforms at 6 and 8 MPRx the effect of recurrence of a T2 tumour of the anterior two thirds of the right vocal fold, which extended into the right ventricle. At 1 MPRx subject 3 spoke very softly. Sustained phonation at mid and low pitch resulted in a poorly defined but recognizable Lx waveform. The poorly defined contact, low amplitude, short contact duration and long open phase were the likely result of phonation at very low intensity.

Five months later (Fig. 56 b, p. 231)(6 MPRx) subject 3 complained of a deterioration in voice quality for two weeks. His voice was produced with effort in conversation. Sustained phonation, however, indicated ability to sustain good vocal fold contact periodicity at all pitches, but Lx waveforms of sustained [i] at high and mid pitch reveal both amplitude and frequency perturbation. This is possibly the result of increased stiffness caused by the recurring tumour in the vocal fold. Lx is better defined and there is longer duration of the contact phase than on the first occasion.

On the last assessment occasion at 8 MPRx (Fig. 56 c), phonation a high pitch reveals extremely abnormal and irregular contact pattern. Mid [i] has very limited tissue contact as evidenced by very low amplitude and short contact phase of the waveform. Low pitch is still well controlled with no evidence of interference. Fundamental frequency is higher than on the previous occasion, however. This increase in pitch is confirmed in a dramatic increase in Speech and Reading Fx Mean (Appendix 9 A and B). Subject 3 had a cordectomy two months later.

The effect of irradiation and recurrence of a T2 tumour of the Right vocal cord, involving the anterior commissure and extending into the subglottis is also seen in Fig 110 a. The patient was referred for voice therapy as his voice remained very husky almost one year after the end of radiotherapy (11 MPRx).

Indirect laryngoscopy showed evidence of approximation of the Left false vocal cord against the right true cord on phonation. This may be the explanation of the very abnormal Lx waveforms on this occasion, showing very delayed contact-closing phase at mid and low pitch and irregularity in the contact-opening phase (Fig. 110 a). After voice therapy Lx shows a more normal appearance at 12 MPRx on attempted high and mid pitch (Fig. 110 b), but the latter illustrates alternating long and short period variation. The abnormal waveforms



Subject 24, Lx

a) 11 MPRx b) 12 MPRx c) 22 MPRx d) 25 MPR

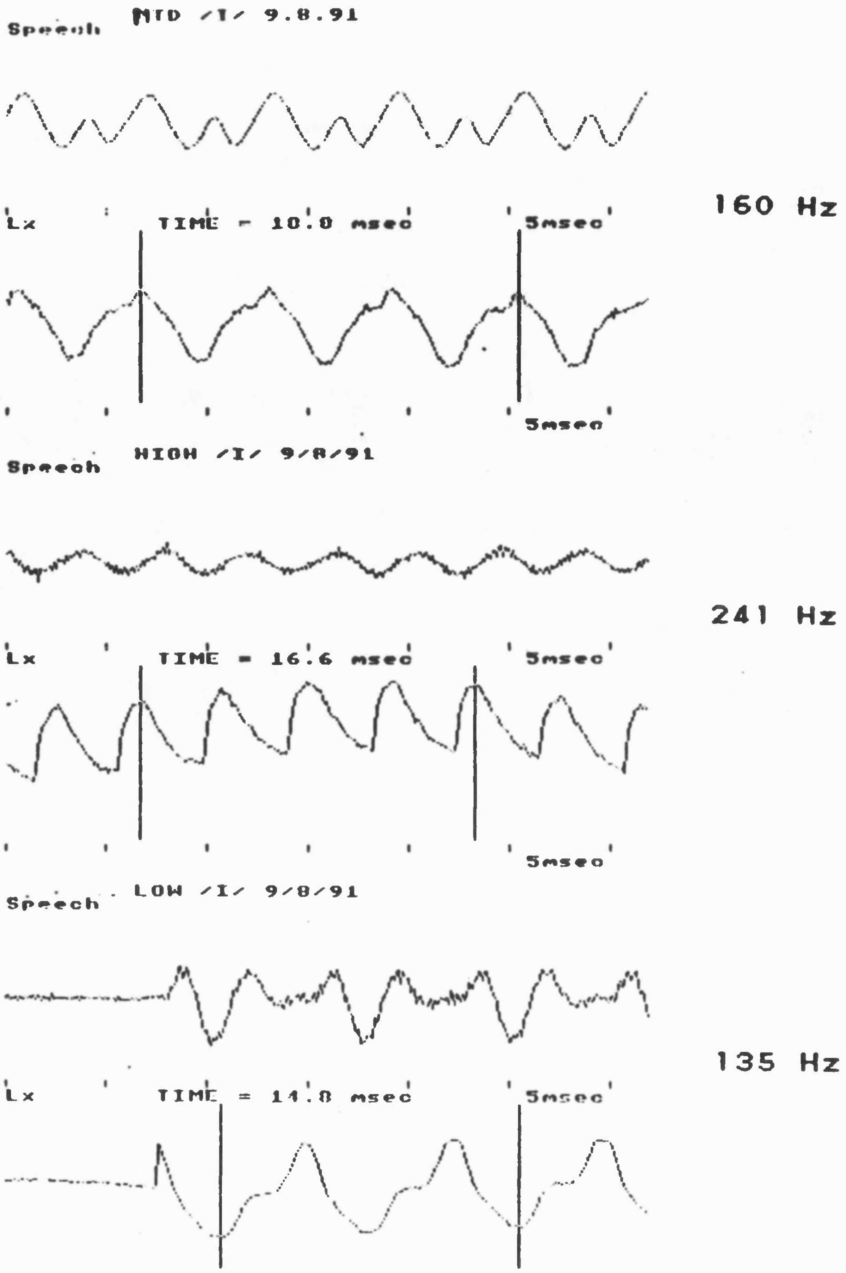
Figure 110

at low pitch may reflect approximation of the true and false vocal cords whereas the higher pitches seemed to result in, at least temporarily, an approximation of the true folds. Speech Fx and %TS on this occasion (Appendix 9 A) seems to indicate a slight increase in both measures.

On subsequent recording occasions at 22 and 25 MPRx (Fig. 110 c and d) subject 24's Lx, despite the poor polaroid quality due to oscilloscope trigger problems, shows increasing abnormality, reduction in the amount of contact and duration and extreme irregularity. At 44 MPRx (Fig. 111) recurrence has been confirmed and Lx shows abnormally long contact-closing phase and one may hypothesise the abnormal waveforms at mid and low pitch are the result of contact interference by the tumour. Subject 24 had a total laryngectomy five months later.

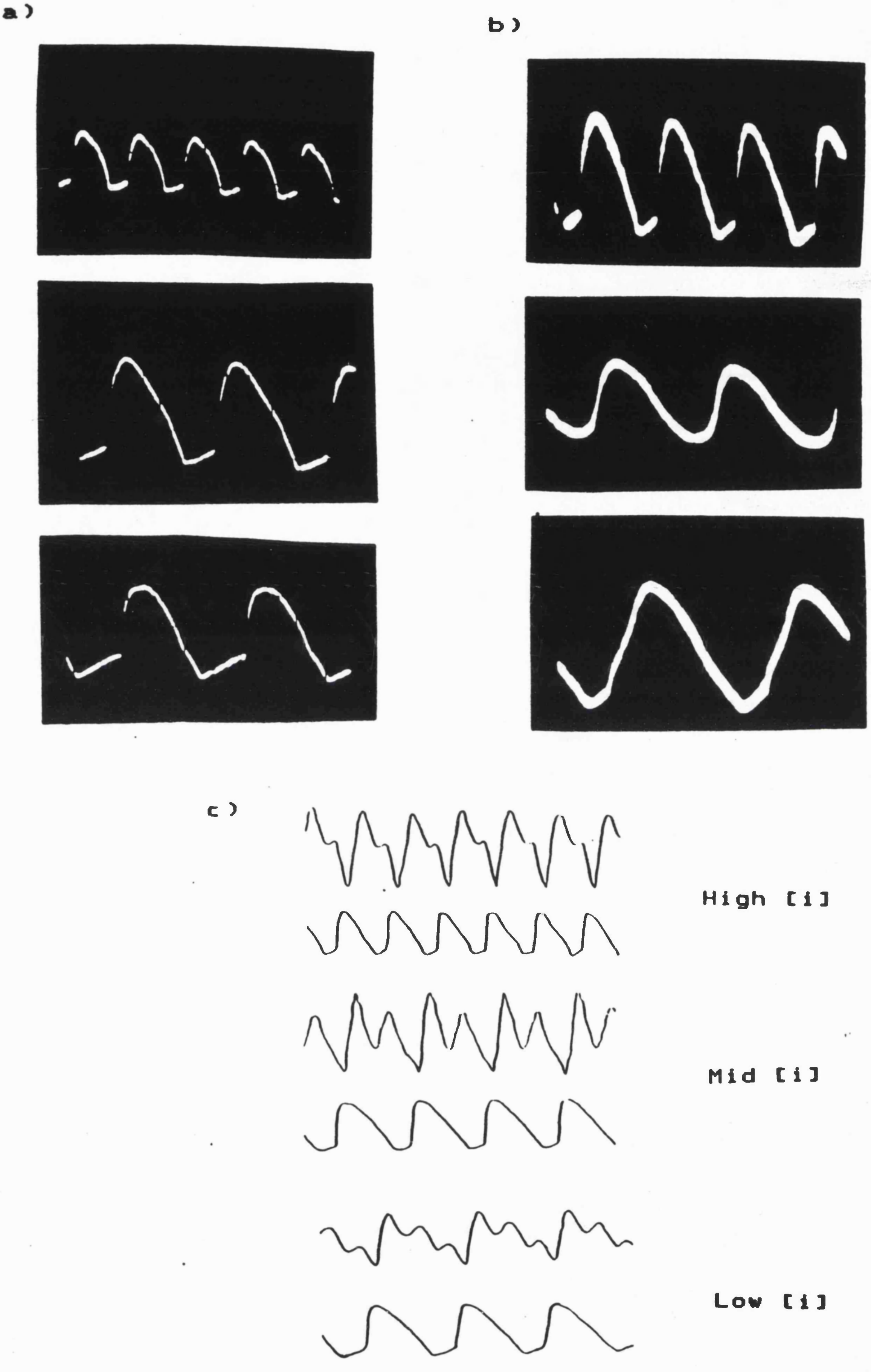
Figure 51 (p. 218), showed the effect on Lx of an undiagnosed tumour on the right vocal fold. Subject 20 had been referred for voice therapy in September after a biopsy in July had not revealed any malignancy of the vocal fold. His voice was extremely harsh and whispery only intermittently modal, and produced with massive laryngeal tension. Voice therapy using Lx for feedback resulted in some slight improvement in vocal function, but the Lx waveforms in Fig. 51 a showed such extreme limitation in the amount and duration of vocal fold contact, despite improved phonatory technique, that the patient was referred back to ENT for review.

Nasendoscopy revealed a large patch of leukoplakia on the anterior two thirds of the right vocal fold. Biopsy in October 1984 showed malignant changes consistent with a well to moderately differentiated T1 tumour. Lx waveforms in January 1985 (Fig. 51 b), just before the start of radiotherapy confirm the considerable improvement in voice quality. Lx waveforms, despite poor polaroid quality, show much improved amount, duration and quality contact periodicity. A course of radiotherapy in 3 fractions per week for one month was carried out in January - February 1985.



Subject 24, Lx 44 MPRx
after diagnosed recurrence.

Figure 111



Subject 20, Lx,

a) 9 MPRx b) 14 MPRx c) 92 MPRx.

Figure 112

For comparison purposes Figure 112 a and b, illustrates subject 20's Lx waveforms at 9 and 14 months post radiotherapy. He was using a rather low pitched voice and spoke very softly but with increasing control and regularity %TS (Appendix 8 A and B). He remained rather anxious and protective of his voice and reported considerable limitations in voice quality and function on the Questionnaire (Appendix 6 A) His voice quality remained Harsh and Whispery with increased laryngeal tension on the last assessment occasion (Appendix 5, Table 1).

On this occasion, more than seven years after the end of radiotherapy (92 MPRx), nasendoscopy showed hypertrophied false vocal cords almost obscuring the view of the true vocal folds, which were described as 'floppy'. Lx waveforms of sustained phonation on this occasion with corresponding Speech pressure waveforms (Fig. 112 c) show well defined contact periodicity of long duration and high amplitude. Speech waveforms illustrate the decay in the complex acoustic wave during the opening and open phase of the cycle as acoustic energy is absorbed in the subglottis.

vi Lx waveforms after radiotherapy and voice therapy.

Fig. 93 a and b, (p. 374), illustrated the waveforms produced by subject 22 at 5 and 12 MPRx. In the intervening period he had received eleven voice therapy sessions. His vocal folds on the first occasion were described as 'red' and his arytenoids as oedematous. Perceptually his voice was extremely creaky and produced with excessive laryngeal tension (Appendix 5, Table 1).

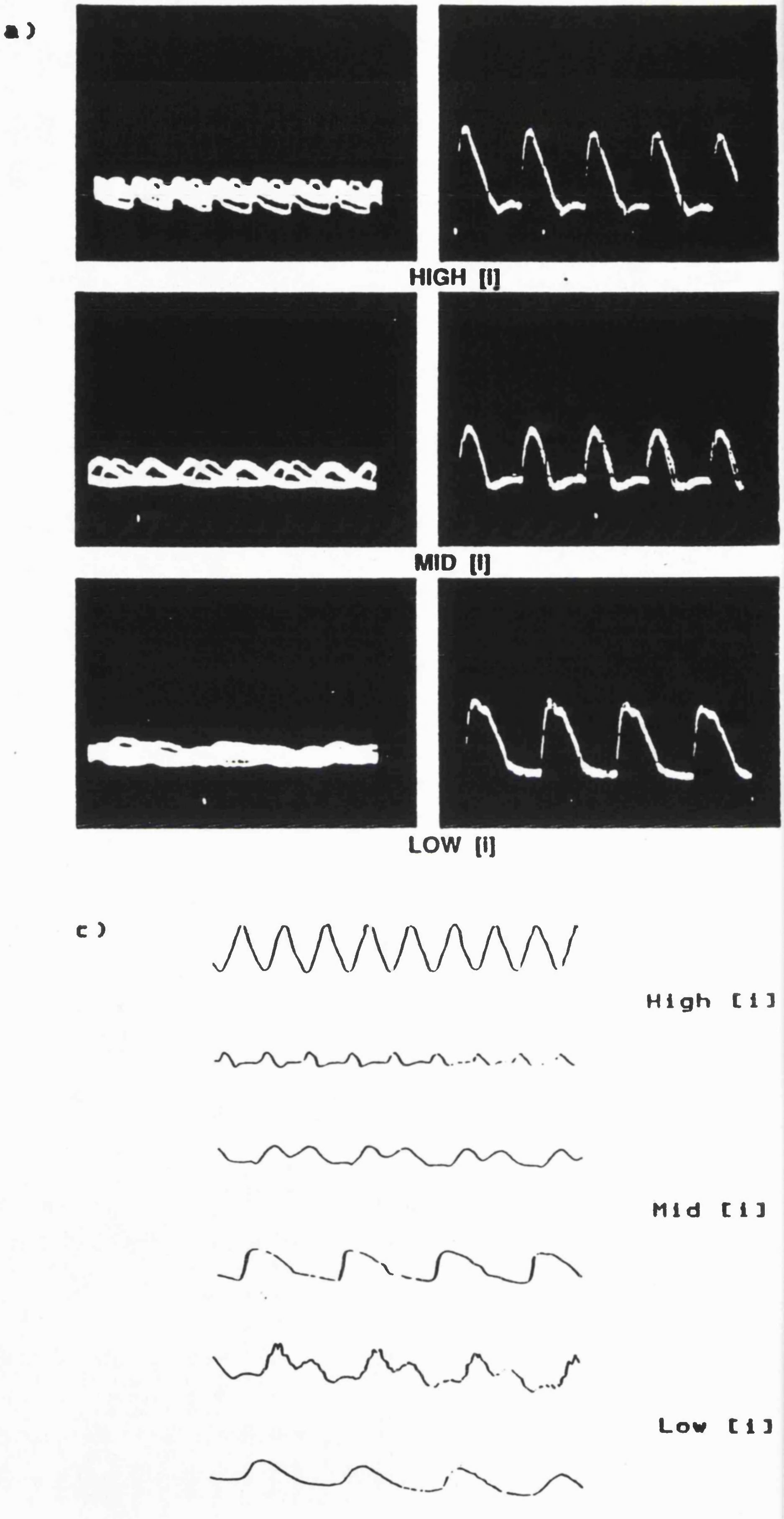
Lx waveforms at 5 MPRx illustrate the effect of such hyperfunctional voice production in the low amplitude, long contact phase and delay in achieving maximum contact at mid and low pitch. The waveforms could also be described in terms of what Titze (1990) calls 'skirt bulging' or 'knee', both in the closing and opening contact phase. He suggests this is the effect of medial surface bulging on vocal fold adduction, typical of male EGG. The waveforms also illustrate 'pulse widening' resulting from 'Pressed' phonation. Titze (1990) produced the latter feature on modelled EGG by decreasing the abduction quotient (Fig. 27 d, p. 143).

Therapy was aimed at reducing laryngeal effort and muscular compression of the vocal folds on phonation and to teach subject 22 to achieve a balance between subglottal pressure and laryngeal resistance using breath support as described in the previous chapter. The effect of such a change in phonatory technique is shown in Fig. 93 b, which illustrated increased speed of closure of the vocal folds, increased amplitude as a result of increased tissue contact between the vocal folds and no evidence of 'medial surface bulging'. If anything, there is evidence of a decrease in the duration of the contact phase on this occasion as subject 22 was overcompensating and producing a 'Whispery' voice quality in his effort to reduce laryngeal tension and creaky voice.

Titze (1990) explains increased 'Peak skewing' to the left as the result of increased 'convergence' and 'vertical phasing' on vocal fold adduction and abduction (Fig. 28 b, p. 144). Closure produces an abrupt increase in contact over the full thickness of the vocal folds, but there is a gradual release of contact on opening. This seems to fit in with the theory of what voice therapy using the Accent method achieves in terms of increased speed of vocal fold adduction through enhancement of the Bernoulli effect (Frøkjær-Jensen, 1983).

Examination of subject 22's vocal folds a few weeks later revealed no arytenoid oedema, the likely result of a reduction in hyperfunctional phonation and increased speaking pitch and regularity, %TS (Appendix 8 A and B).

Similar evidence of improved vocal fold function after voice therapy is shown in Fig. 113 b. This subject (16), described in greater detail in the previous chapter, requested help with his voice five years after completed radiotherapy for a T1 tumour of his right vocal fold.



Subject 16, Lx

a) Before voice therapy (61 MPRx)
b) After voice therapy (63 MPRx)
c) 92 MPRx.

Figure 113

Inspection of Appendix 8 A and B illustrates a remarkable stability of subject 16's Mean Fx and %TS from the recording only two weeks after the end of radiotherapy, to this occasion five years later (61 MPRx). Both parameters are stabilised at a very low level, however. Voice therapy between 61 and 63 MPRx has the effect of raising his fundamental frequency to the level of 'Normal average' and also of increasing his %TS. Recordings on later occasions reveal, despite some further fluctuation in both measures, that subject 16 has maintained the improvement over time.

Subject 16 was grossly overweight, with a very thick neck. His Lx waveforms at 61 MPRx (Fig. 113 a) were consequently thought to be the result of a poor signal. However, Lx was used for visual feedback in therapy and with improved control of subglottal pressure, he was able to produce much improved Lx showing good control and mode of vibration at 63 MPRx (Fig. 113 b) compared to the pre therapy occasion. Fig. 113 c confirms maintenance of this ability 2 ½ years later (93 MPRx). Examination of his larynx never showed any major evidence of abnormality except 'redness' of the irradiated vocal fold. However, in the stroboscopy examination on the last occasion, thinning of the irradiated cord and a reduced mucosal wave were observed, which may account for the long open phase and limited Lx amplitude in Fig. 113 c.

viii <u>Lx waveforms in subjects with persistent voice problems after radiotherapy.</u>

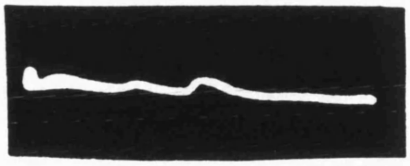
Figure 58 (p. 234), shows the Lx waveforms of the second subject (2) to be referred to the Pilot study. He had had radiotherapy 4½ years earlier for a poorly differentiated T1 tumour of his anterior left vocal cord, involving the anterior commissure. His voice had never returned to normal and an anterior glottic web was observed on indirect laryngoscopy on the occasion at 53 MPRx (Fig. 58 a, p. 234) when the first Lx waveforms were recorded. His voice was extremely abnormal in quality with high degrees of Harshness, Whisper, Creak and Laryngeal tension (Appendix 5, Table 1).

The most notable feature of the waveforms is the high fundamental frequency at all pitches, indicating poor control of Fx and very low amplitude indicating limited tissue contact on adduction. Both features are the likely result of the anterior glottic web tethering the cords and restricting the amount of contact and increasing stiffness resulting in higher than normal pitch.

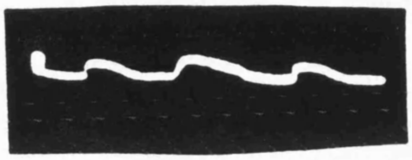
Appendix 8 A and B illustrating Speaking and Reading Mean Fx and %TS, show remarkably low but fluctuating %TS and extremely high Reading Fx in recordings between 53 and 112 MPRx.

On the third assessment occasion at 65 MPRx (Fig. 58 c), waveforms showed extremely erratic Lx at high pitch, very limited duration and amount of contact at mid [i] and almost normal Lx at low pitch. The latter feature gave the impression that a different vibratory pattern may be possible, and voice therapy aimed at teaching improved control of the damaged vocal instrument, was offered between the recording occasions at 65 and 70 MPRx (Fig. 58 d). Lx waveforms on the latter occasion confirm better control of pitch and vibratory mode, longer duration contact phase at high and low pitch, but a delay in achieving maximum contact.

Mean Fx and %TS (Appendix 8 A and B and Fig. 57, p. 233) show an increase in %TS, more dramatic in Reading than in Speech, a slight lowering of pitch in Speech and an increase in Reading at 70 MPRx. On subsequent recording occasions, however, there is evidence of further abnormal increase in Fx. Regularity, %TS, remains very low. Lx 18 months later (Fig. 114, 89 MPRx) indicates a deterioration in vocal fold vibratory control on sustianed phonation with evidence of both frequency and amplitude perturbation. Perceptual voice quality ratings confirm the extremely poor voice quality (Appendix 5, Table 1). The likely explanation is increased stiffness of laryngeal tissues as a result of late fibrosis and mucosal dryness coupled with the anterior glottic web, which is likely to have arisen due to the original lesion involving the anterior commissure.



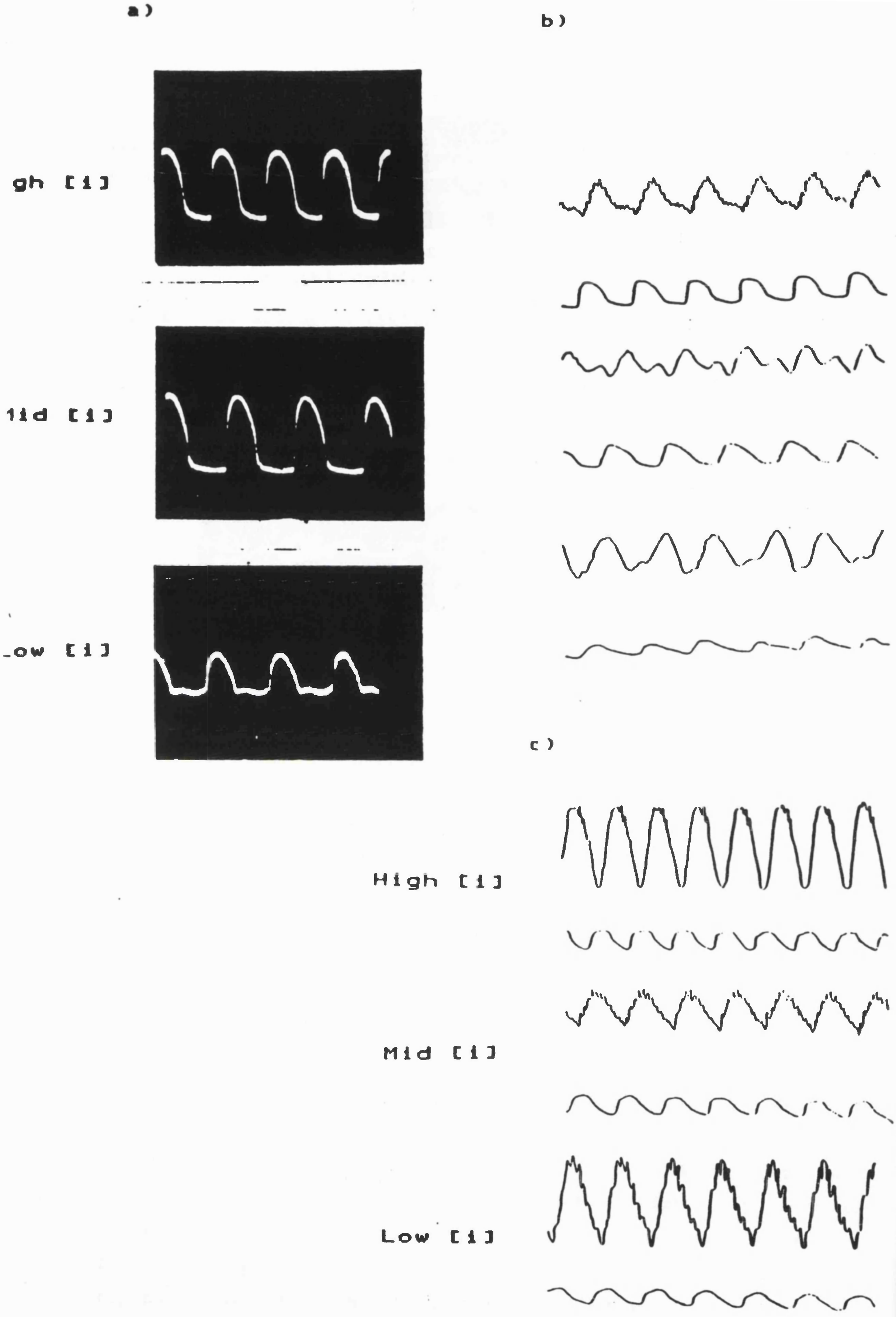
High [i]



Low [i]

Subject 2, Lx, 89 MPRx.

Figure 114



Subject 14, Lx

a) 120 MPRx b) 149 MPRx c) 156 MPRx

Figure 115

Another subject who developed late radiation fibrosis and whose voice remains a problem is subject 14. Lx waveforms ten years after radiotherapy (Fig. 115 a, 120 MPRx) are well defined, indicating rapid closing of the vocal folds and regular contact periodicity at all pitches. Opening seems quite abrupt, however, at mid and low [i]. This may be evidence of increased stiffness of the vocal fold 'cover' as suggested by Hirano, also giving rise to abnormally high speaking and reading Mean Fx (Appendix 8 A and B). His voice quality was perceived as very harsh and whispery and produced with excessive laryngeal tension (Appendix 5, Table 1).

Two years later (Fig. 115 b, 149 MPRx) Lx at mid pitch is not so clearly defined but regular. At low pitch there is evidence of alternating short and long periods and a delay in achieving maximum contact. The Speech pressure waveform (Sp) also shows aperiodicity reflecting subject 14's poor voice quality at low pitch. High pitch shows a long contact duration with abrupt opening.

Videostroboscopy a few months later showed good right vocal fold movement and mucosal wave but an atrophied left fold with no mucosal wave. The supralarynx looked stiff and the false vocal cords 'bulky'. Lx on this occasion show a long contact-closing phase, particularly at high and low pitch and long contact duration (Fig. 115 c, 156 MPRx). Fundamental frequency is overall very high. Relative Lx amplitude is very low indicating limited tissue contact a likely result of the lack of mucosal wave in the 'cover' of the irradiated vocal fold as suggested above and confirmed by stroboscopy. This has lead to further abnormal increase in Mean Fx and on subsequent recording occasions a dramatic decrease in %TS (Appendix 8 A and B). Voice therapy between the latter two occasions did not have any beneficial effect as described in the previous chapter XIII.

ix Conclusion.

This chapter illustrates the manner in which Lx waveforms can be used to reflect variation in and quality of vocal fold contact as a result of acute and late tissue reactions to radiotherapy, undiagnosed and recurring mass lesions, and to changing laryngeal behaviour as a result ov voice therapy.

Orlikoff's (1991) suggestion that his Contact Index (CI) (Fig. 35 b, p. 154) reflects vocal fold tonus and vertical fold dynamics may possibly be illustrated in the changes observed in the duration of the 'contact-closing' (ccp) and 'contact opening' (cop) phases with time after radiotherapy (Fig. 108), with recurrence of tumour (Fig. 56 and 110 a,b) and with the development of late fibrosis (Fig. 115).

Fig. 115 b and c showing increased duration of the Contact Phase (CP) in the same subject, may be the result of increased compression of the vocal folds in the horizontal plane, as suggested by Orlikoff, possibly as a response to increased fibrosis and the complete loss of mucosal wave motion in the irradiated cord observed on stroboscopy at 156 MPRx (Fig. 115 c). The waveforms resemble Titze's modelled EGG typical for 'Pressed' phonation with medial surface bulging and 'Pulse widening'. The other example of this and of change with voice therapy in mucosal and compression dynamics were illustrated in Fig. 93 a and b (p. 374).

Lx waveforms can not replace direct observation for diagnosis of laryngeal pathology, nor should waveforms produced by different individuals be compared to each other. However, despite variation in recording techniques and lack of strict control of loudness of phonation on different recording occasions in this study, it was still possible to detect improvement or deterioration in vocal fold function reflected in Lx, and direct examination confirmed abnormalities observed in the waveform (cf. Fig. 51, p. 218).

CHAPTER XV - DISCUSSION

This discussion will refer to conclusions drawn at the end of chapters and will be divided into three parts:

a) Issues concerning the viability and reliability of Electrolaryngography (ELG) as a tool for routine clinical voice assessment.

b) Issues concerning the particular group of subjects examined and the methodology used.

c) Lastly, suggestions for future research will be made.

i Electrolaryngography as a means of routine clinical voice assessment.

The findings in this study have confirmed the usefulness of Electrolaryngography for routine clinical assessment of voice patients and for indirect examination and monitoring of vocal fold vibratory function.

The instrumentation and software used produced good agreement and repeatablility between ELG based voice parameters provided care was taken to adjust and monitor the gain on the recording and on the Laryngograph processor to avoid overload on analysis. The vast majority of subjects produced strong enough Lx signals for Fx analyses to be carried out despite sustained phonation resulting in 'non-analysable' Lx waveforms for a few subjects at some pitches.

Second order Dx distributions were found to produce marginally better agreement and repeatability than 1st order, particularly for the irradiated subjects. The technique is entirely applicable and practical in a clinic situation, the great advantage being the non-invasiveness and imperviousness of Lx to ambient noise and the ability to record and analyse long stretches of natural speech. The instrumentation has become increasingly 'user friendly' over time.

Lx waveforms were shown to reflect improvement in vocal fold function with time after radiotherapy and with changing laryngeal aerodynamics as a result of voice therapy. Lx also showed evidence of abnormality or deterioration in subjects whose tumours were recurring or who had persistent voice problems.

Confirmation of the usefulness of Lx in monitoring progress in voice therapy and revealing suspect, persistent limitations in vocal fold function as a result of an organic lesion was obtained with the referral of subject 20 with a six year history of voice problems including removal of vocal nodules. His very abnormal looking waveforms (Fig. 51, p. 218) after a period of therapy resulted in referral back to the ENT surgeons and confirmation of malignant changes in his vocal fold mucosa after biopsy.

ii Longterm effects on voice quality after radiotherapy for early glottic tumours.

Problems with the methodology chosen for this longitudinal study have been mentioned in the introductions to the Pilot and Main studies and were due to:

a) A non-random sample of subjects whose only common denominator was the fact that they had received radiotherapy for early glottic tumours.

b) Small numbers of T1 and T2 subjects treated in different fractions

c) Difficulties in the timing of assessments

Considering the findings in normative studies of speaking fundamental frequency referred to in Chapter II, and the degree of variation in voice quality parameters found within and between speakers without laryngeal abnormality, performing different speaking and reading tasks and/or sustained phonation, the wisdom of a longitudinal study recording the same speakers' voices several times after radiotherapy, spanning several years may be questioned.

Statistical analysis of the total data would not have been appropriate due to the above mentioned methodological difficulties. Therefore graphic plots were used to give a general overview of trends over time in the Mean Fx and %TS data (Appendix 3, 8, 9 and 10 A and B), self-rated voice quality features (Appendix 6 A, B and C) and tables of the examiner's ratings of auditory perceptual features using the Vocal Profile Analysis protocol on each recording occasion (Appendix 5, Tables 1 and 2). Some individual subjects were then followed and described in terms of the above features.

The difficulties reported in the literature for raters to reach agreement using perceptual voice evaluation systems (Kreiman, 1993) were confirmed in Chapter IX and there is still a need to find and develop objective measurements, which will reduce the need for, or at least corroborate, subjective rating scales. This has been an aim of studies reporting acoustic correlates of hoarseness and other descriptive voice quality terms (Chapter III, iii) although most conclude that for a complete description of voice quality several objective measures will be needed.

The most useful objective measure used in this study was %TS, which incorporates features of fundamental frequency perturbation and was significantly lower in irradiated speakers compared to normals. So was also the case with the more recently developed % Irregularity measure (Chapter X). %TS correlated significantly with both clinicians' perceptual and subjects' self-rated voice quality features (Chapter XI).

Fx and %TS plots in the group of Pilot subjects (Appendix 3 A and B) seemed to demonstrate a trend of a reduction in Fundamental frequency (Fx) during the first six months after the end of radiotherapy and then a stabilisation of Fx and an increase in regularity of vocal fold vibration (%TS), except in subjects whose tumours recurred (Subj.3) or had persistent voice problems (Subjects 2, 13 and 14).

These trends were confirmed to some extent in the Main study subjects, most of whom had treatment in 5F/week, however. They showed less strong trend towards an increase and greater variability in regularity (%TS) over time (Appendix 8-10, A and B).

Among the Pilot subjects, subject 10 had received voice therapy between the recording occasions at 5 and 7 MPRx from which he benefitted as evidenced in both perceptual, self-rated and objective voice quality measures (Table 11, Appendix 3 A and B). He complained on the latter occasion of feeling depressed and tired and was found to suffer from myxoedema for which he received treatment before the next occasion at 10 MPRx.

In the Main study, subject 40 had also been found to have myxoedema, which had been diagnosed six months after the end of radiotherapy. He was receiving thyroxine treatment for this when he was referred for voice therapy at 9 MPRx, as his voice was still not satisfactory.

Subjects 22 and 25 were referred for voice therapy due to poor voice recovery at 5 and 9 MPRx respectively. There is no evidence that they had their thyroid function tested but evidence of extremely low pitch and %TS when they were first recorded (Appendix 8 and 10, A and B). After voice therapy, in common with subject 10, there seems to be an increase and stabilisation in both parameters for both subjects. Subject 20 also shows a slight reduction in Fx and %TS between 4 and 6 MPRx after which both show a steady increase and follow closely the pattern of subject 22.

Other subjects who seem to show this pattern of a reduction in fundamantal frequency and regularity around 4-6 MPRx followed by an increase and stabilisation are T1 (5F/week) subjects 26, 31 and 37 and T2 subjects 8, 29, 33 and 35.

A female patient referred recently for voice therapy after irradiation of a T2 tumour reported feeling extremely tired and depressed in a therapy session five months after the end of radiotherapy. She had also put on weight. She was found to be hypothyroid.

One of the symptoms of Hypothyroidism is a lowering of the pitch of the voice. It is not inconceivable that irradiation of the larynx for early glottic carcinoma using radiation fields of between 5*5 and 6*7 cm^2 may include the thyroid gland, particularly in people with short necks and may affect the production of thyroxine. As mentioned above, subjects 10 and 40 were diagnosed as having myxoedema, having felt unwell for some time before diagnosis at 6 MPRx. Other subjects showing the pattern of lowering of Fx and %TS until 4-6 MPRx and a later increase may have had mild hypothyroidism of which a low pitched voice was the main symptom but not recognized, as a low pitched voice and tiredness may be expected after radiotherapy.

Patients about to undergo total laryngectomy are routinely referred pre-operatively to the Speech and Language therapist to get to know the therapist, who is going to assist him in vocal rehabilitiation after the operation and for the therapist to explain how the operation will affect his ability to communicate, swallow, taste and smell.

Similarly, it seems reasonable to assume that patients with lesions likely to be cured by radiotherapy but to result in vocal limitations, particularly towards the end and immediataly after treatment, are likely to benefit from expert advice and guidance on how to manage vocal symptoms and limitiations before the commencement of radiotherapy and advise on how to avoid secondary vocal abuse.

Psychological support and counselling, where needed, is part of the normal speech and language therapy process with voice patients. In this study it was found that many subjects were quite unconcerned about their vocal limitations after radiotherapy, and did not consider these a problem. However, explanations of laryngeal structure and function, of expected symptoms and suggestions how to manage these were always received with interest and relief by all subjects and allowed them to ask questions and be reassured by the therapist at any time before during and after radiotherapy. Many subjects referred at different times after the end of radiotherapy, particularly to the Pilot project but also to the Main study, expressed a wish that such information and advice had been available early in their management and treatment.

Fallowfield (1990) suggests:

> *"The fact that psychological distress, which profoundly affects quality of life, is still so apparent in cancer patients who have an extremely good prospect of complete cure may seem counterintuitive, but demonstrates the need for good counselling support, irrespective of cancer site."*

In the case of patients who are about to have or have had radiotherapy for T1 and T2 laryngeal carcinoma, which will inevitably affect their ability to communicate for a period during and after treatment, the most obvious professional to take on the counselling role may be considered to be the Speech and Language Therapist. The present study shows evidence that direct therapy for alleviation of vocal symptoms and laryngeal discomfort can also have longterm, measurable beneficial effects on voice quality.

iii Future research.

Recent developments and means of corroboration, identification and quantification of phases of vocal fold contact reflected in the Lx waveform by means of modeling experiments and synchronised EGG and stroboscopy has greatly advanced the clinical usefulness of this technique for phonatory assessment.

Orlikoff's (1991) well documented evidence of significant CQ differences between different loudness levels indicate that future research and clinical use of ELG for assessment and diagnostic purposes should incorporate means of recording the intensity of phonation. This would improve intrasubject Lx waveform comparisons at different times after intervention. Thereby also, allowing quantification of phase relationships and calculation of for instance CQ and CI as demonstrated by Orlikoff.

Such developments are already under way and will, when more widely available, improve diagnosis and clinicians' ability to reliably measure and monitor the effects of both surgical intervention and vocal rehabilitation on natural speech phonation without resort to invasive procedures. These are neccessary for identifying organic causes of vocal malfunction but interfere with habitual laryngeal gesture during articulated speech.

A recent report by Cannon (1994) provides an extensive clinical and experimental study of hypothyroidism in Head and Neck cancer patients and a historical overview of such reports in the literature from 1961 to 1991. He reports an incidence of hypothyroidism after irradiation for head and neck cancer of 15% and recommends thyroid function testing before and after treatment, irradiation and/or surgery, of all such patients particularly those who develop a second primary tumour in whom the incidence of hypothyroidism is higher still.

Biel and Maisel (1985) report an incidence of 38 % hypothyroidism in a sample of 216 laryngeal cancer patients after radiotherapy alone.

An explanation of the apparent delay in effect on pitch and regularity of vocal fold vibration between 4-6 months after the end of treatment may be suggested by Rubin and Cassarett (1968). They postulate a *'biphasic mode'* of thyroid injury as a result of radiotherapy; the first occurring a few hours up to a few days after treatment injuring the *'endothelium of the nutrient vessels'*. Subsequent doses of radiation, suggest Rubin and Cassarett, damage the residual vasculature giving rise to secondary degeneration of the follicular epithelium resulting in decreased thyroid function.

An interesting future study would be a closer examination of the relationship between thyroxine levels and voice fundmental frequency and regularity with time post radiotherapy.

The development of stroboscopy triggered by Lx enables detailed observation of the vocal folds for signs of oedema, which is usually the cause of lower pitch in Hypothyroidism. The combination of oedema with the inevitable well documented reduction in mucosal wave over the irradiated vocal fold would explain why such subjects may be observed as having poor voice quality and complain of more vocal problems than those who do not develop signs of myxoedema. Treatment with thyroxine would be initiated with improvement both in patients' general well being and voice quality.

REFERENCES.

ABBERTON, E. and FOURCIN, A. (1972). Laryngographic analysis and intonation. British Journal of Disorders of Communication 7, 24-29.

ABBERTON, E. (1976). A Laryngographic study of voice quality. Ph.D. thesis, University of London.

ABBERTON, E. and FOURCIN, A. (1978). Intonation and speaker identification. Language and Speech 21, 305-317.

ABBERTON, E. and FOURCIN, A.J. (1984). Electrolaryngography in Clinical Phonetics. Code, C and Ball, M. Eds. Croom Helm, London.

ABBERTON, E.A., HOWARD, D.M. and FOURCIN, A.J. (1989). Laryngographic assessment of normal voice: a tutorial. Clinical Linguistics and Phonetics. Vol. 3, No. 3, 281-296.

ALTMAN, Douglas, G. (1991). Some common problems in medical research. In Practical Statistics for Medical Research. Publ. Chapman and Hall.

AMERICAN JOINT COMMITTEE FOR CANCER STAGING AND END RESULTS REPORTING. Manual for staging of Cancer, 1978. AJC, Chicago, 39-41.

ANASTAPLO, S. KARNELL, M.P.(1988) Synchronized videostroboscopic and electroglottographic examination of glottal opening. Journal of the Acoustic Society of America, 83. 1883-90.

ANDREWS, S., WARNER, J. and STEWART, R. (1986). EMG biofeedback and relaxation in the treatment of hyperfunctional dysphonia. British Journal of Disorders of Communication, 21, No.3; 353-369

ARONSON, A.E.(1980). Clinical Voice Disorders. Thieme Stratton Inc. New York.

ASKENFELT, A. and HAMMARBERG, B. (1980). Speech waveform perturbation analysis. Speech Transmission Laboratory - Quarterly Progress and Status Report, 4, 40-49.

ASKENFELT, A. GAUFFIN, J. SUNDBERG, J. KITZING, P. (1980). A comparison of contact microphone and electroglottograph for the measurement of vocal fundamental frequency. Journal of Speech and Hearing Research, 23, 258-273.

ASKENFELT, A. and HAMMARBERG, B. (1986). Speech waveform perturbation analysis: A perceptual acoustical compariosn of seven measures. Journal of Speech and Hearing Research 29, 50-64.

ATKINSON, J. (1976). Inter- and intraspeaker variation in fundamental voice frequency. Journal of the Acoustic Society of America 60, 440-445.

BAER, T. (1979). Vocal jitter. A neuromuscular explanation. In: Transcripts of the eighth symposium: Care of the Professional Voice Part II, Lawrence, V. ed. New York: The Voice Foundation, pp. 19-22.

BAER, T., LOFQUIST, A. and Mc GARR, N.S. (1983 a). Laryngeal vibrations: a comparison between high-speed filming and glottographic techniques. Journal of the Acoustic Society of America 73 (4), 1304.

BAER, T., TITZE, I. and YOSHIOKA, H. (1983 b). Multiple simultaneous measures of vocal fold activity. In Vocal Fold Physiology: Contemporary Research and Clinical Issues. Bless, D.M. and Abbs, J.H. Eds. College Hill Press, San Diego, 229.

BAKEN, R.J. (1987). Clinical Measurement of Speech and Voice. Publ. Taylor and Francis Ltd. London.

BALL, V., FAULKNER, A. and FOURCIN, A. (1990). The effects of two different speech-coding strategies on voice fundamental frequency control in deafened adults. British Journal of Audiology. 24, 393-409.

BARRY, W.J., GRICE, M., HAZAN, V. and FOURCIN, A.J. (1989). Excitation distributions for synthesised speech. In: Tubach, J.P. and

Mariani, J.J. Eurospeech 89, European Conference on Speech Communication and Technology, Paris. Volume 1; 353-356.

BARRY, W.J., GOLDSMITH, M., FOURCIN, A.J. and FULLER, H. (1990 a). Larynx Analyses of Normative Reference Data. ALVEY Project MMI/132, Speech Technology Assessment Final Report, Part I. August.

BARRY, W.J., GOLDSMITH, M., FOURCIN, A.J. and FULLER, H. (1990 b). Stability of Laryngeal Measures in Speech. ALVEY Project MMI/132, Speech Technology Assessment Final Report, Part 2, August.

BARRY, W.J., GOLDSMITH, M., FOURCIN, A.J. and FULLER, H. (1991). Stability of voice frequency measures in Speech. Proc. 12th Internal Congress of Phonetic Sciences, University of Provence. Volume 2, pp. 38-41.

BASSICH, C.J. and LUDLOW, C.L. (1986). The use of perceptual methods for assessing voice quality. Journal of Speech and Hearing Disorders; 51: 125-33.

BECKETT, R.L. (1969). Pitch perturbation as a function of subjective vocal constriction. Folia Phoniatrica, 21; pp.416-425.

BECKFORD, N.S., MAYO, R., WILKINSON, A. and TIERNEY, M. (1990). Effects of short-term endotracheal intubation on Vocal function. Laryngoscope, Volume 100, No. 4, 331-6.

BENNINGER, M.S., GILLEN, J., THIEME, P., JACOBSON, B. and DRAGOVICH, J. (1994). Factors associated with recurrence and voice quality following radiation therapy for T1 and T2 Glottic Carcinomas. Laryngoscope, 104, March; 294-298

BERG van den, J. (1968). Mechanism of the larynx and the laryngeal vibrations. In; Malmberg, B. ed. Manual of Phonetics. North-Holland, London.

BERKE, G.S. and GERRATT, B.R. (1993). Laryngeal Biomechanics: An overview of mucosal wave mechanics. Journal of Voice, Volume 7, No. 2, pp. 123-128.

BIEL, M.A. and MAISEL, R.H. (1985) Indications for performing hemithyroidectomy for tumours requiring total laryngectomy. Archives of Otolaryngology, Head and Neck Surgery. 150: 435-439.

BIEVER, D. and BLESS, D. (1989). Vibratory characteristics of the vocal folds in young adults and geriatric women. Journal of Voice, 3; pp. 120-131.

BLAND, M.J. and ALTMAN, D.G. (1986). Statistical methods for assessing agreement between two methods of clinical measurement. The Lancet, February 8, 307-310.

BOTNIK, L.E., ROSE, C.M., GOLDBERG, I. and RECHT, A. (1984). The role of radiation therapy in the treatment of laryngeal cancer. In: Otolaryngologic Clinics of North America, Volume 17, No. 1, February.

BRADLEY, K. (1985). An acoustic analysis of voice quality following radiotherapy for laryngeal cancer. Unpublished Research Project, Faculty of Medicine, University of Toronto.

BROWN, B.B. (1975). New Mind New Body: Biofeedback; New directions for the mind. Harper Row, New York, N.Y.

BROWN Jr, W.S., MORRIS, R.J., HOLLIEN, H. and HOWELL, E. (1991). Speaking fundamental frequency characteristics as a function of age and professional singing. Journal of Voice, Volume 5, No. 4, pp. 310-315.

BRUGERE, J., GUENEL, P., LECLERC, A. and RODRIGUEZ, J. (1986). Differential effects of tobacco and alcohol in cancer of the larynx, pharynx and mouth. Cancer 57: 391-95.

BURCH, P.R. (1981). Passive smoking and lung cancer. British Medical Journal, 282:1393.

CANN, C.I. and FRIED, M.P. (1984). Determinants and Prognosis of Laryngeal Cancer. Otolaryngologic Clinics of North America, Volume 17, No. 1.

CANNON, C.R. (1994) Hypothyroidism in Head and Neck Cancer Patients: Experimental and clinical observations. Laryngoscope, 104, November.

CARDING, P.N. and HORSLEY, I.A. (1992). An evaluation study of voice therapy in non-organic dysphonia. European Journal of Disorders of Communication, Vol.27, No.2; 137-158

CARLSON, E. (1986). Watching the voice puts you in the picture. Speech Therapy in Practice, Volume 2, No. 1.

CARLSON, E.I. (1988 a). Electrolaryngography as a tool for routine voice assessment and therapy. The Voice Research Society Newsletter, Volume 2, No. 2.

CARLSON, E.I. (1988 b). Voice characteristics and quality of life after radiotherapy for glottic carcinoma. Paper presented at the National Head and Neck Oncology Conference, Nottingham.

CARLSON, E.I. (1988 c). The Laryngograph for Assessment and visual Feedback in Voice Therapy. In: 'Selected Papers from the 4th Voice Conservation Symposium'. The Voice Research Society.

CARLSON, E.I. (1993) a. Accent method plus direct visual feedback of electroglottographic signals. In: Voice Therapy, Clinical Studies. Ed. Joseph C. Stemple. Publ. Mosby Year Book Inc., St Louis.

CARLSON, E.I. (1993) b. Electrolaryngography in the assessment and treatment of puberphonia. Paper presented at the British Voice Association International Symposium. London.

CASTELIJNS, J.A. and SNOW, G.B. (1991). General aspects of laryngeal cancer. Chapter 1 in 'MR Imaging of Laryngeal Cancer', J.A. Castelijns, G.B. Snow and J. Valk (Eds.), Publ. Kluwer Academic Publishers.

CATFORD, J.C. (1964). Phonation Types: the classification of some laryngeal components of speech production. In: Abercrombie, D. et al, pp. 26-37.

CHILDERS, D.G., NAIK, J.M., LARAR, J.N., KRISHNAMURTHY, A.K. and MOORE, G.P. (1983). Electroglottography, speech and ultrahigh speed cinematography presented at the International Conference of Physiology and Biophysics of the Voice. University of Iowa, Iowa City, May 4-7. (Conference proceedings).

CHILDERS, D.G., SMITH, A.M. and MOORE, G.P. (1984). Relationships between electroglottograph, speech and vocal cord contact. Folia Phoniatrica; 36: 105-8.

CHILDERS, D.G. and KRISHNAMURTHY, A.K. (1985). A critical review of electroglottography. CRC, Critical Reviews in Biomedical Engineering, 12, 131-161.

CHILDERS, D.G., HICKS, D.M., MOORE, G.P. and ALSAKA, Y.A. (1986 a). A model for vocal fold vibratory motion, contact area and the electroglottogram. Journal of the Acoustic Society of America; 80 (5): 1309-20.

CHILDERS, D.G., ALSAKA, Y.A., HICKS, D.M. and MOORE, G.P. (1986 b). Vocal fold vibrations in dyphonia: model vs measurement. Journal of Phonetics, 14, pp. 429-434.

CHILDERS, D.G., HICKS, D.M., MOORE, G.P., ESKENAZI, L. and LALWANI, A.L. (1990). Electroglottography and vocal fold physiology. Journal of Speech and Hearing Research. Volume 33, 245-254.

COLEMAN, R.F. (1971). Effect of waveform changes upon roughness perception. Folia Phoniatrica 23: 314-322.

COLEMAN R.F. and MARKHAM, I.W. (1991). Normal variations in habitual pitch. Journal of Voice, Vol.5, No.2; 173-177.

COLTON, R.H. (1987). The Role of pitch in the discrimination of voice quality. Journal of Voice, Volume 1, No. 3; pp. 240-245.

COLTON, R.H., SAGERMAN, R.H., CHUNG, C.T., YU, Y.W. and REED, G.F. (1978). Voice change after radiotherapy: some preliminary results. Radiology; 127: 821-4.

COLTON, R.H. and CONTURE, E.G. (1990). Problems and pitfalls of Electroglottography. Journal of Voice, Volume 4. Raven Press Ltd, New York: pp. 10-24.

COLTON, R.H. and CASPER, J.K. (1990). Understanding Voice Problems. A physiological perspective for diagnosis and treatment. Publ. Williams and Wilkins.

COMINS, J.A. (1988). A Comparison of Smokers' and Non-smokers' voices in two age groups using F x and Carbon Monoxide data. Thesis for the College of Speech Therapists' Advanced specialist course in Voice Therapy.

CRANEN, B. (1991). Simultaneous modeling of EGG, PGG and glottal flow. In: Gauffin, J., Hammarberg, B. Eds. Vocal Fold Physiology: Acoustic, Perceptual and Physiological Aspects of the Voice Mechanism. San Diego, Singular Publishing Group; 57-64.

DEJONCKERE, P.H. (1981). Comparison of two methods of photoglottography in relation to electroglottography. Folia Phoniatrica, 33: 338-347.

DEJONCKERE, P.H. and LEBACQ, J. (1985). Electroglottography and vocal nodules: An attempt to quantify the shape of the signal. Folia Phoniatrica 37; pp. 327-335.

DE RIENZO, D.P., GREENBERG, S.D. and FRAIRE, A.E. (1991). Carcinoma of the larynx. Changing incidence in women. Archives of Otolaryngology- Head and Neck Surgery June, 117 (6), pp. 681-4.

DICKENS, V.J., CASSISI, N.J., MILLION, R.R. and BOVA, F.J. (1983). Treatment of early vocal cord carcinoma. A comparison of apples and apples. Laryngoscope, 93; pp. 216-19.

ESKENAZIT, L., CHILDERS, D.G. and HICKS, D.M. (1990). Acoustic correlates of vocal quality. Journal of Speech and Hearing Research, Volume 33, June; pp. 298-306.

ESLING, J.H. (1978). Voice Quality in Edinburgh: a sociolinguistic and phonetic study. Ph D. dissertation, University of Edinburgh.

FABRE, P. (1957). Un procédé électrique percutané d'inscription de l'accolement glottique au cours de la phonation: glottographie de haute fréquence. Premiers résultats. Bulletin de l'Academie Nationale de Médecine 141; pp. 66-69.

FABRE, P. (1958). Etude compareé des glottogrammes et des phonogrammes de la voix humaine. Annales d'Oto- Laryngologie 75; pp. 767-775.

FAIRBANKS, G. (1960). Voice and articulation drillbook, New York: Harper and Brothers.

FALLOWFIED, L. (1990). The Quality of Life. The missing measurement in health care. Publ. Souvenir Press, Human Horizons Series.

FANT, G. (1956). On the predictability of formant levels and spectrum envelopes from formant frequencies. In: Halle, M., Lunt, H. and MacLean, H. (Eds.), pp. 109-20. For Roman Jacobson Mouton, The Hague.

FANT, G., ONDRAÇKOVA, J., LINDKVIST, J. and SONESSON, B. (1966). Electrical glottography. Quarterly Progress and Status Report. 4, 15-21. Speech Transmission Laboratory, Royal Institute of Technology, Stockholm.

FANT, G. (1979). Glottal source and excitation analysis. Speech Transmission Laboratory. Quarterly Progress and Status Report, Kungliga Tekniska Hogskolan, Stockholm 1: 85-107.

FEX, S. and HENRIKSSON, B. (1969). Phoniatric treatment combined with radiotherapy of laryngeal cancer for the avoidance of radiation damage. Acta Otolaryngologica Supplement. 263.

FEX, B. (1987). Treatment of voice by the accent method. Working papers in Logopedics and phoniatrics. Lund University, 4: pp. 4-8.

FINKELHOR, B.K., TITZE, I.R. and DURHAM, P.L. (1988). The effect of viscosity changes in the vocal folds on the range of oscillation. Journal of Voice, Volume 1, No. 4; pp. 320-325. Raven Press Ltd, New York.

FITCH, J.L. and HOLBROOK, A. (1970). A modal vocal fundamental frequency of young adults. Archives of Otolaryngology 92, pp. 379-382.

FLANDERS, W.D. and ROTHMAN, K.J. (1982). Interaction of alcohol and tobacco in laryngeal cancer. American Journal of Epidemiology 115: pp. 371-379.

FLETCHER, G.H. and KLEIN, R. (1964). Dose-time-volume relationship in squamous-cell carcinoma of the larynx. Radiology, 1964, 82, 1032-1042.

FLETCHER, G.H. (1980). Textbook of Radiotherapy; 3rd edition. Churchill Livingstone, p. 291.

FOG PEDERSEN, M.F. (1977). Electroglottography compared with synchronised stroboscopy in normal persons. Folia Phoniatrica 29; pp. 191-199.

FOURCIN, A. and ABBERTON, E. (1971). First Application of a New Laryngograph. Medical Biology Illustrated 21; pp. 172-182.

FOURCIN, A. (1974). Laryngographic Examination of Vocal Fold Vibration. Chapter 19 in Ventilatory and Phonatory Control Systems. Ed. B. Wyke; Publ. Oxford University Press.

FOURCIN, A.J. and ABBERTON, E. (1976). The Laryngograph and the Voiscope in Speech Therapy. XVIth International Congress of Logopedics and Phoniatrics, (pp. 116-122) Ed. Loebell, E., Bern; Publ: S. Karger A.G, Basel.

FOURCIN, A.J. (1981). Laryngographic assessment of phonatory function. In: C.L. Ludlow and M.O. Hart (Eds.) Proceedings of the Conference of the Assessment of Vocal Pathology, Maryland.

FOURCIN, A.J. (1982). Electrolaryngographic assessment of vocal fold function. Journal of Phonetics 14, pp. 435-442.

FOURCIN, A. (1989)a. An aspect of voice quality in speech pathology. Voice Research Society Newsletter Volume 4, No. 1, August 1989; pp. 32-38.

FOURCIN, A. (1989)b. An aspect of voice quality in speech pathology. University College London, Department of Phonetics and Linguistics Progress Report.

FRITZELL, B., HAMMARBERG, B., GAUFFIN, J., KARLSSON, I. and SUNDBERG, J. (1986). Breathiness and insufficient vocal fold closure. Journal of Phonetics, 14; pp. 549-553.

FRØKJÆR-JENSEN, B. and PRYTZ, S. (1974). Registration of voice quality. Bruel and Kjaer Technical Review, 3; 3-17.

FRØKJÆR-JENSEN, B. and THORVALDSEN, P. (1968). Construction of a Fabre glottograph. Annual Report of the Institute of Phonetics of the University of Copenhagen (3), 1-8.

FRØKJÆR-JENSEN, B. (1979). Akustisk-statistisk analyser of stemmen. Dansk Audiologopedi 15; pp. 64-79.

FRØKJÆR-JENSEN, B. (1983). "Can Electroglottography be used in the Clinical Practise?" Folia Phoniatrica 35, 126.

FU, K.K., WOODHOUSE, R.J., QUIVEY, J.M., PHILIPS, T.L. and DEDO, H.H. (1982). The significance of laryngeal oedema following radiotherapy of carcinoma of the vocal cord. Cancer 15: pp. 655-658.

FUJIMURA, O. (1981). Body-Cover Theory of the vocal fold and its phonetic implications. Chapter 19 in Vocal Fold Physiology, Eds K.N Steveus and M. Hirono, University of Tokyo Press; pp. 271-281.

GARDNER M.J. and ALTMAN, D.G. (1989). Statistics with confidence, confidence intervals and statistical guidelines. Publ. British Medical Journal, Tavistock Square, London WC1H 9JR.

GARRETT, K. and HEALEY, C. (1987). An acoustic analysis of fluctuations in the voices of normal adult speakers across three times of day. Journal of the Acoustic Association of America 82; pp. 58-62.

GAUFFIN, J. and SUNDBERG, J. (1989). Spectrum correlates of glottal voice source waveform characteristics. Journal of Speech and Hearing Research 32: pp. 556-65.

GERRATT, B.R., HANSON, D.G., BERKE, G.S. and PRECODA, K. (1991). Photoglottography: A Clinical Synopsis. Journal of Voice, Volume 5, No. 2; pp. 98-105.

GERRITSEN, G.J. and SNOW, G.B. (1991). The patterns of growth and spread of laryngeal cancer. Chapter 2 in 'MR Imaging of Laryngeal Cancer', J.A. CASTELIJNS, G.B. SNOW and J. VALK (Eds.), Kluver Academic Publishers, The Netherlands.

GILBERT, H.R. and WEISMER, G.G. (1974). The effects of smoking on the speaking fundamental frequency of adult women. Journal of Psycholinguistic Research, Volume 3; pp. 225-31.

GILBERT, H.R., POTTER, C.R. and HOODIN, R. (1984). The Larygograph as a measure of vocal fold contact area. Journal of Speech and Hearing Research 27; pp. 178-182.

GRADDOL, D. (1986). Discourse specific pitch behaviour. In: C. Johus-Lewis (ed.) Intonation in discourse, London: Croom-Helm; pp. 221-237.

GRADDOL, D. and SWANN, J. (1983). Speaking fundamental frequency: Some physical and social correlates. Language and Speech 26; pp. 351-366.

GRAMMING, P. (1988). Non-organic dysphonia: Phonetograms for pathological voices before and after therapy. Nordisk Tidskrift for Logopedi og Foniatri, 1; 3-16.

GREEN, N. (1972). Automatic speaker recognition using pitch measurements in conversational speech. JSRU Report 1000, Joint Speech Reasearch Unit, Ruislip, Middlesex.

GREENE, M. and MATHIESON, L. (1989). The Voice and its Disorders. Whurr Publishers.

GROSJEAN, F. and COLLINS, M. (1979). Breathing, pausing and reading. Phonetica, 36; pp. 98-114.

GUENEL, P., CHASTANG, J.F., LUCE, D., LECLERC, A. and BOUGERE, J. (1988). A study of the interaction of alcohol drinking and tobacco smoking among French cases of laryngeal cancer. Journal of Epidemiology and Community Health 42; pp. 350-354.

HAJI, T., HORIGUCHI, S., BAER, T. and GOULD, W. (1986). Frequency and amplitude perturbation analysis of electroglottograph during sustained phonation. Journal of the Acoustic society of America 80 (1), July.

HALL, E.J. (1988). Radiobiology for the Radiologist. Third edition. Publ. J.B. Lippincott Company, Philadelphia.

HAMLET, S.L. (1971). Location of slope discontinuities in glottal pulse shapes during vocal fry. Journal of the Acoustic society of America 50; pp. 1561-1562.

HAMLET, S.L. and PALMER, J.M. (1974). Investigation of laryngeal trills using the transmission of ultrasound through the larynx, Folia Phoniatrica, 26; pp. 362-377.

HAMLET, S.L. (1980). Ultrasonic measurement of larynx height and vocal vibratory pattern. Journal of the Acoustic Society of America 68; pp. 121-126.

HAMMARBERG, B., FRITZELL, B., GAUFFIN, J., SUNDBERG, J. and WEDIN, L. (1980). Perceptual and Acoustic Correlates of Abnormanl Voice Qualities. Acta Otolaryngologica Supplement 90; pp. 441-51.

HAMMARBERG, B., FRITZELL, B. and SHIRATZKI, H. (1984). Teflon injection in 16 patients with paralytic dysphonia; perceptual and acoustic evaluations. Journal of Speech and Hearing Disorders 49; pp. 72-82.

HAMMARBERG, B., FRITZELL, B., GAUFFIN, J. and SUNDBERG, J. (1986). Acoustic and Perceptual analysis of vocal dysfunction. Manuscript accepted for publication in Journal of Phonetics 14; pp. 533-547.

HAMMARBERG, B. (1986). Pitch and quality characteristics of mutational voice disorders, before and after therapy. Manuscript submitted for publication.

HAMMARBERG, B. and ASKENFELT, A.G. (1986). Speech waveform perturbation analysis; a perceptual-acoustical comparison of seven measures. Journal of Speech and Hearing Research 29; pp. 50-64.

HAMMARBERG, B. (1986). Perceptual and Acoustic Analysis of Dysphonia. Studies in Logopedics and Phoniatrics No. 1, Huddinge University Hospital.

HAMMARBERG, B. (1988). Pathological voice qualities- perceptual and acoustic aspects. In: 'Selected papers from the voice conservation symposium' 1988. The Voice Reseach Society, London; pp. 15-26.

HAMMARBERG, B. (1992). Pathological voice qualities- perceptual and acoustic characterstics of a set of Swedish "reference" voices. Karolinska Institutet, Huddinge Hospital; Phoniatric and Logopedic Progress Report 8; pp. 15-22.

HANLEY, T., SNIDECOR, J. and RINGEL, R. (1966). Some acoustic differences among languages. Phonetica 14; pp. 97-107.

HANSON, D.G., GERRATT, B.R. and WARD, P.H. (1983). Glottographic Measurement of Vocal Dysfunction. A Preliminary Report. Annals of Otology, Rhinology and Laryngology, 92; 413-420.

HARDCASTLE, W.J. (1976). The Physiology of Speech Production. Academic Press, New York.

HARWOOD, A.R. and TIERE, A. (1979). Radiotherapy of early glottic cancer II. Int. J. Rad. Oncol. Biol. Phys. 1979; 5: pp. 477-82.

HARWOOD, A.R and DE BOEHR, G. (1980)a. Prognostic factors in T2 glottic cancer. Cancer 45; pp. 991-995.

HARWOOD, A.R., BRYCE, D.P. and RIDER, W.D. (1980)b. Management of T3 Glottic Cancer. Archives of Otolaryngological Research, 106; pp. 697-699.

HARWOOD, A.R., HAWKINS, N.V., KEANE, T., CUMMINGS, B., BEALE, F.A., RIDER, W.D. and BRYCE, D.P. (1980)c. Radiotherapy for early glottic cancer. The Laryngoscope, 90; pp. 465-470.

HARWOOD, A.R. and RAWLINSON, E. (1983). The Quality of Life of Patients following treatment for Laryngeal Cancer. Int. J. Radiation Oncol. Biol. Phys. Volume 9; pp. 335-338.

HECKER, M. and KREUL, E. (1974). Description of the speech of patients with cancer of the vocal folds, Part I. Measures of fundamental frequency. Journal of the Acoustic Society of America 44; pp. 1275-1282.

HENTON, C. (1983). Critique of Vocal Profile Analysis Protocol. Comments compiled in Phonetics Laboratory, Oxford University.

HERITY, B., MORIARTY, M., BOURKE, G.J. and DALY, L. (1981). A Case-Control Study of Head and Neck Cancer in the Republic of Ireland. British Journal of Cancer 43; pp. 177-182.

HERRINGTON-HALL, B.L., LEE, L., STEMPLE, J.C., NIEMI, K.R. and McHONE, M.M. (1988). Description of Laryngeal pathologies by age, sex and occupation in a treatment seeking sample. Journal of Speech and Hearing Disorders 53; pp. 57-64.

HERTEGARD, S., GAUFFIN, J. and KARLSSON, I. (1992). Physiological correlates of the inverse filtered flow waveform. Journal of Voice, Volume 6, No 3; pp. 224-234.

HILLER, S., LAVER, J. and MACKENZIE, J. (1984). Durational aspects of long-term measurements of fundamental frequency perturbations in connected speech. Work in Progress 17; pp. 59-76. Department of Linguistics, University of Edinburgh.

HILLMAN, R.E., HOLMBERG, E.B., PERKELL, J.S., WALSH, M. and VAUGHAN, C. (1989). Objective assessment of vocal hyperfunction: an experimental framework and initial results. Journal of Speech and Hearing Research, Volume 32, no 3; pp. 73-92. June 1989.

HIRANO, M. (1977). Structure and vibratory behaviour of the vocal folds. In: Sawashima, M. and Cooper, F.S.(Eds.) Dynamic aspects of speech production, University of Tokyo Press.

HIRANO, M., KURITA,S. MATSUO,K. NAGATA,K.(1980) Laryngeal tissue reaction to stress. Transcripts of the ninth symposium: Care of the professional voice, Part II (pp. 10-20) New York, The Voice Foundation.

HIRANO, M. (1981). Clinical Examination of Voice. Chapter 7. Publ. Springer-Verlag.

HIRANO, M., KIYOKAWA, K. and KURITA, S. (1988). Laryngeal muscles and glottic shaping. In: O. Fujimura (Ed.) Vocal Fold Physiology: Volume 2. Voice Production Mechanisms and Functions; pp. 49-65. Raven Press, New York.

HIRANO, M., TANAKA, S., FUJITA, M. and TERASAWA, R. (1991). Fundamental frequency and sound pressure level of phonation in pathological states. Journal of Voice, Volume 5, No 2; pp. 120-127.

HIROSE, H. and SAWASHIMA, M. (1981). Functions of the Laryngeal muscles in Speech. Chapter 11 of vocal Fold Physiology; pp. 137-154. Tokyo University Press.

HIROTO, I. (1981). Introductory remarks to 'Vocal fold Physiology'. University of Tokyo Press; pp. 3-9.

HIRSON, A. and ROE, S. (1993). Stability of voice and periodic fluctuations in voice quality through the menstrual cycle. Voice, 2, 1993, British Voice Association; 78-88.

HIXON, T., KLATT, D. and MEAD, J. (1971). Influence of forced transglottal pressure changes on vocal fundamental frequency. Journal of the Acoustic Society of America 49, 105 (A).

HOLBROOK, A., ROLNICK, M. and BAILEY, C.W. (1974). Treatment of vocal abuse disorders using a vocal intensity controller. Journal of Speech and Hearing Disorders, 39.

HOLLIEN, H. (1962). The relationship of vocal fold thickness to absolute fundamental frequency of phonation. In: 4th International Congress of Phonetic Sciences, The Hague: Mouton, pp. 173-177.

HOLLIEN, H., MOORE, P., WENDAHL, R.W. and MICHEL, J.F. (1966). On the nature of vocal fry. Journal of Speech and Hearing Research 9; pp. 245-247.

HOLLIEN, H. and MICHEL, J.F. (1968). Vocal Fry as a phonational register. Journal of Speech and Hearing Research 11; pp. 600-4.

HOLLIEN, H. (1971). Three major vocal registers: a proposal. Proceedings of the 7th International Congress of Phonetic Sciences, Montreal; pp. 320-31.

HOLLIEN, H., DEW, D. and PHILIPS, P. (1971). Phonational frequency ranges of adults. Journal of Speech and Hearing Research 14; pp. 755-760.

HOLLIEN, H. and SHIPP. (1972). Speaking fundamental frequency and chronologic age in males. Journal of Speech and Hearing Research 15; pp. 155-159.

HOLLIEN, H. and JACKSON, B. (1973). Normative data on the speaking fundamental characteristics of young adult males. Journal of Phonetics 1; pp. 117-120.

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

HOLLIEN, H. (1974). On vocal registers. Journal of Phonetics 2; pp. 125-43.

HOLLIEN, H. (1981). In search of vocal frequency control mechanisms. Chapter 31 in 'Vocal Fold Physiology'.

HOLMBERG, E.B., HILLMAN, R.E. and PERKELL, J.S. (1987). Glottal airflow and transglottal air pressure measurements for male and female speakers in soft, normal and loud voice. Journal of the Acoustic Society of America; 84; 511.

HOLMER, N.G. and RUNDQVIST, H.E. (1975). Ultrasonic registration of the fundamental frequency of a voice during normal speech. Journal of the Acoustic Society of America 58; pp. 1073-1077.

HONJO, I. and ISSHIKI, N. (1980). Laryngoscopic and voice characteristics of aged persons. Archives of Otolaryngology 106, 149.

HORII, Y. (1975). Some statistical characteristics of voice fundamental frequency. Journal of Speech and Hearing Research 18; pp. 192-201.

HORII, Y. (1979). Fundamental frequency perturbation observed in sustained phonation. Journal of Speech and Hearing Research 22, pp. 5-19.

HORII, Y. (1980). Vocal shimmer in sustained phonation. Journal of Speech and Hearing Research 23; pp. 202-209.

HOWARD, D.M., LINDSEY, G.A. and ALLEN B. (1990). Toward the Quantification of Vocal Efficiency. Journal of Voice, Volume 4, No 3; pp. 205-212.

HOWARD, D.M., LINDSEY, G. and PALMER, S. (1991). Larynx closed quotient measures for the female singing voice. Proceedings of the XIIth International Congress of Phonetic Sciences. August 19-24th, Aix en Provence, France.

HOYT, D.J., LETTINGA, J.W., LEOPOLD, K.A. and FISHER, S.R. (1992). The Effect of Head and Neck Radiation Therapy on Voice Quality. Laryngoscope Volume 102, No 5, May; pp. 477-480.

HURME, P. and SONNINEN, A. (1985). Normal and disordered voice quality. Listening tests and long term average spectrum analysis. Papers in Speech Research. University of Jyväskylä, 6; pp. 49-72.

INOUE, T., INOUE, T., CHATANI, M. and TESHIMA, T. (1992). Irradiated volume and arytenoid oedema after radiotherapy for T1 glottic carcinoma. Strahlentherapie-Onkologie, January, 168 (1); pp. 23-26.

INTERNATIONAL UNION AGAINST CANCER, UICC (1968, 1978) TNM Classification of Malignant Tumours. 1st and 3rd Ed. UICC, Geneva.

INTERNATIONAL UNION AGAINST CANCER, UICC,(1987) TNM Classification of Malignant Tumours. 4th Ed. Springer Verlag, London

ISHIZAKA, K. (1981). Equivalent lumped-mass models of vocal fold vibration. Chapter 17; pp. 231-241. In: Vocal Fold Physiology. Ed. K.N.S Leveus and Minoru Hirano.

ISON, K.T. (1985). Dx- user guide. Phonetics Department, University College, London.

ISSHIKI, N. and von LEDEN, H. (1964). Hoarseness: Aerodynamic studies. Archives of Otolaryngology, Chicago, (80) 206.

ISSHIKI, N., YANAGIHARA, N. and MORIMOTO, M. (1966). Approach to the objective diagnosis of hoarseness. Folia Phoniatrica 18; pp. 393-400.

ISSHIKI, N., OKAMURA, H., TANABE, M. and MORIMOTO, M. (1969). Differential Diagnosis of Hoarseness. Folia Phoniatrica 21; pp. 9-19.

IWATA, S. and von LEDEN, H. (1970). Pitch perturbation in normal and pathological voices. Folia Phoniatrica 22; pp. 413-424.

JOHNS, M.E., SLAUGHTER FITZ-HUGH, G., BOYD, J.C., COPLEY Mc LEAN, V., CLARK, D.A. and CANTRELL, R.W. (1983). Stage II Glottic Carcinoma: Prognostic Factors and Management. Laryngoscope 93, June.

JOHNS, M.E., NEAL, D.A. and CANTRELL, R.W. (1984). Staging of cervical lymph mode metastases. Comparison of two systems (S-stage group). Annals of Otology, Rhinology and laryngology 93; pp. 330-332.

JOHNSON, K.W. and MICHEL, J.F. (1969). The effect of selected vowels on laryngeal jitter. Asha 11, 96.

JUDD, M.W. (1983). Correction of low-frequency phase distortion of Electro-laryngograph waveform recording using a BBC microcomputer. Speech, Hearing and Language: Work in progress. University College, London, No 1.

KAGAN, A.R., CALCATERRA, T., WARD, P. and CHAN, P. (1974). Significance of oedema of the endolarynx following curative irradiation for carcinoma. American Journal of Roentgenology 120; pp. 169-172.

KAHANE, J. (1978). A morphological study of the human prepubertal and pubertal larynx. American Journal of Anatomy 151; pp. 11-20.

KAHANE, J.C. (1983). A survey of age related changes in the corrective tissues of the human larynx. In: D.M. Bless and J.H. Abbs Eds. Vocal Fold Physiology, San Diego, College Hill Press.

KAKITA,Y. HIRANO, M. MATSUSHITA, H. HIKI, S IMAIZUMI,S. (1977) Differentiation of laryngeal diseases using acoustical analysis. Practical Otolaryngology. (Kyoto) 70; 729-739.

KANEKO, T., UCHIDA, K., SUZUKI, H., KOMATSU, K., KOMESAKA, T., KOBAYASHI, N. and NAITO, J. (1981). Ultrasonic observations of vocal fold vibration. Chapter 9; pp. 107-117; Vocal Fold Physiology, Ed. Stevens and Hirano, University Tokyo Press.

KAPLAN, H.M. (1960). Anatomy and Physiology of Speech, McGraw-Hill, New York.

KAPLAN, M.J., JOHNS, M.E., SLAUGHTER FITZ-HUGH, G., BOYD, J.C., McLEAN, W.C., CLARK, D.A. and CANTRELL, R.W. (1983). Stage II Glottic Carcinoma: Prognostic Factors and Management. Laryngoscope 93: pp. 725-728.

KARIM, A.B.M.F., SNOW, G.B., DIEK, H.T.H. and HANJO, K.H. (1983). The Quality of Voice in Patients Irradiated for Laryngeal Carcinoma. "Cancer" 51, January 1; pp. 47-49.

KARNELL, M.P. (1989). Synchronized videostroboscopy and electroglottography. Journal of Voice 3; pp. 68-75.

KARNELL, M.P. (1991). Laryngeal perturbation analysis: minimum length of analysis window. Journal of Speech and Hearing Research June, 34(3); pp. 544-8.

KEIDAR, A. (1983). An acoustic-perceptual study of vocal fry using synthetic stimuli. Journal of the Acoustic Society of America 73, (Supplement 1).

KEIDAR, A., HURTIG, R.R, TITZE, I.R. (1987). The perceptual nature of vocal register change. Journal of Voice, Volume 1, No 3; pp. 223-233. Raven Press Ltd, New York.

KELMAN, A.V. (1981). Vibratory pattern of the vocal folds. Folia Phoniatrica 33; pp. 73-99.

KIM, K.M., KAKITA, Y., HIRANO, H. (1982). Sound spectographic analysis of the voice of patients with recurrent laryngeal nerve paralysis. Folia Phoniatrica 34; pp. 124-133.

KITZING, P. (1979). Glottografisk Frekvensindikering. diss. Lund University, Malmo.

KITZING, P. (1982). Photo and electroglottographical recording of the laryngeal vibratory pattern during different registers. Folia Phoniatrica 34; pp. 234-241.

KITZING, P., CARLBORG, LOFKVIST, A. (1982). Aerodynamic and glottographic studies of the laryngeal vibratory cycle. Folia Phoniatrica 34; pp. 216.

KITZING, P. (1986). Glottography - The Electrophysiological investigation of phonatory biomechanics. Acta Otol. Rhinol. Laryngolocica Belgica 40; pp. 863-875.

KITZING, P. and AKERLIND (1992). Long time average spectrograms of dysphonic voices before and after therapy. Bullelund Audiophonologie. Ann. Sc. Univ. Franche-Comte VIII: lang 2; 25-38.

KLATT, D. and KLATT, L. (1990). Analysis, synthesis and perception of voice quality variations among female and male talkers. Journal of the Acoustical Society of America 87; pp. 820-857.

KLEINSASSER, O. (1992). Revision of classification of laryngeal cancer, is it long overdue? Journal of Laryngology and Otology, March. Volume 106; pp. 197-204.

KLICH, R.J. (1982). Relationships of vowel characterisitics to listener ratings of breathiness. Journal of Speech and Hearing Research 25; pp. 574-580.

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

KOIKE, Y. and HIRANO, M. (1973). Glottal area time function and subglottal pressure variation. Journal of the Acoustical Society of America 54; pp. 1618-1627.

KOIKE, Y. and TAKAHASHI, H. and CALCATERRA, T.C. (1977). Acoustic measures for detecting laryngeal pathology. Acta Otolaryngologica 84; pp. 105-117.

KOOPMAN, J.S. (1981). Interaction between discrete causes. American Journal of Epidemiology, 113; 716-24.

KORCOK, M. (1983). Irradiation for early vocal cord cancer saves life and voice. Journal of the American Medical Association 249; pp. 1241.

KOTBY, M.N., EL-SADY, S.R., BASIOUNY, S.E., ABOU-RASS, Y.A. and HEGAZI, M.A. (1991). Efficacy of the accent method of voice therapy. Journal of Voice, Vol.5, No.4; 316-320.

KRAMER, E. (1989). A Voiscopic Analysis of Male Voices: Non-smokers, previous smokers and smokers. Aged 50-75. Unpublished MSc thesis, City University, London.

KREIMAN, J., GERRATT, B.R. and PRECODA, K. (1990). Listener experience and perception of voice quality. Journal of Speech and Hearing Research 33; pp. 103-15.

KREIMAN, J., GERRATT, B.R. and BERKE, G.S. (1992). The multi dimensional nature of pathological voice quality. Unpublished manuscript. V.A. Medical Centre, West Los Angeles and UCLA School of Medicine, L.A., CA.

KREIMAN, J., GERRATT, B.R., KEMPSTER, G.B., ERMAN, A. and BERKE, G.S. (1993). Perceptual evaluation of voice quality: Review, tutorial and a framework for future research. Journal of Speech and Hearing Research, Vol.36, February; 21-40.

KUNZE, L.H. (1964). Evaluation of methods of estimating sub-glottal air-pressure. Journal of Speech and Hearing Research 7; pp. 151-64.

LAVER, J. (1980). The Phonetic Description of Voice Quality. Cambridge University Press.

LAVER, J. (1981). Users Manual for Vocal Profile Analysis Protocol. A Perceptual Guide. Draft, October.

LAVER, J., WIRZ, S., MacKENZIE, J. and HILLER, S. (1981). A Perceptual Protocol for the analysis of Vocal Profiles In: Work in Progress, University of Edinburgh: Department of Linguisitics 14; pp. 139-155.

LAVER, J. and HANSON, R. (1981). Describing the normal voice. In: John Darby ed. Evaluation of Speech in Psychiatry. Publ. Grune and Stratton, New York.

LAVER, J., HILLER, S. and MacKENZIE, J. (1984). Acoustic Analysis of Vocal Fold Pathology. "Proceedings of the Institute of Acoustics", Volume 6, Part 4.

LAVER, J., HILLER, S. and MacKENZIE-BECK, J.V. (1992). Acoustic waveform perturbations and voice disorders. Journal of Voice, Volume 6, No 2; pp. 115-126. Raven Press Ltd., New York.

LECLUSE, F.L.E., BROCAAR, P.M. and VERSHUURE, J. (1975). The electroglottograph and its relation to glottal activity. Folia Phoniatrica 27; pp. 215-224.

LECLUSE, F.L.E. (1977). Elektroglottografie. M.D. thesis, Erasmus University, Rotterdam, Dukkeerijel in Kwijk BV: Utrecht.

von LEDEN, T. MOORE, P. and TIMKE, R. (1960). Laryngeal vibrations: measurements of the glottic wave, Part III. The Pathologic Larynx. Archives of Otolaryngology, Chicago. 71, 26.

de LEEUW, I.M. (1990). The relation between perceptual and clinical parameters of voice quality of patients with early glottic cancer before and after radiotherapy and of normal speakers. Proceedings 14, Institute of Phonetic Sciences, University of Amsterdam; pp. 27-38.

de LEEUW, I. (1991). Perceptual evaluation of voice quality before and after radiotherapy of patients with early glottic cancer and of normal speakers. Proceedings 15; pp. 109-120. Institute of Phonetic Sciences, University of Amsterdam.

LEFF, J. and ABBERTON, E. (1981). Voice pitch measurements in schizophrenia and depression. Psychological Medicine, 11; pp. 849-852.

LEHMAN, J.J., BLESS, D.M. and BRANDENBURG, J.H. (1988). An objective assessment of voice production after radiation therapy for stage I squamous cell carcinoma of the glottis. Otolaryngology, Head and neck Surgery, Volume 98, No 2, February.

LESSER, T.H.J., WILLIAMS, G. and HODDINOTT, C. (1986). Laryngographic changes following endotracheal intubation in adults. British Journal of Disorders of Communication 21 (2), August; pp. 239-244.

LIEBERMANN, P. (1961). Perturbation in vocal pitch. Journal of the Acoustical Society of America 33; pp. 597-602.

LIEBERMANN, P. (1963). Some acoustic measures of the fundamental periodicity of normal and pathologic larynges. Journal of the Acoustic Society of America 35; pp. 344-53.

LINKE, C.E. (1973) a study of pitch characteristics of female voices and their relationship to vocal effectiveness. Folia Phoniatrica. 25; pp. 173-185.

LINVILLE, S.E. (1988). Intraspeaker variability in fundamental frequency stability: Pro age related phenomenon. Journal of the Acoustic Society of America, 83; 741-745

LLEWELLYN-THOMAS, H.A., SUTHERLAND, H.J., HOGG, S.A., CIAMPI, A., HARWOOD, A.R., KEANE, T.J., TILL, J.E. and BOYD, N.F. (1984) a. Linear analogue self-assessment of voice quality in laryngeal cancer. Journal of Chronic Disorders 37; pp. 917-24.

LLEWELLYN-THOMAS, H.A., SUTHERLAND, H., CIAMPI, A., ETEZADI-AMOLI, J., BOYD, N.F. and TILL, J.E. (1984)b. The Assessment of Values in Laryngeal Cancer: Reliability of Measurement Methods. Journal of Chronic Disorders, Volume 37, No 4; pp. 283-291.

LO, T.C.M., SALZMAN, F.A. and SCHWARTZ, M.R. (1985). Radiotherapy for Cancer of the Head and Neck. In: Otolaryngologic Clinics of North America, Volume 18, No 3, August; pp. 521-531.

LYNDES, K.O. (1975). The application of biofeedback to functional dysphonia. Biofeedback Research Society, 6th annual meeting, Monterey CA.

MacCURTAIN, F. and FOURCIN, A.J. (1982). Applications of the electrolaryngograph waveform display. Lawrence Van, L. (ed.) Transcripts of the tenth Symposium on Care of the Professional Voice, New York. The Voice Foundation, Part III; 51-57.

McCOY, G.D., WYNDER, E.L. (1979). Etiological and preventative implications in alcohol carcinogenesis. Cancer Research. 39;2844-50.

McILWAIN, J.C. (1991). The posterior glottis. Journal of Otolaryngology, 20. Supplement, p. 1-24.

MAIER, H., DIETZ, A., HELLER, V.D.and JüNEMANN, K.H. (1990). The Role of tobacco, ethanol consumption and occupation as risk factors for laryngeal carcinoma. Otorhinolaryngology, Head and Neck Surgery; pp. 2433-2436. Proc. XIV World Congress of Otorhinolaryngology Head and Neck Surgery, Madrid, Spain, September, 1989.

MANTRAVADI, R.V.P., LIEBNER, E.J., HAAS, R.E., SKOLNIK, E.M. and APPLEBAUM, E.L. (1983). Cancer of the glottis: Prognostic factors in radiation therapy. Radiation 149; pp. 311-314.

MARKEL, J.D. and DAVIS, S.B. (1979). Text independent speaker recognition from a large linguistically unconstrained time- spaced data base. IEEE Transactions, Acoustics Speech and Signal Processing ASSP-25; pp. 330-337.

MARKS, R.D., FITZ-HUGH, G. and CONSTABLE, V.C. (1971). Fourteen years experience with Cobalt 60 Radiation Therapy. In: The Treatment of Early Cancer of the True Vocal Cords. Cancer 28; pp. 571-576.

McGLONE, R.E. and SHIPP, T. (1971). Some physiologic correlates of vocal fry phonation. Journal of Speech and Hearing Research 14; pp. 769-75.

MEAD, K.O. (1974). Identification of speakers from fundamental frequency contours in conversational speech. JSRU Report 1002, Joint Speech Research Unit, Ruislip, Middlesex.

MENDENHALL, V.M., PARSOUS, J.T., STRINGER, S.P. (1988). T1-T2 vocal cord carcinoma: A basis for comparing the results of irradiation and surgery. Head and Neck Surgery 10; pp. 373-377.

MENDONCA, D.R. (1975). State of the patient after successful irradiation for laryngeal cancer. Laryngoscope, 85, I; pp. 534-539.

MICHEL, J.F. (1968). Fundamental frequency investigation of vocal fry and harshness. Journal of Speech and Hearing Research 11; pp. 590-4.

MICHEL, J.F. and HOLLIEN, H. (1968). Perceptual differentiation in vocal fry and harshness. Journal of Speech and Hearing Research 11; pp. 439-43.

MILISEN, R. (1957). Methods of evaluation and diagnosis of speech disorders. In: Travis, L.E., Handbook of Speech Pathology. Appleden Century Crofts, New York.

MILLER, S. HARRISON, L.B. SOLOMON, B. and SESSIONS, R.B. (1990) Vocal changes in patients undergoing radiation therapy for glottic carcinoma. Laryngoscope. 100, No 6; pp. 603-606.

MONSEN, R.B. and ENGEBRETSON, A,M. (1977). Study of variations in the male and female glottal wave. Journal of the Acoustical Society of America 65; pp. 981-93.

MOOK YOON, K., KAKITA, Y. and HIRANO, M. (1984). Sound spectrographic Analysis of the Voice of Patients with Glottic Carcinoma. Folia Phoniatrica 36; pp. 24-30.

MOORE, G.P. (1971). Organic Voice Disorders. Prentice-Hall, Englewood Cliffs.

MOORE, P. and von LEDEN, H. (1958). Dynamic variations of the vibratory pattern in the normal larynx. Folia Phoniatrica 10; pp. 205-38.

MORGAN. (1988). Voice quality ten years after radiotherapy for early glottic cancer. Clinical Radiology, May; 39(3); pp. 295-

MOTTA, G., CESARI, M., IENGO, G. and MOTTA Jr. (1990). Clinical application of Electroglottography. Folia Phoniatrica 42; pp. 111-117.

MURPHY, C.H. and DOYLE, P.C. (1987). The effects of cigarette smoking on voice fundamental frequency. Otolaryngology, Head and Neck Surgery, Volume 97, No 4, October.

MURRY, T., BONE, R.C. and von ESSEN, C. (1974). Changes in Voice Production During Radiotherapy for Laryngeal Cancer. Journal of Speech and Hearing Disorders 39, 2.

MURRY, T., SINGH, S. and SARGENT, M. (1977). Multidimensional classification of abnormal voice qualities. Journal of the Acoustical Society of America 61; pp. 1630-1635.

MURRY, T. (1978). Speaking fundamental frequency characteristics associated with voice pathologies. Journal of Speech and Hearing Disorders, Volume 43, No 3, 3 August; pp. 374-379.

MURRY, T. and DOHERTY, E.T. (1981). Selected acoustic characteristics of pathologic and normal speakers. Journal of Speech and Hearing Research 23(2); pp. 361-369.

MYERSON, M.C. (1964). The Human Larynx. Springfield Illinois, Publ. Charles C. Thomas; pp. 186-90.

MYSAK, E.D. (1959). Pitch and duration characteristics of older males. Journal of Speech and Hearing Research 2; pp. 46-54.

NEEL, III H.B., DEVINE, K.D. and De SANTO, W. (1980). Laryngofissure and cordectomy for early cordal carcinomas: Outcome in 182 patients. Otolaryngology, Head and Neck Surgery 88; pp. 79-84.

NEIL, W.F., WECHSLER, E. and ROBINSON, J.M.P. (1977). Electrolaryngography in laryngeal disorders. Clinical Otolaryngology 2; pp. 33-40.

NETSELL, R., LOTZ, W.K. DUCHANE A.S. and BARLOW S. M. (1991) Vocal tract aerodynamics during syllable production: Normative data and theoretical implications. Journal of Voice, Vol. 5, No 1; pp. 1-9.

NOLAN, F. (1983). The Phonetic Bases of Speaker Recognition. Cambridge University Press.

NOSCOE, N.J., FOURCIN, A.J., BROWN, N.J. and BERRY, R.J. (1983). Examination of vocal fold movement by ultra-short pulse x-radiography. The British Journal of Radiology 56; pp. 641-645.

OGLE, S.A., MAIDMENT,J.A. (1993) Laryngographic analysis of child directed speech. European Journal of Disorders of Communication 28, No. 3, 289-297.

OHLSSON, A.C., JÄRVHOLM, B., LÖFQVIST, A., NÄSLUND, P.E. and STENBORG, R. (1987)a. Vocal behaviour in Welders- a Preliminary Study. Folia Phoniatrica 39; pp. 98-103.

OHLSSON, A.C., JARVHOLM, B. and LOFQVIST, A. (1987)b. Vocal symptoms and vocal behaviour in teachers. Scandinavian Journal of Logopedics Phoniatrics, Volume 12, No 2; pp. 61-69.

OHLSSON, A.C. (1988). Voice and Work Environment Towards and ecology of vocal behaviour. Publ. Department of Logopedics and Phoniatrics, Gothenburg University, Sweden.

ORLIKOFF, R.F. (1991). Assessment of the dynamics of vocal fold contact from the electroglottogram: Data from normal male subjects. Journal of Speech and Hearing Research, Volume 34, October; pp. 1066-1072.

ORLIKOFF, R.F. and KAHANE, J.C. (1991). Influence of mean sound pressure level on jitter and shimmer measures. Journal of Voice, Volume 5, No 2; pp. 113-119.

PAINTER, C. (1988). Electroglottogram waveform types. Archives of Otolaryngology, Rhinology and Laryngology, 245; pp. 116-21.

PALMER, J.M. (1972) Anatomy for Speech and Hearing. Publ. Harper and Row.

PEGORARO-KROOK, M.I. (1986). Speaking Fundamental Frequency Characteristics of Normal Swedish Subjects obtained by Glottal

Frequency Analysis (GFA). Unpublished paper, Department of Phoniatrics, Malmö General Hospital, University of Lund, Sweden. (Publ. (1988) in Folia Phoniatrica, 40; pp. 82-90.)

PRYTZ, S. (1977). Long - time - average - spectra (LTAS) analyses of normal and pathological voices. Proceedings of the International Association of Logopedics and Phoniatrics Congress, Copenhagen, Volume 1; p. 457.

PTACEK, P.H., SANDER, E.K., MALONEY, W.H. and JACKSON, C.C.R. (1966). Phonatory and Related Changes with advanced age. Journal of Speech and Hearing Research 9; pp. 353-60.

RAINBOW, D. (1985). An Investigation into the voice of male smokers. Final year project, National Hospitals College of Speech Sciences, London.

RAMIG, L.A. and RINGEL, R.L. (1983). Effects of physiological aging on selected acoustic characteristics of voice. Journal of Speech and Hearing Research 26; pp. 22-30.

RAMMAGE, L.A., PEPPARD, R.C. and BLESS, D.M. (1992). Aerodynamic, laryngoscopic and perceptual acoustic characteristics in dysphonic females with posterior glottal chinks: A retrospective Study. Journal of Voice, 6, (1); 64-78.

RISKA, T.B. and LAUERMA, S. (1966). Die stimmfunktion nach der Behandlung von Stimmbandskarzinan im Stadium I. Acta Otolaryngologica, Suppl. 224; pp. 501-514.

RONTAL, E., RONTAL, M., JACOB, H.J. and ROHNICK, M.I. (1983).Quantitative and Objective Evaluation of Vocal Cord Function. Annals of Otolaryngology, Rhinology and Laryngology, 92; pp. 421-423.

ROTHENBERG, M. (1981) Some relationships between glottal airflow and vocal fold contact area. Chapter 8, Proceedings from the conference on Assessment of Vocal Pathology. Report No 11; pp. 88-96. American Speech, Language and Hearing Association.

ROTHENBERG, M. and MAHSHIE, J. (1988). Monitoring vocal fold abduction through vocal fold contact area. Journal of Speech and Hearing Research, 31; 338-51.

ROTHENBERG, S. and ROTHENBERG, M. (1989). Variation in EGG output among average sized men and women and heavy and muscular men. Unpublished manuscript.

ROTHENBERG, M. (1992). A Multichannel electroglottograph. Journal of Voice, Volume 6, No 1; pp. 36-44.

ROTHMAN, K.J., CANN, C.I. and FLANDERS. (1980). Epidemiology of laryngeal cancer. Epidemiological Review, 2; pp. 195-209.

RUBIN, P. and CASSARETT, G.W. (1968) The Endochrine Glands in Clinical Radiation Pathology. W.B. Saunders, Philadelphia, pp. 749-764.

RUGG, T., SAUNDERS, M.I. and DISCHE, S. (1990). Smoking and mucosal reactions to radiotherapy. British Journal of Radiology, 63; pp. 554-556.

SALLINEN-KUPARINEN, A. (1985). Pitch level and type of oral task. Papers in Speech Research, No 6; pp. 81-90. Department of Communication, University of Jyväskylä, Finland.

SASAKI, Y., OKAMURA, H. and YUMOTO, E. (1991). Quantitative analysis of hoarseness using a digital sound spectograph. Journal of Voice, Volume 5, No1; pp. 36-40, Raven Press, New York.

SAXMAN, J.H. and BURK, K.W. (1967). Speaking Fundamental frequency characteristics of middle aged females. Folia Phoniatrica, 19; pp. 167-172.

SCHAEDEL, G. (1979). Spectographic Analysis in Case Management of Vocal Pathology. Kay Elmetrics Corporation.

SCHIPPER, H., CLINCH, J., McMURRAY, A. and LEVITT, M. (1984). Measuring the quality of life of cancer patients. The functional living index- cancer: development and validation. In: Journal of Clinical Oncology, 2; pp. 472-83.

SCHULTZ-COULON, H.J. (1975). Bestimming und Beurteilung der individuellen mittleren sprechstimmlage. Folia Phoniatrica 27; pp. 375-386.

SEDERHOLM, E., McALLISTER, A., SUNDBERG, J. and DALKVIST, J. (1992). Perceptual analysis of child hoarseness using continuous scales. Speech Transmission Laboratory, Quarterly Progress and Status Report, April 15; pp. 99-113.

SHAPSAY, S.M. and HYBELS, R.L. (1985). Treatment of cancer of the larynx. Analysis of success and failure. In: Otolaryngologic Clinics of North America, Volume 18, No 3, August; pp. 461-468.

SHAW, H. (1979). Tumours of the larynx. In: Scott-Brown's Diseases of the Ear, Nose and Throat. Ed. J. Ballantyne and J. Groves. Volume 4, The Pharynx and Larynx, 4th Edition. Publ. Butterworths.

SHIPP, T. and HOLLIEN, H. (1969). Perception of the ageing male voice. Journal of Speech and Hearing Research, 12; pp. 703-710.

SHIPP, T. and McGLONE, R.E. (1971). Laryngeal dynamics associated with voice frequency change. Journal of Speech and Hearing Research, 14; pp. 761-68.

SHVILI, Y., ZOHAR, Y. and RAHIMA, M. (1990). Anterior commssure carcinoma of the larynx - Irradiation versus conservation surgery. Otorhinolaryngology, Head and Neck Surgery; pp. 2467-2469. Proceedings of the XIV World Congress of Otorhinolaryngology, Head and Neck surgery, Madrid, Spain, 1989. Ed. Sacristan, T., Alararez-Vinvent, J.Z., Bartual, J., Antoli-Candela, et al. Publ. Kugler and Ghedini Publicaltions, Amsterdam, Berkley, Milano.

SIKORA, K. and HALNAN, K.E. (Eds.). (1990). Treatment of Cancer. 2nd Edition, 1990, Chapman and Hall Medical, London.

SISSON, G.A. and PELZER, H.J. (1985). Staging system by sites. Problems and Refinements. In: Otolaryngologic Clinics of North America, Volume 18, No 3, August; pp. 397-402.

SKOLNIK, E.M., KING, F.Y., WHEATLEY, M.H. and MARTIN, L.O. (1975). Carcinoma of the Laryngeal Glottis. Therapy and end results. Laryngoscope, Volume 85, No 9; pp. 1453-1466, September.

SLEVIN, M.L., PLANT, H., LYNCH, D. et al. (1988). Who should measure quality of life, the doctor or the patient? In: British Journal of Cancer, 57; pp. 109-12.

SMITH, S. and THYME, K. (1976). Statistic Research on changes in speech due to Pedagogic Treatment, (the Accent method). Folia Phoniatrica, 28; pp. 98-103.

SMITH, S. and THYME, K. (1978). Accentmetoden. Special-pedagogisk forlag A.S. Herning.

SMITHERAN, J.R. and HIXON, T.J. (1981). A clinical method for estimating laryngeal airway resistance during vowel production. Journal of Speech and Hearing Disorders, 46, May; pp. 138-146.

SNIDECOR, J.C. (1943). A comparative study of the pitch and duration characteristics of impromptu speaking and oral reading. Speech Monographs, 10; pp. 50-57.

SöDERSTEN, M. and LINDESTAD, P.A. (1990). Glottal Closure and Perceived Breathiness during Phonation in normally speaking Subjects.

Journal of Speech and Hearing Research, Volume 33, September; pp. 601-611.

SÖDERSTEN, M. and HAMMARBERG, B. (1992). Vocal fold closure, perceptual and fundamental frequency characteristics in normally speaking females before and after voice training. Karolinska Institutet, Huddinge Hospital, Phoniatric and Logopedic Progress Report, 8; pp. 23-29.

SÖDERSTEN, M., HERTEGARD, S. and HAMMARBERG, B. (1994). Glottal closure, airflow and voice quality in middleaged women as related to changes in loudness. Phoniatric and Logopedic Progress Report, 9, Karolinska Institutet, Huddinge Hospital; 3-20.

SOPKO, J. (1986). Zur Objecktivierung der stimmlippenschwingungen mittels synchroner elektrographischer und stroboskopischer untersuchung. Sprache-Stimme-Gehör 10; pp. 83-87. Georg Thieme Verlag Stuttgart, New York.

SORENSEN, D. and HORII, Y. (1982). Cigarette smoking and voice fundamental frequency. Journal of Communication Disorders, 15; pp. 135-44.

SORENSEN, D. and HORII, Y. (1984). Directional Perturbation factors for jitter and for shimmer. Journal of Communication Disorders, 17; pp. 143-151.

de STEFANI, E. CORREA, P., OREGGIA, F. LEIVA, J.RIVERO, S. FERNANDEZ, G. DENEO-PELLEGRIN, H. ZAVALA, D. FONTHAM, E. (1987). Risk factors for Laryngeal Cancer. Cancer, 60 III; pp. 3087-3091.

STELL, P.M. and MORRISON, M.D. (1973). Radiation necrosis of the larynx. Archives of Otolaryngology 98; pp. 111-113.

STEMPLE, J., WEILER, E., WHITEHEAD, W. and KOMRAY, R. (1980). EMG biofeedback training with patients exhibiting a hyperfunctional voice disorder. Laryngoscope 90; pp. 471-476.

STEMPLE, J.C. (ed.) (1993). Voice Therapy: Clinical Studies. Mosby-Year Book Inc.

STOICHEFF, M.L. (1975). Voice following radiotherapy. Laryngoscope, 85:I; pp. 608-617.

STOICHEFF, M.L. (1981). Speaking Fundamental Frequency characteristics of non-smoking female adults. Journal of Speech and Hearing Research, 24; pp. 437-41.

STOICHEFF, M.A., CIAMPI, A., PASSI, J.E. and FREDRICKSON, J.M. (1983). The irradiated larynx and voice: a perceptual study. Journal of Speech and Hearing Research, 26; pp. 482-5.

STRONG, M.S. (1975). Laser excision of carcinoma of the larynx. Laryngoscope, 85; 1286-1289.

STURLAUGSSON, W.R. (1975). Biofeedback and Psychogenic Voice Disorders: A general review and case study. ASHA Convention Presentation.

SUNDBERG, J. and NORDSTRöM, P.E. (1976). Raised and lowered larynx-the effect on vowel formant frequencies. Quarterly Progress and Status Report of the Speech Transmission Laboratory, RIT Stockholm, Volumes 2-3; pp. 35-9.

SUNDBERG, J. and GAUFFIN, J. (1979). Waveform and spectrum of the glottal voice source. In: Frontiers of speech communication research. Festschrift for Gunnar Fant, ed. B. Lindblom and S. öhman, London Academic Press; pp. 301-20.

SUNDBERG, J. (1987). The Science of the Singing Voice. Northern Illinois University Press.

SUTHERLAND, H.J., LLEWELLYN-THOMAS, H.A., HOGG, S.A. et al. (1984). Do patients and physicians agree on the assessment of voice quality in laryngeal cancer? Journal of Otolaryngology, 13; pp. 325-30.

TAKAHASHI, H. and KOIKE, Y. (1975). Some perceptual dimensions and acoustical correlates of pathological voices. Acta Oto-Laryngologica, Supplement 338; pp. 1-24.

TITZE, I.R. and STRONG, W.J. (1975). Normal modes in vocal cord tissues. Journal of the Acoustical Society of America, 57; pp. 736-44.

TITZE, I.R. (1980). Comments on the Myoelastic-Aerodynamic Theory of Phonation. Journal of Speech and Hearing Research, 23, September; pp. 495-510.

TITZE, J. and TALKIN, D. (1981). Simulation and Interpretation of Glottographic waveforms. ASHA reports, No 11; pp. 48-55.

TITZE, J. (1984). Parameterization of the glottal area, glottal flow and vocal fold contact area. Journal of the Acoustical Society of America, 75(2), February; pp. 570-580.

TITZE, I.R., JIANG, J. and DRUCKER, D.G.A. (1988). Preliminaries to the Body- Cover theory of pitch control. Journal of Voice, Volume 1, No 4; pp. 314-319, Raven Press Ltd, New York.

TITZE, I.R. (1989)a. Physiologic and acoustic differences between male and female voices. Journal of the Acoustical Society of America, 85; pp. 1699-1707.

TITZE, I.R. (1989)b. On the relation between subglottal pressure and fundamental frequency. Journal of the Acoustical Society of America, 85(2); pp. 901-06.

TITZE, I.R. (1989)c. A four-parameter model of the glottis and vocal fold contact area. Speech Communication, Volume 8, No 3, September; Elsevier Science Publishers B.V, North Holland.

TITZE, I.R. (1990). Interpretation of the Electroglottographic signal. Journal of Voice, Volume 4, No 1; pp. 1-9.

TRUDGILL, P. (1974). The social differentiation of English in Norwich. Cambridge University Press.

TUYUS, A.J. (1988). Cancer of the larynx, hypopharynx, tobacco and alcohol. IARC, International Case Control Study in Turin and Varese, Zaragoza and Navarra, Geneva and Calvados. In: Journal of Cancer, 41; pp. 483-491.

Van MICHEL, C.L. (1967). Morphologie de la courbe glottographique dans certains troubles fonctionnels du larynx. Folia Phoniatrica, 19 pp. 192-202.

Van MICHEL, C. PFISTER, K.A. LUCHSINGER, R. (1970) Elctroglottographie et cinematographie laryngee ultra-rapide comparison des resultats. Folia Phoniatrica 22; 81-91.

Van RIPER, C. and IRWIN, J.V. (1958). Voice and articulation Prentice Hall, Englewood Cliffs.

VERDOLINI- MARSTON, K. TITZE, I.R. DRUKER, D.G. (1990) Changes i phonation threshold pressure with induced conditions of hydration. J of Voice Vol. 4, (2), pp. 142-151.

VERDOLINI - MARSTON,K. SANDAGE, M. TITZE,I. (1994) Effect o hydration treatments on laryngeal nodules and polyps and relate voice measures. J. of Voice, Vol. 8, (1) pp. 30-47.

WANG, C.C. (1974). Treatment of Glottic Carcinoma by Megavoltag Radiation Therapy Results. American Journal of Roentgenology, 120 pp. 157-163.

VATERHOUSE, J. et al (Ed.) (1971). Cancer Incidence in Five Continents, Lyon JARC, Scientific Publications, Volume III.

VECHSLER, E. (1975). The use of the laryngograph in the study of some patients with voice disorders. Unpublished M.Sc. thesis in the University of London.

VECHSLER, E., NEIL, V.F. and FOURCIN, A.J. (1976). Laryngographic analysis of pathological vocal fold vibration. Proceedings of the Institute of Acoustics.

VECHSLER, E. (1977). A Laryngographic Study of voice disorders. British Journal of Disorders of Communication. Vol. 12, No 1, pp. 9-22.

WENDAHL, R.V. (1963). Laryngeal analog synthesis of harsh voice quality. Folia Phoniatrica, 15; pp. 241-50.

VENDAHL, R.W. (1964). The role of amplitude breaks in the perception of vocal roughness. American Speech and Hearing Association Convention Abstracts, 6; 406.

VENDAHL, R.W. (1966). Laryngeal analog synthesis of jitter and shimmer auditory parameters of harshness. Folia Phoniatrica, 18; pp. 98-108.

VENDLER, J., RAUHUT, A. and KRÜGER, H. (1985). Classification of voice qualities by means of LTAS statistical approaches. Paper at the Symposium on Voice Acoustics and Dysphonia, Gotland, Sweden.

WERNER-KUKUK, E., Von LEDEN, H. and YANAGIHARA, N. (1968). The effects of Radiation Therapy on Laryngeal Function. Journal of Laryngology and Otology, June.

WEWERS, M.E. and LOWE, N.K. (1990). A critical review of visual analogue scales in the measurement of clinical phenomena. Nursing and Health, 13; pp. 227-236.

WHITTET, H.B., LUND, V.J., BROCKBANK, M. and FEYERABEND, C. (1991). Serum cotiniue as an abjective marker for smoking habit in head and neck malignancy. Journal of Laryngology and Otology, December; 105(12); pp. 1036-9.

VIERNIK, G. (1981). Three or five times weekly fractionation evidence from the B.I.R. fractionaltion survey of laryngeal carcinoma. British Journal of Radiology, 54; pp. 170-171.

VIERNIK, G., BATES, T.D., BLEHEEN, N.M., BRINDLE, J.M., BULLIMORE, J., FOWLER, J.F., HAYBITTLE, J.L., HOWARD, N., LAING, A.H., LINDUP, R., McGURK, F., PHILLIPS, D.L., and REZVANI, M. (1990). Final report of the general clinical results of the British Institute of Radiology fractionation study of 3F/week versus 5F/week in radiotherapy of carcinoma of the laryngopharynx. British Journal of Radiology, 63; pp. 169-180.

WILLIAMS, C.E. and STEVENS, K.N. (1981). Vocal correlates of Emotional States in Speech Evaluation in Psychiatry. Ed. John K. Darby; Publ. Grume and Stratton.

WOLFE, V.I., STEINFATT, T.M. (1987). Prediction of vocal severity within and across voice types. Journal of Speech and Hearing Research, 30; pp. 230-240.

WOODHOUSE, R.J., QUIVEY, J.M., FU, K.K., SIEN, P.S., DEDO, H.H. and PHILIPS, T.L. (1981). Treatment of carcinoma of the vocal cord. A review of 20 years experience. The Laryngoscope, 91; pp. 1151-1162.

WYNDER, E.L., COVEY, L.S., MABUCH, K. and MUSHINSKI, M. (1976). Environmental factors in cancer of the larynx: A second look. Cancer, 38; pp. 1591-1601.

WYNDER, E.L. and STELLMAN, S.D. (1977). Comparative epidemiology of tobacco-related cancers. Cancer Research, 37; pp. 4608-4622.

WYNTER, H. (1974). An investigation into the analysis and terminology of voice quality and its correlation with assessment reliability of speech therapists. British Journal of Disorder Communication.

YANAGIHARA, N. (1967). Significance of harmonic changes and noise components in hoarseness. Journal of Speech and Hearing Research, 10; pp. 531-541.

YOUNG, J.L. Jr., PERCE, C.L. and ASIRE, A.J. (Ed.) (1981). Surveillance, Epidemilogy and End results: Incidence and Mortality Data 1973-77. NCI Monograph 57 NIH, PHS, DHEW Publication No (NIH), 81-2330.

YUMOTO, E., SASAKI, Y. and OKAMURA, H. (1984). Harmonics to noise ratio and psychophysical measurement of the degree of hoarseness. Journal of Speech and Hearing Research, 27; pp. 2-6.

YUMOTO, E., GOULD, W.J. and BAER, T. (1982). Harmonics to noise ratio as an index of the degree of hoarseness. Journal of the Acoustical Society of America, 71; pp. 1544-1550.

ZEMLIN, W.R. (1964). Speech and Hearing Science, Stipes, Champaign, Illinois.

SUBSIDIARY MATTER:

Carlson E.I. Accent method plus direct visual feedback of electroglottographic signals.
From: Voice Therapy; Clinical studies. Ed. Joseph C. Stemple.
Publ. Mosby Year Book Inc. St Louis, MO. 1993.

ACCENT METHOD PLUS DIRECT VISUAL FEEDBACK OF ELECTROGLOTTOGRAPHIC SIGNALS

Contributed by Eva Carlson, M.Sc.

In the next case study, Eva Carlson, M.Sc., combines the Accent Method of therapy with direct visual feedback of electroglottographic signals to modify vocal hyperfunction.

LARYNGOGRAPHY

For the past 7 years the Fourcin "laryngograph" has been in daily use for routine voice assessment and visual feedback in treatment, in the department of speech and language therapy at St. Thomas' Hospital London (Fourcin and Abberton, 1971; Fourcin, 1974, 1981).

The laryngograph is an electroglottograph (EGG), but Fourcin chose to call his version a "laryngograph," and this particular EGG technique is therefore called electrolaryngography (ELG). The resulting waveform is called "Lx" (larynx excitation) and is customarily shown in a positive-going direction (on y axis) for in-

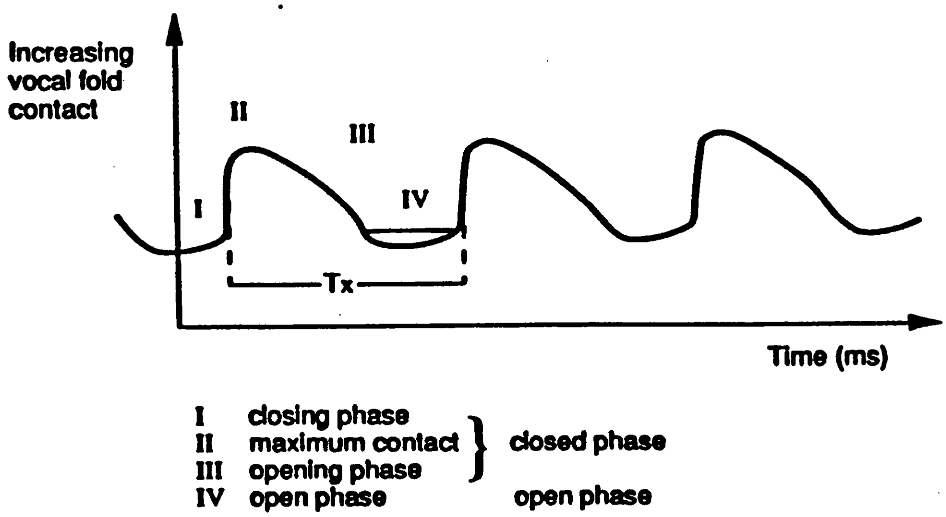


FIG 3–5.
Laryngographic waveform (Lx).

creasing vocal fold closure (Fig 3–5). The opposite tends to be the case in most other reports of EGG data, where the EGG waveform is shown in a negative-going direction for increasing vocal fold closure.

Two surface electrodes are placed either side of the thyroid cartilage at the level where the strongest Lx signal is achieved on phonation. The Lx signal reflects the variation in electrical conductance between the electrodes, as the vocal folds open and close. The amplitude of Lx indicates the amount of vocal fold contact. The greater the tissue contact the greater the current flow between the electrodes and the higher the amplitude (Gilbert et al., 1984). The Lx signal does not differentiate between the degree of vertical or horizontal closure but gives an integrated measure of the total current flow between the electrodes.

Voice therapy is often aimed at teaching patients to improve their breath support to achieve the right balance between subglottal pressure and laryngeal adduction forces, rather than just varying laryngeal resistance for voice production. The resulting aerodynamic changes and laryngeal tension characteristics give rise to increased Lx amplitude and steeper rise in Lx waveform. It is hypothesized to be due to the Bernoulli effect playing a greater part in glottic closure, resulting in increased tissue contact and a faster rate of closure of the vocal folds.

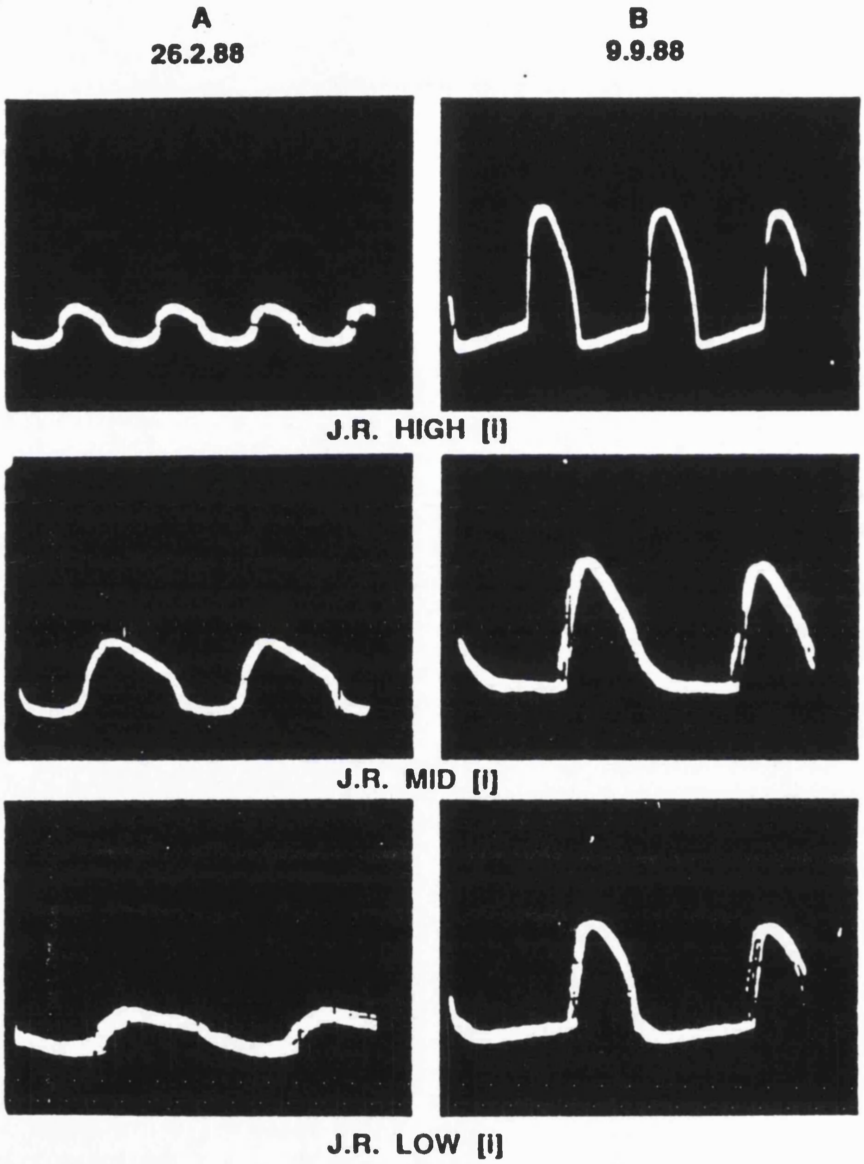
A male patient with an extremely "pressed" voice assessed in February 1988, was helped to produce voice with considerably less effort using Lx for visual feedback. At the end of a period of voice therapy (Fig 3–6, B), he sustained waveforms of higher amplitude and with steeper closing phase gradients, indicating faster closure rate. If anything, he was producing a voice with rather "whispery" quality in his effort to reduce laryngeal tension as evidenced by the long open phase, particularly at "high [i]" (see Fig 3–6, B).

The speed of vocal fold closure is an important factor in determining voice quality (Fourcin, 1981; Kelman, 1981; Sundberg, 1987) and can be monitored by looking at the steepness of the rising portion of the Lx waveform (Fig 3–6). Colton and Conture (1990) measured the "closing time" of the EGG waveform, defined as the time from the start of the closing phase of the EGG signal (Fig 3–5) to the point of maximum closure. They found that lesions on one or both vocal folds often increased the closing time.

The duration of the closed phase is another indicator of the efficiency of the voice source (Fourcin, 1981). During the open phase (Fig 3–5), acoustic energy is lost in the subglottis. However, in "pressed" and hyperkinetic voice, the closed phase occupies an excessive amount of the whole period and gives rise to characteristic Lx waveshapes (Fig

6, A). In patients with a lot of subcutaneous overlying the thyroid cartilage, common in omen, or in older men with short necks, ere the larynx has dropped to a point where thyroid alae are difficult to access, the Lx gnal is full of "noise" and is not a reliable urce of information of glottal dynamics. lton and Conture (1990) report that they re unable to get reliable EGG signals from % of their patients.

However, in clinical practice a poor signal can sometimes be useful for comparison purposes. Figures 3–7, A and B show the waveforms from a subject who had radiation therapy for glottic carcinoma and later asked for help with his voice, as it was deteriorating despite no evidence of recurrence of the tumor. The Lx waveforms from December 8, 1989 (Fig 3–7, A) show some periodicity but so much noise in the signal that a scientist would no doubt consider



IG 3–6.
, pretherapy pressed or hyperkinetic waveform. B, posttherapy waveform demonstrating reduced laryngeal ension.

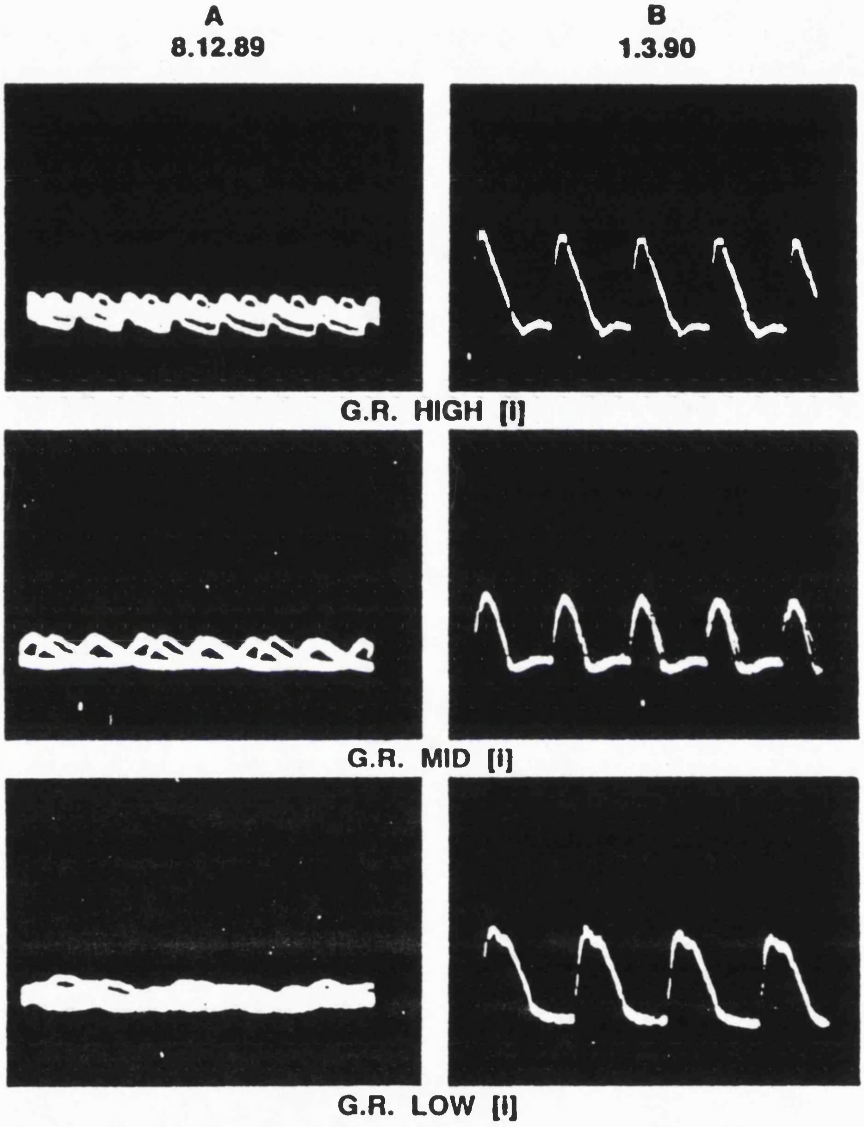


FIG 3–7.
A, Lx waveform demonstrating signal noise in patient after vocal cord radiation therapy. B, Lx waveform of same patient after voice therapy.

them worse than useless. However, as a basis for clinical comparison of the patient's later ability to produce and sustain Lx with good amplitude and no evidence of signal interference (Fig 3–7, B), they serve a useful purpose. The amplitude of the Lx waveform was used for feedback during therapy, to confirm that his new vocal habits were resulting in improved glottal dynamics. He used the waveform to monitor this and gain confidence in his ability to vary pitch and volume without effort, in a variety of vocal exercises, conversation, and reading aloud.

This case study aims to describe how the Lx waveform, and fundamental frequency parameters derived from Lx, can be used to give visual feedback of changing glottal dynamics as a result of vocal therapeutic and surgical intervention. It is not used for diagnosis but may, as will be shown later, reflect observed organic vocal fold changes that interfere with vocal fold contact.

Abberton and Fourcin (1984) and Abberton et al. (1989) give a detailed description of the laryngograph and the measurements derived from the Lx signal. Limitations of the technique are well summarized by Baken (1987) and more recently by Colton and Conture (1990) and by Childers et al. (1990). Suggested improvements and development of the technique are described by Rothenberg (1991). Finally, simulation of EGG waveforms from computer modeling of glottal configurations, incidentally with the same orientation in space as Lx and strongly reminiscent of waveforms produced in the clinic, are reported by Titze (1990).

Equipment Used

The Laryngograph Processor, which is a development of the original Laryngograph, is linked to a BBC Master microcomputer, a Cumana dual discdrive, and an Epson FX80 printer. A Telequipment S61 oscilloscope is used to show the Lx waveform in real time. The input from the laryngeal surface electrodes and an RS professional dynamic microphone is recorded on a Sony TC 144CS stereo cassette tape recorder.

Objective Voice Parameters Derived From Lx

The ELG parameters that we use most for voice assessment and visual feedback during treatment are the following:

1. The Lx waveform itself (Fig 3–5), showing a closing phase, a point of maximum contact, an opening phase, and an open phase (Abberton et al. 1989). It is used as a basis for calculation of fundamental frequency, Fx, by measuring the time, Tx, between successive vocal fold closures (Fig 3–5). The Laryngography Processor allows the simultaneous recording and display of Lx and the corresponding speech waveform recorded through a small electrode microphone (see Figs 3–8, 3–10, and 3–12).

2. The recorded Fx data from several minutes of conversation or reading aloud, plotted in a probability histogram, Dx (Distribution of excitation), and giving statistical information on the recorded data (see Figs 3–9, 3–11, and 3–13). The frequency scale is logarithmic to correspond to pitch perception and divided into 128 equally spaced "bins" between 30.52 Hz and 1,000 Hz (Abberton et al., 1989). Second- and third-order Dx plots give an indication of the amount of regularity in the sample in that they only admit the instances where two and three adjacent Fx samples, respectively, fall into the same frequency "bin." A harsh, "irregular" voice will show a dramatic decrease in the "total sample" carried into second and third-order Dx distributions. An estimate of voice regularity can be expressed as a proportion of the total sample carried into second and third order (% TS).

3. A different way of illustrating first-order Dx data is in the form of a Cx plot (cross plot of FX) or "scatter plot" (Figs 3–9, 3–11, and 3–13). This plots adjacent pairs of Fx values against each other. The frequency scales are divided into 64 logarithmically equally spaced "bins." The density of the markings reflect the number of occurrences of transitions at that position. A smooth, regular voice will show a dense, narrow plot along the diagonal, the length of which reflects the frequency range of the voice. A thick, "cigar"-shaped diagonal and significant scatter away from it, indicate an irregular, rough voice.

The Cx plot gives the visually most striking display of improvement in voice quality, and patients often ask for copies of their "before" and "after" speech or reading Cx plots as souvenirs.

Rationale for Using ELG for Visual Feedback in Voice Therapy

It is the ability to noninvasively and continuously, reflect glottal dynamics during a variety of phonatory tasks, that makes an EGG device an invaluable asset in the voice clinic. As the technique is easily applied, it can be used routinely for assessment of most voice patients, although there are a minority of patients for whom the technique is not suitable (Colton and Conture, 1990).

The use of laryngeal electrodes for recording and analysis of voice fundamental frequency allows the clinician to interact with the patient in a natural conversation, several min-

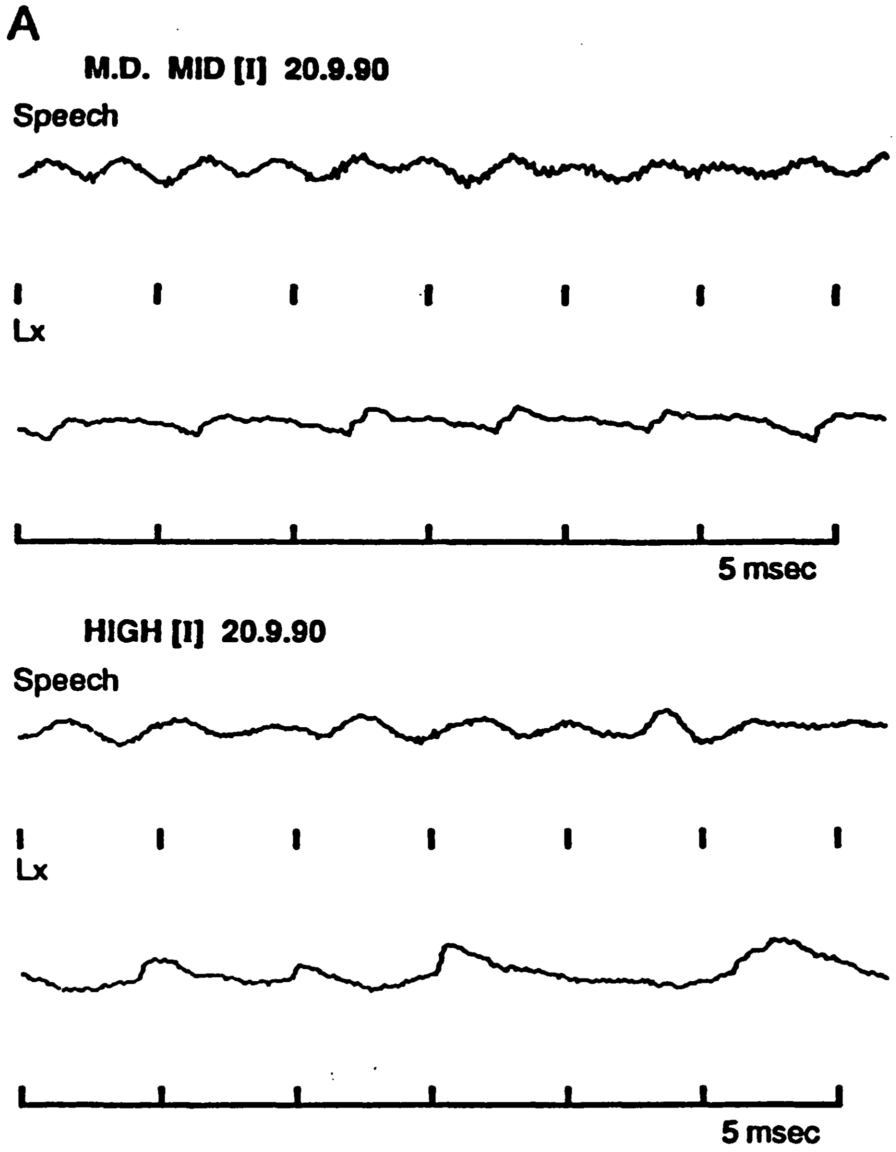


FIG 3–8.
A, Lx waveforms for /I/, demonstrating massive laryngeal effort. (*Continued.*)

utes in duration. Many of our patients are not confident readers, and a reading sample is not always obtainable, nor does it always give a true impression of the patient's habitual voice production. The electrodes are impervious to extraneous noise, and the recording can be carried out in an ordinary clinic environment, provided certain minimum standards are adhered to regarding sound attenuation. The patient is instructed not move his or her head, and is seated on a swivel chair to allow the person to turn toward the therapist without turning the head. During recording, the therapist can observe patterns of vocal fold contact during conversation and other phonatory tasks, as reflected in the shape of the Lx waveform.

A tendency to excessive laryngeal movement during speech, common in hyperkinetic dysphonias and in vocal fold palsy, gives rise to gross movement of the Lx baseline during speech. This can in itself be used for visual

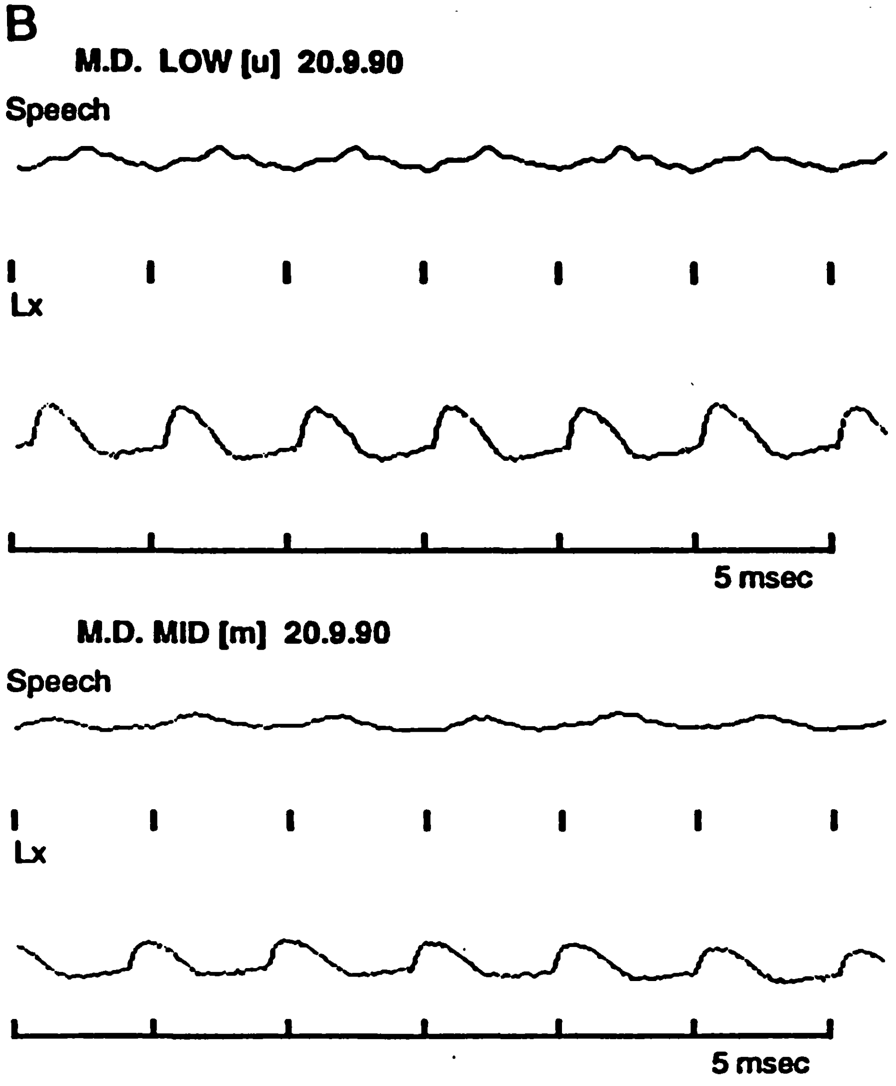


FIG 3–8 (cont.).
B, Lx waveforms for /u/ and /m/ with reduced laryngeal effort, reduced intensity, and abdominal breath support.

feedback, to heighten the patient's awareness of excessive laryngeal activity. The effect of reduction of this is clearly observable in the cessation of major gross baseline movement.

This feature, and other potential sources of "noise," are obvious problems in the scientific application of EGG and ELG (Childers et al., 1990; Colton and Conture, 1990; Rothenberg, 1991) and are often a reason given for not using the technique. However, the increasing body of quantitative research into EGG waveforms produced by healthy and dysfunctional subjects (Kelman, 1981; Colton and Conture, 1990; Orlikoff, 1991) provide confirmation of the validity of many EGG features that we have found useful in therapeutic application of ELG for visual feedback.

Many of the reservations expressed in the literature about the drawbacks of EGG as a research tool need not be seen as reasons for not using it in a clinical situation. There is less need for scientific stringency in the application of ELG as a therapeutic tool, for comparison of the same patient with himself at different stages in the therapeutic process, and for use as immediate visual feedback of glottal dynamics.

CASE STUDY: PATIENT K

Patient K was a 47-year-old woman who worked as a cashier at a staff restaurant. She started smoking 7 years before presentation and smoked 25 cigarettes a day. She was happily married and had two grown children and two grandchildren. Patient K was referred by her general practitioner to the otolaryngologist in April 1990. She reported progressive hoarseness over a 3-year period which had recently been aggravated by an upper respiratory tract infection. Indirect laryngoscopy showed an irregularity of the anterior end of her right vocal fold, and she was admitted for microlaryngoscopy on July 7, 1990. A hemorrhagic polyp was removed. A small nodule on the left vocal fold, opposite the polyp, was observed but not removed on this occasion. She was sent home on "voice rest" and referred for voice therapy. On October 10, what was described as a large polyp on the anterior end of her left vocal fold was removed. Her progress in voice therapy between Sept 17, 1990, and Nov 19, 1990, is described in the following sections.

Description of Voice

Patient K's voice quality on her first visit for assessment and treatment on September 20, 1990, was extremely "harsh" and "whispery" and produced with massive "laryngeal tension" in the terminology used by Laver et al. She had poor breath control and was continuously clearing her throat of excessive mucus. She was distressed over comments on her extremely low-pitched, rough voice that were continuously made at work.

The Lx waveforms for /i/ recorded on September 20, 1990 (Fig 3–8, A) were produced with massive laryngeal effort. There is some contact periodicity but very limited tissue contact at mid /i/, a gradual closure, and an extremely long closed phase. Each period corresponds to a very damped speech waveform. The attempt at a high pitch /i/ resulted in a extremely irregular pattern of closure.

Instruction, on the same occasion, to reduce laryngeal effort, to reduce intensity, and to use abdominal breath support, immediately resulted in more normal appearing periodic Lx waveforms on production of /u/ and /m/ at mid-pitch (Fig 3–8, B). There was increased tissue contact as evidenced by the increased Lx amplitude, and this time a long open phase indicating a "whispery" quality. This was used as proof of the patient's potential ability to produce a better voice quality with less effort, despite the continuing presence of a left vocal fold polyp. The weak speech waveforms of /u/ and /m/ are due to the rather inexpert therapist not having increased the gain on asking the patient to reduce intensity.

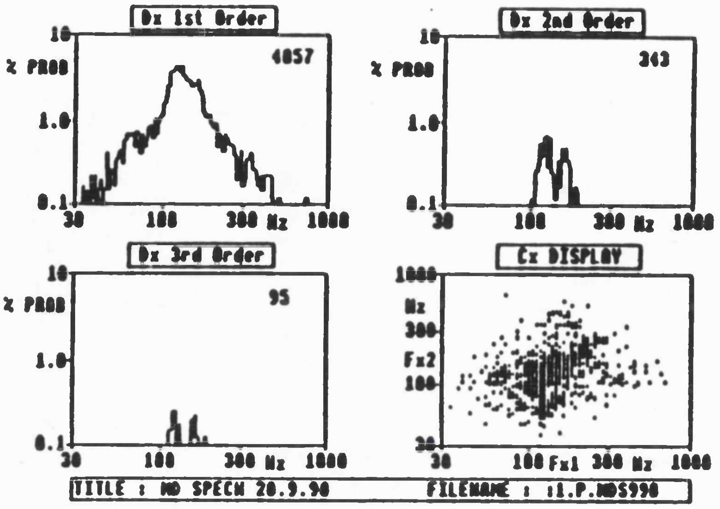
The Dx and Cx plots and fundamental frequency statistics of conversational speech on this occasion (Fig 3–9) show extremely abnormal distribution of Fx values. Central measurements, in this case the mode, is the only usable one, are extremely low for a 47-year-old woman.

There was a lot of "noise" in the signal because of gross laryngeal movement throughout the recording. Continuous need to clear mucus and possibly interference of the polyp on phonation must also account for some of the extreme scatter of Fx values. The second- and third-order Dx plots based on only 8.5% and 2.3% of the original sample (% TS), respectively, indicate an extremely small amount of regularity of vocal fold vibration in the sample, in second order within a 90% range of 104.3 Hz to 201.3 Hz. Another feature of the second-order Dx plot is its marked bimodality. One might hypothesize this reveals a "diplophonia" possibly due to different parts of the vocal folds vibrating at different frequency, which would not be implausible, knowing the presence of the vocal fold polyp. Some of this limited "regularity" of vibration within a similar frequency range is carried into third order. Perceptually the voice did not sound diplophonic, as there was so much other "noise" present.

The Cx plot reveals a very wide scatter along the diagonal throughout the range, illustrating a very limited amount of regularity of vocal fold vibration.

Description of Therapy Approach

The voice therapy offered in our department is preceded by a detailed explanation of the vocal mechanism using a model of the



% TS 8.5

% TS 2.3

STATISTICS TABLE

TITLE : MD SPECH 20.9.90 FILENAME : :1.P.MDS990

DISTRIBUTION TYPE	Dx 1st Order	Dx 2nd Order	Dx 3rd Order
SAMPLE TOTAL	4057	343	95
MEAN	133.5 Hz	137.2 Hz	144.9 Hz
MODE	122.9 Hz	126.4 Hz	122.9 Hz
MEDIAN	133.5 Hz	133.5 Hz	141 Hz
STANDARD DEVIATION	0.20 LOG-Hz	0.10 LOG-Hz	0.08 LOG-Hz
80% RANGE	75.1 237.3 Hz	113.2 185.3 Hz	119.6 190.6 Hz
90% RANGE	60.3 329.7 Hz	104.3 201.3 Hz	116.4 201.3 Hz

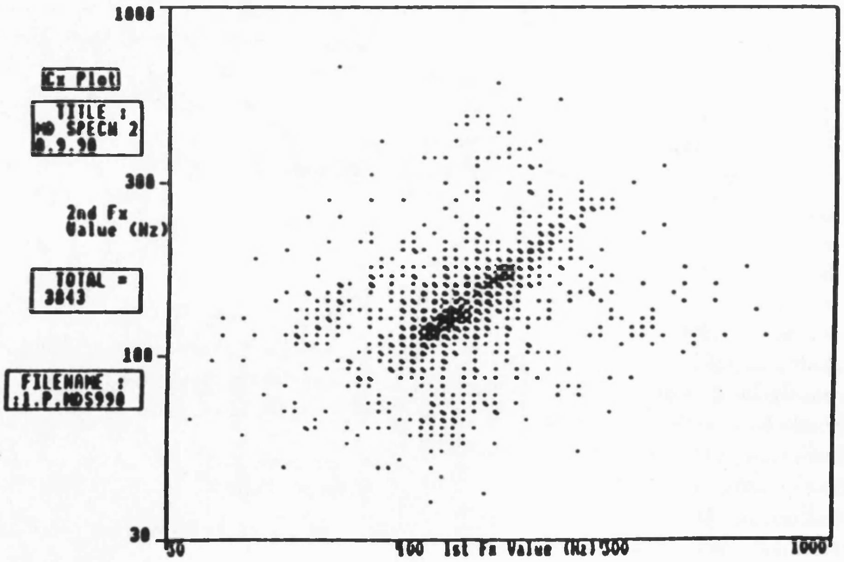


FIG 3–9.
Dx and Cx plots and fundamental frequency of conversational speech.

larynx and xeroradiograms of the vocal tract at rest and during phonation. The aim is to give the patient as much understanding of the mechanism as possible to "demystify" the voice problem. The therapy is loosely based on the "Accent Method" described by Smith and Thyme (1976, 1978). This approach emphasizes the control of subglottal pressure to "drive" the voice source to achieve increased speed and duration of vocal fold closure through the Bernoulli effect. This results in less air wastage and the acoustic effect of increased number and intensity of overtones (Smith and Thyme, 1978). The amplitude of the Lx waveform feeds back this change in glottal aerodynamics during voice exercises, as it responds to increased amount of tissue contact and reduction in force of adduction, as described previously. Accentuated breath pulses, controlled by abdominal muscles, are taught in a progressive series of exercises, which start with control of expiratory airflow during production of voiceless fricatives. The intensity of the sounds is varied by varying the expiratory force in different rhythms. Gradually, vowel sounds are introduced with soft attack and produced in increasing sequences and also in varying rhythmical patterns. The advantage with the method is the emphasis on reduction of laryngeal effort and concentration on the abdominal muscles as the power source for voice production. We also emphasized in this case the use of an open vocal tract. Chewing, humming, "playing on the lips," and "chanting" on a monotone were used to help the patient gain confidence and control over breath support for voice production, before embarking on gradually more speech-like exercises. The aim is to develop the control of a relaxed, comfortable, and flexible voice in speech and reading, appropriate in pitch, quality, and volume to the patient's sex and vocal needs.

Results of Therapy

Patient K was seen for voice therapy twice weekly and practiced recorded exercises at home between visits. The aim of therapy was to reduce volume, reduce laryngeal effort, and increase pitch using the breath control that was taught in clinic. Reassessment on October 3, 1990, shows improved control of vocal fold vibration in well-defined Lx waveforms showing improved tissue contact and steeply rising closing portions of the wave (Fig 3–10, A–C). The open phase occupies still the major portion of the cycle, however, resulting in a poorly differentiated acoustic output.

An interesting feature of the closing portion of the Lx waveform appears at low /i/ (Fig 3–10, C). The high amplitude indicates increased tissue contact, and the closed phase occupies more of the total period than at mid and high pitch. There is, however, a recurring irregularity at the same point in the rising portion of the waveform, which has the effect of delaying complete closure. This illustrates the claim of Colton and Conture (1990) that the closing time is affected by organic changes of the vocal folds and that ECG and ELG sometimes reflect this interference. Because of different modes of vibration of the vocal folds at different pitches and with different types of voice production, the effect of organic lesions on the Lx waveform varies. In this case it would seem that the vocal fold polyp mainly interferes with vibrations at low pitch. Both at mid and low pitch, however, there is also evidence of variability in tissue contact from one cycle to the next as seen in the slight variation in amplitude from one cycle to the next.

The Dx and Cx plots illustrating conversational speech on this occasion (Fig 3–11) show a marked improvement in regularity and range of Fx values. Central measurements are markedly higher than on the previous occasion, although the 90% range is still somewhat low and narrow. There is a marked increase in the % TS carried into second and third order, showing an increased amount of regularity of vocal fold vibration as a result of the changes made by patient K in her vocal habits. The best proof of her improvement was comments on her voice from her manager at work, who told her she was "beginning to sound like a human being." Her husband had commented that she did not "sound like a navvy off a building site any longer." These comments also serve to illustrate what she once sounded like to others, and the understandable distress it caused her.

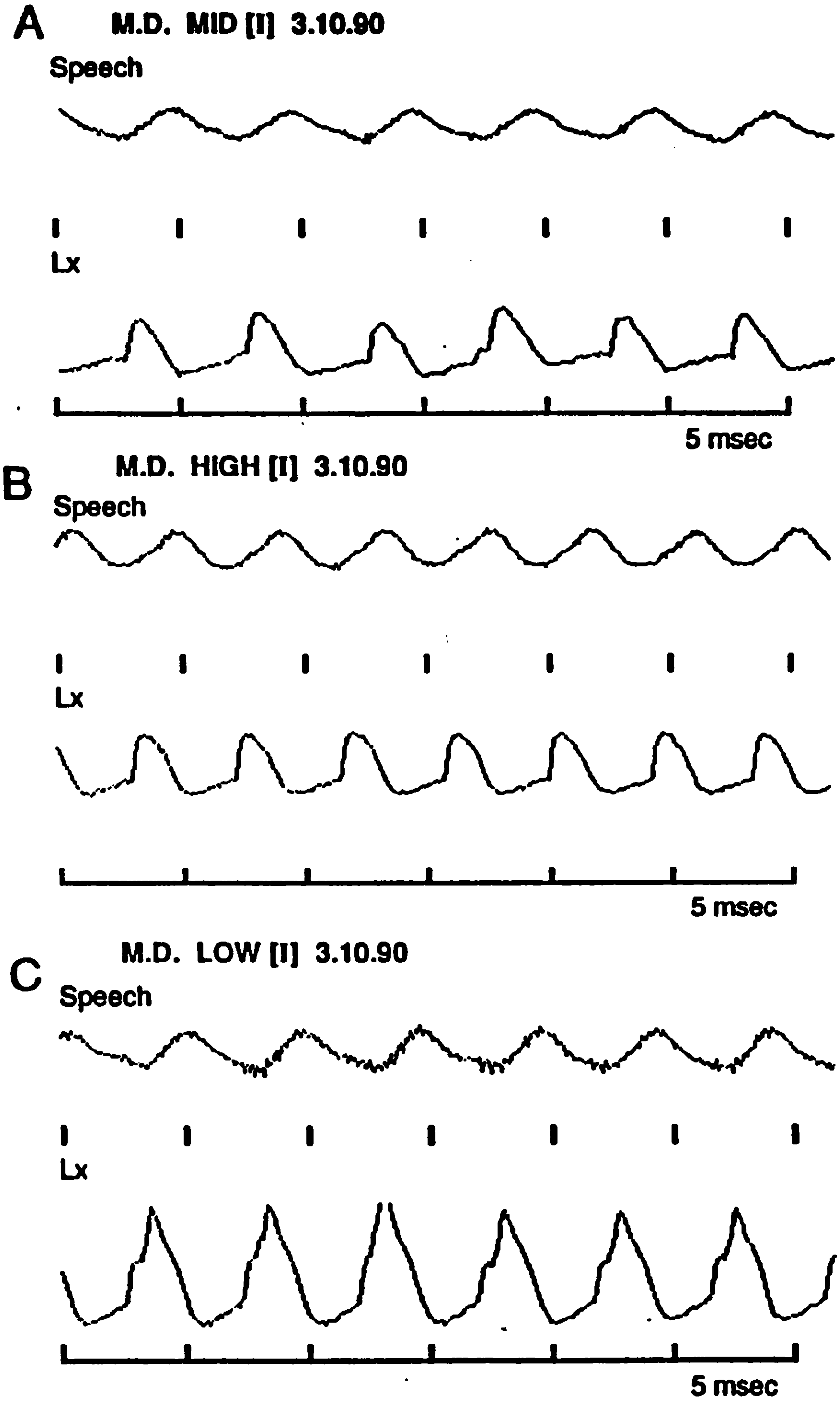
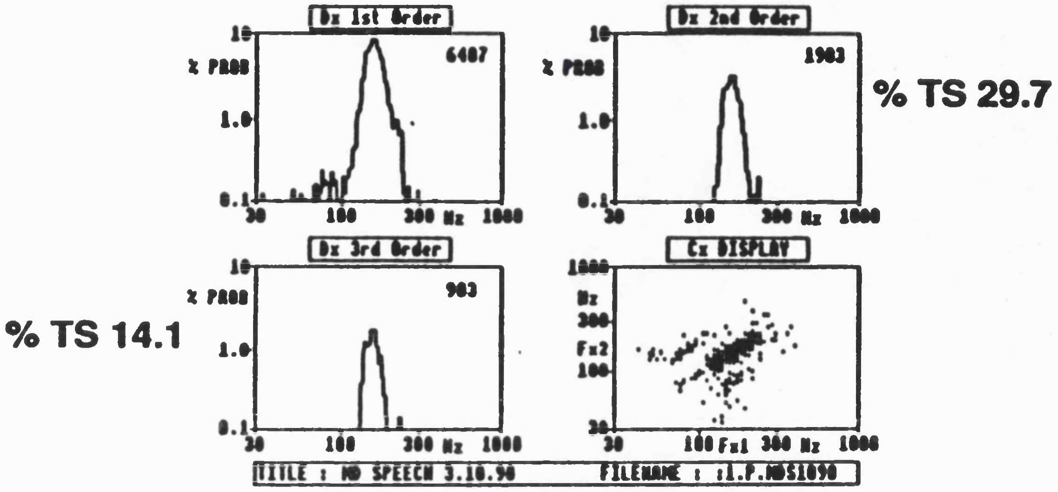


FIG 3-10.
A, Lx waveform for mid /I/ posttherapy. B, Lx waveform for high /I/ posttherapy. C, Lx waveform for low /I/ posttherapy.



STATISTICS TABLE

TITLE : MD SPEECH 3.10.90 FILENAME ::1.P.MDS1090

DISTRIBUTION TYPE	Dx 1st Order	Dx 2nd Order	Dx 3rd Order
SAMPLE TOTAL	6407	1903	903
MEAN	157.3 Hz	157.3 Hz	157.3 Hz
MODE	161.7 Hz	157.3 Hz	157.3 Hz
MEDIAN	161.7 Hz	161.7 Hz	161.7 Hz
STANDARD DEVIATION	0.11 LOG-Hz	0.05 LOG-Hz	0.04 LOG-Hz
80% RANGE	133.5 195.0 Hz	141.0 185.3 Hz	141.0 180.3 Hz
90% RANGE	113.2 218.6 Hz	137.2 195.0 Hz	141.0 190.6 Hz

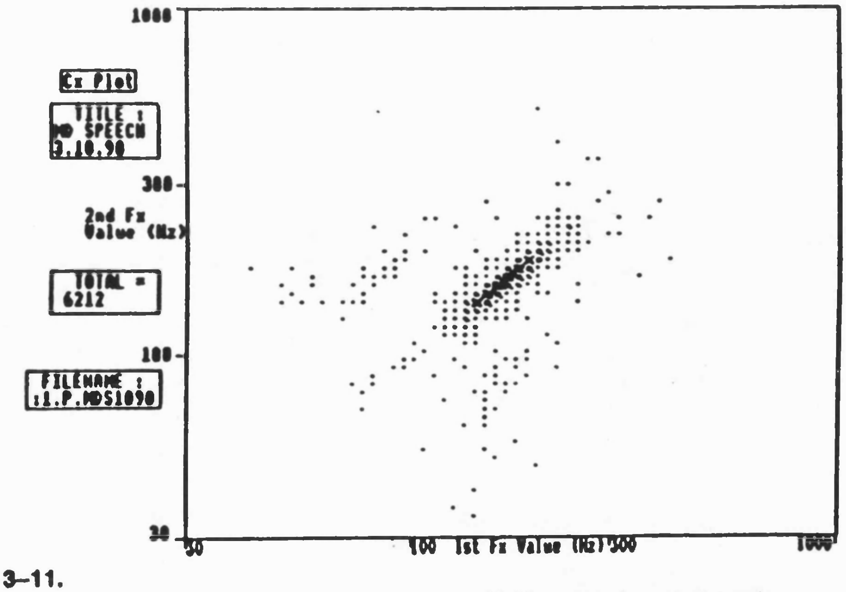


FIG 3–11.
Dx and Cx plots of conversational speech prior to vocal fold surgery for anterior polyp.

Patient K had another microlaryngoscopic treatment on October 10, 1990, for removal of the anterior polyp from her left vocal fold. Both vocal folds were observed to be edematous. Voice therapy continued after an initial postoperative period of voice rest. She stopped smoking at this time, and this was still the case on the last voice assessment, 5 weeks later on November 19, 1990.

On this occasion (Fig 3–12), the Lx plot shows overall, regular periodicity and higher fundamental frequency, well-defined closure, and longer closed phase compared with October 3, 1990 (Fig 3–10, A–C).

Fundamental frequency analysis on this occasion (Fig 3–13) shows a marked increase in central measurements and a higher and wider 90% range. A strong tendency to use hard glottal attacks and glottal stops, characteristic for certain London accents, which tends to lower the range is probably responsible for the bimodality of the Dx distribution.

The therapist remarked in her notes that the patient's voice sounded "lighter and was produced with considerably less laryngeal tension than previously." Her husband now described her as sounding "gentler."

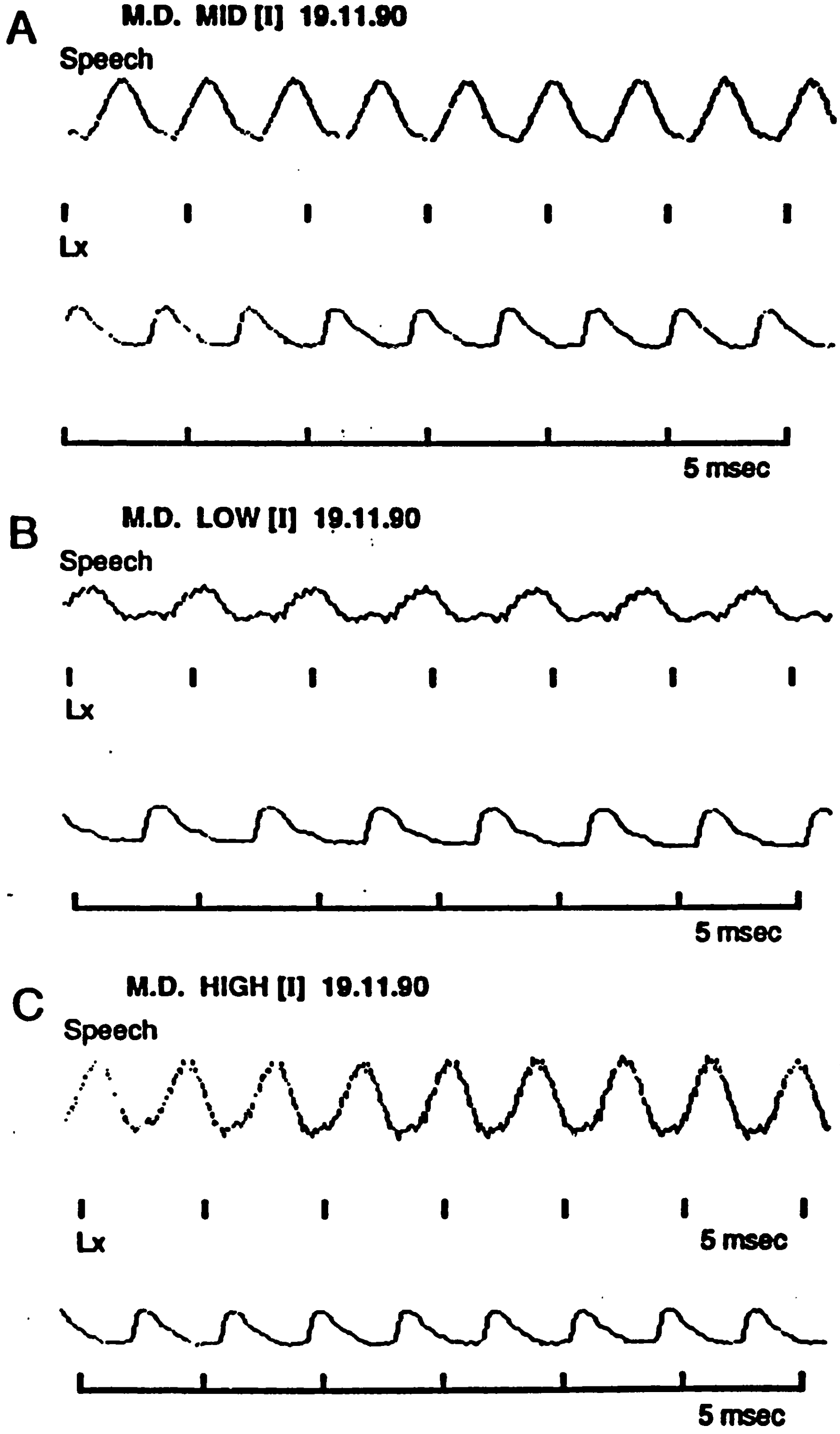
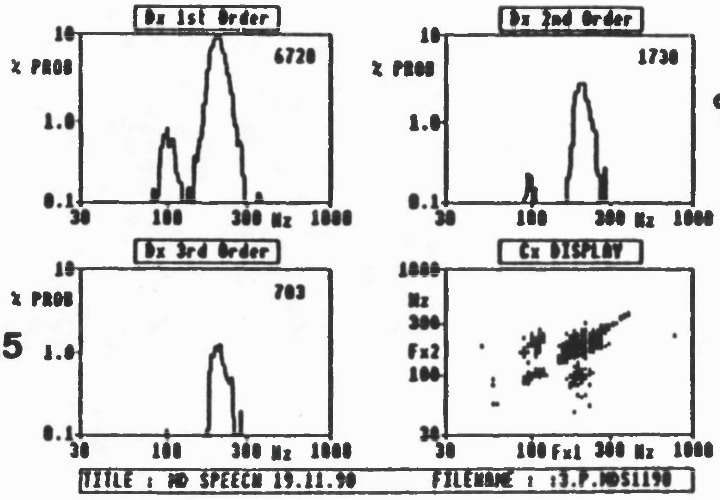


FIG 3–12.
Lx waveforms for mid, low, and high /i/ 6 weeks after surgery.



% TS 25.7

% TS 10.5

STATISTICS TABLE

TITLE : MD SPEECH 19.11.90 FILENAME ::3.P.MDS1190

DISTRIBUTION TYPE	Dx 1st Order	Dx 2nd Order	Dx 3rd Order
SAMPLE TOTAL	6720	1730	703
MEAN	195.9 Hz	201.3 Hz	206.9 Hz
MODE	206.9 Hz	201.3 Hz	212.7 Hz
MEDIAN	201.3 Hz	206.9 Hz	212.7 Hz
STANDARD DEVIATION	0.11 LOG-Hz	0.08 LOG-Hz	0.07 LOG-Hz
80% RANGE	166.2 237.3 Hz	105.3 243.9 Hz	105.3 250.7 Hz
90% RANGE	104.3 257.6 Hz	161.7 250.7 Hz	180.3 257.6 Hz

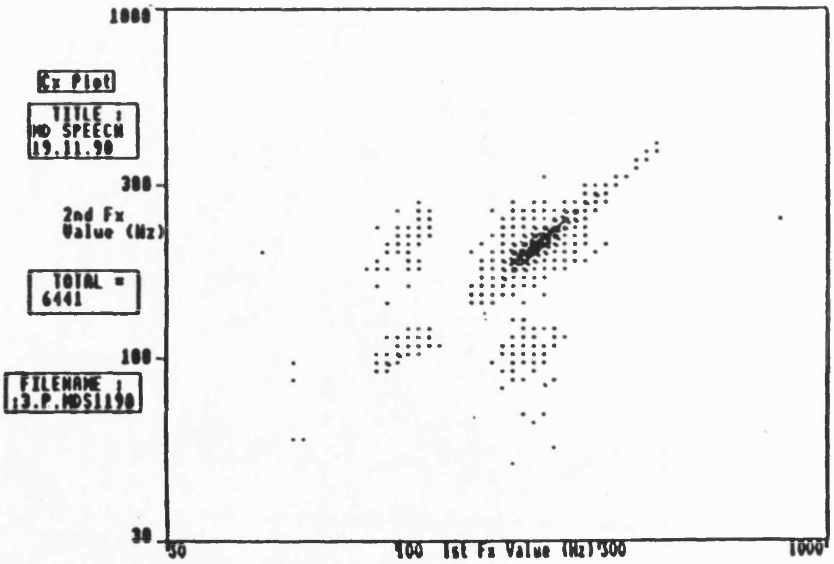


FIG 3–13.
Dx and Cx plots of conversational speech following surgery.

A STUDY OF VOICE QUALITY IN A GROUP OF IRRADIATED LARYNGEAL CANCER PATIENTS TUMOUR STAGES T1 AND T2.

by

Eva Ingrid Carlson

VOLUME TWO

APPENDICES

A thesis submitted for the degree
of Ph.D in the University of London.

1995

VOLUME TWO.

LIST OF APPENDICES.

APPENDIX 1 A

VOISCOPE STUDY- QUESTIONNAIRE.

May 1983

Name: Age: Date:

1 What is/was your job?

2. a) Are you working now? YES NO

b) Were any changes neccessary in your old job because of your voice ? YES NO

c) If so, give details:

d) Did you have to change your job? YES NO

3. a) If you are back at work, what does your job entail?

b) How long before you could return to work?

4. What effect did the radiotherapy have on your voice?

5. How much talking did you do while coming for treatment?

1	2	3	4	5	6	7
As little as possible			As much as usual			More than usual

6. Roughly how long did your voice take to recover after the end of treatment?

1	2	3	4	5	6	7
1 mth	2 mths	4 mths	6 mths	9 mths	12 mths	More than 12 months

7 Were you a smoker before your symptoms developed?

If so, how many cigarettes, oz of tobacco etc. did you smoke per day?

8. Are you smoking now? YES NO

If so how many/much do you smoke per day?

9. How would you rate your voice at the moment?

	1	2	3	4	5	6	7	
Back to normal								Very hoarse, an effort to speak at all times.

10 a How much do you use your voice in your daily life?

	1	2	3	4	5	6	7	
Very little because of discomfort/ have no one to talk to.			Normal use*					Talk a lot throughout the day, at work, at home socially.

* Daily talking to family, friends etc. but with periods of rest throughout the day.

10 b Is that more or less than before you had radiotherapy?

	1	2	3	4	5	6	7	
Less				Same as before				More

11. How much of a problem, if any, is your voice to you at the moment?

	1	2	3	4	5	6	7	
No problem								Major problem

12. **What would you say the problems with your voice are?**
Rate whichever symptom/s is/are appropriate in your case:

a) 1 2 3 4 5 6 7

Voice does not get 'tired' even after a lot of talking.

Voice gets 'tired' very easily.

b) 1 2 3 4 5 6 7

I'm able to 'sing' any tune.

I cannot 'sing' at all.

c) 1 2 3 4 5 6 7

I'm well able to shout

I'm unable to shout

d) 1 2 3 4 5 6 7

I'm able to talk in a noisy environment.

I'm unable to make myself heard over noise.

e) 1 2 3 4 5 6 7

I use the telephone with ease

Telephoning is impossible.

13. Describe any other symptom or problem you may have:

14. What do you try to do about your voice when it is rough e.g. keep quiet, clear your throat, drink something, suck something?

THANKYOU FOR YOUR COOPERATION!

APPENDIX 1 B

VOICE QUESTIONNAIRE.

Name: Age: Date:

1 a) What is/was your job?

b) Did you have to change/leave your job because of voice problems? YES NO Retired

2 a) Were you a smoker? YES NO

If so how many cigarettes/cigars/etc did you smoke per day?

b) Are you smoking now? YES NO

If so how many cigarettes/cigars/etc do you smoke per day?

Please mark the point on the scales below which you feel best describes the state of your voice at the moment. THERE ARE NO RIGHT OR WRONG ANSWERS!

3. How would you rate your voice at the moment?

	1	2	3	4	5	6	7	
Back to normal								Very hoarse, an effort to speak at all times.

4. How much do you use your voice in your daily life?

	1	2	3	4	5	6	7	
Very little because of discomfort/ have noone to talk to.				Normal use				Talk a lot throughout the day, at work, at home, socially.

5. How much of a problem, if any, is your voice to you at the moment?

	1	2	3	4	5	6	7	
No problem								Major problem

Mean P-score:

What would you say the problems with your voice are?
Rate whichever symptom/s is/are appropriate in your case.

6. Does your voice get 'tired' after a lot of talking?

1	2	3	4	5	6	7
No tiredness						Tires very easily

7. Are you able to sing or 'hum' a tune?

1	2	3	4	5	6	7
I'm well able to 'sing'.						I cannot sing at all.

8. Can you shout/call out?

1	2	3	4	5	6	7
I'm well able to shout						I'm unable to shout

9. Can you make yourself heard in a noisy environment?

1	2	3	4	5	6	7
No problem						I'm unable to make myself heard over noise.

10. Do you have any difficulty using the telephone?

1	2	3	4	5	6	7
I use the telephone with ease						Telephoning is impossible.

11. Do you experience dryness of the throat?

1	2	3	4	5	6	7
None						Extremely dry

12. To what extent do you think dryness is contributing to your voice problem?

1	2	3	4	5	6	7
Not at all						Major factor

THANKYOU FOR YOUR COOPERATION!

APPENDIX 2 A

SITE OF TUMOUR, HISTOLOGY, TUMOUR DOSE AND FRACTIONATION OF SUBJECTS WITH T1 TUMOURS OF THE VOCAL FOLDS.

Subj	Age	Stage	Tumour Dose	Fractions	Histology	Site
2	66	T1No	5.250	12/29	Poorly diff.	Anter. L v.c. inv anter. commissure
4	64	T1	5.300	13/30	Well diff. invasive	Anter.end Right v.c.
Field size 5 x 5			γ rays			
5	80	T1s		12/28	In situ scc.	Anter.comm.+ middle of the R v.c.
Field size 5 x 5			γ rays			
6	67	T1		14/31		Anter. 1/3 R v.c.
7	62	T1	4.800	11/28	Well diff.	Super.surface L v.c.
Field size 5.5 x 6			γ rays			
10	65	T1s	5.200	12/29	In situ scc.	Anter. 2/3 R v.c.
Field size 6 x 6			γ rays			
14	63	T1	5.960	18/46	Invasive scc.	Left v.c.
Field size 6.5 x 7			γ rays			
16	54	T1	5.300	12/28	Mod. diff.	Right v.c.
Field size 6.5 x 6			X rays			

APPENDIX 2 A

Subj	Age	Stage	Tumour Dose	Fractions	Histology	Site
20	57	T1	5.200	12/27	Well to mod. diff.scc.	Anter. 2/3 R v.c.
Field size 6 x 6			γ rays			
22	49	T1a	5000	12/29		R v.c. in situ, granuloma biopsied
23	61	T1NoMo	5750		Lv.c.mod. diff. inv. R v.c. inv.	Horseshoe lesion of both cords extend. to aryt. on L+anter 2/3 R
25	53	T1	5500	20/28	Mod.diff.scc.	Poster.1/2 L v.c.
Field size 6.5 x 6.5			γ rays			
26	59	T1	5500	20/32	Well diff. scc.	Anter.R v.c., superfical tumour dissect. off R v.c.
Field size 5 x 5			γ rays			
27	69	T1	5500	20/29	Mod.diff.scc.	L v.c.
Field size 5.5 x 5			γ rays			
28	58	T1a	5500	20/28	Well diff. scc.	R Vocal cord, keratotic lesion.
Field size 5 x 5			γ rays			
30	70	T1	5500	20/26	Mod. diff.scc.	Anter. L v.c.
Field size 5.5 x 5			γ rays			
31	48	T1b	5500	20/28	Well diff.scc. invasive	L v.c. ant.comm. anter.1/3 R v.c
Field size 5 x 5			γ rays			

APPENDIX 2 A

Subj	Age	Stage	Tumour Dose	Fractions	Histology	Site
32	64	T1	5500	20/28	Mod.diff.scc.	Ant. L v.c. across anter. commissure
Field	size	5.5 x 5	γ rays			
36	65	T1	5790	20	Well diff. scc.	Anter.mid L v.c
Field	size	5 x 5 6w x 5	CO^{60} Teletherapy			
37	54	T1	5500	20/31	Well diff.	Anter-Mid. Rv.c.
Field	size	6w x 5	CO^{60} Teletherapy			
38	73	T1		12/28	Mod.diff.scc	Mid 1/3 R v.c
Field	size					
39	68	T1	5789	20/28	Mod. diff. inv. scc.	Whole R v.c. up to ant.com.
Field	size	6.5 x 5				
40	53	T1	5500	20/30		R v.c.
Field	size	5.5 x 5				

APPENDIX 2 B

SITE OF TUMOUR, HISTOLOGY, TUMOUR DOSE AND FRACTIONATION OF SUBJECTS WITH T2 TUMOURS OF THE VOCAL FOLDS.

Subj	Age	Stage	Tumour Dose	Fractions	Histology	Site
3	57	T2	5.200	12/32	Well diff.	Anter.2/3 R v.c. extension into R ventr.
Field size 5 x 6			γ rays			
8	56	T2	5.200	12/28		Anter.L v.c. slight subglottic extension
Field size 7 x 7			γ rays			
11	68	T2		12/27		Anter.L v.c. anter. comm.sl. subglott ext.
Field size 5.5 x 6			γ rays			
12	42	T2	6.600	12/19	Ca in situ no invasion	Left v.c. Right subglottis
Field size 9 x 12			4.000	20/26	γ rays	
Field size 5.5 x 7			2.500	12/19	γ rays	
13	39	T2	5.200	12/35	Mod. diff.	Right v.c.
Field size 7 x 7			γ rays			
17	77	T2		12/29		R v.c. with sub- and supraglott. extension.
21	39	T2	3.600	6/18	Poorly diff.	Infer.surface R v.c. subglottic extension
Field size 7.5w x 11						

APPENDIX 2 B

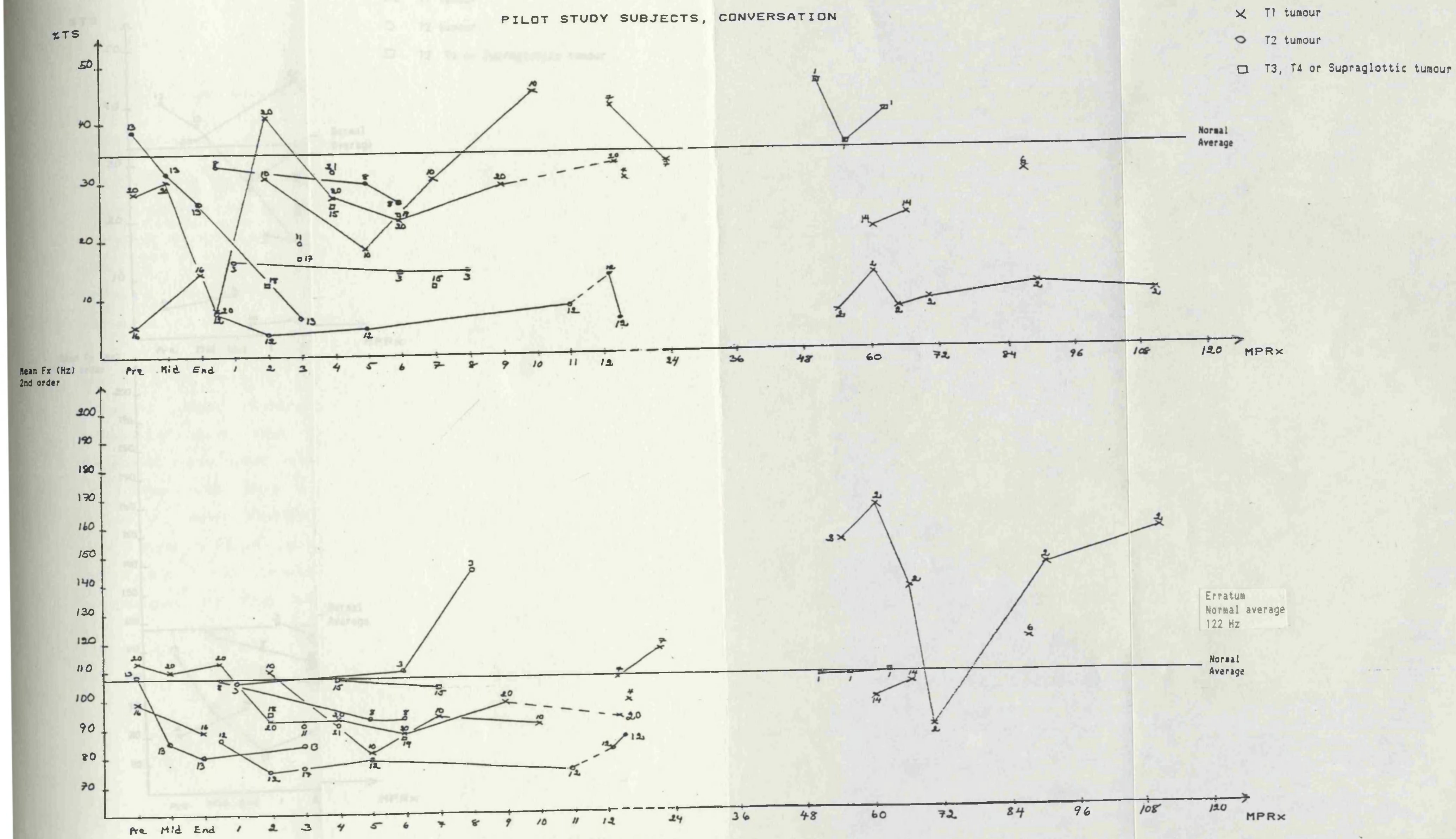
Subj	Age	Stage	Tumour Dose	Fractions	Histology	Site
24	65	T2	5.500	20/30	Mod.diff. scc.	R v.c. into ventric. acro. ant. comm.+ L voc proc. into subglottis
Field size 6.5 x 6.5 γ rays						
29	50	T2N1	6.400	32/49	Mod.diff. scc.	L v.c.extend. into subgl. + towards poster. comm.
Field size 1w x 9						
		1w x 9	4.400	22/28		
Field size 5.5 x 6			2.000	10/14		
33	68	T2		12/20	Poorly diff.+ areas of in situ scc.	All L v.c. + ant. comm. thickened R ventric. band
Field size 7w x 7						
34	78	T2	5.500	20/28	Well diff. invasive	Ant. 1/3 L v.c. extend. subglott. and ant.comm.
Field size 6 x 6.5 γ rays						
35	65	T2	5.500	20/28	Poorly diff. invasive	R. vocal cord
Field size 5w x 5						

APPENDIX 2 C

SITE OF TUMOUR, HISTOLOGY, TUMOUR DOSE AND FRACTIONATION OF SUBJECTS WITH T3, T4 AND SUPRAGLOTTIC TUMOURS OF THE LARYNX.

Subj	Age	Stage	Tumour Dose	Fractions	Histology	Site
1	62			6/19 (in hyperbaric Oxygen)	Poorly diff.	Supraglottic
15	69	T4		14/30		Right v.c. fixed
18	67	T3	6.500	30/43	Mod. diff.	Fixed L v.c. subglottic extension.
19	67	T1N1Mo	6.800	20/28		Supraglottic

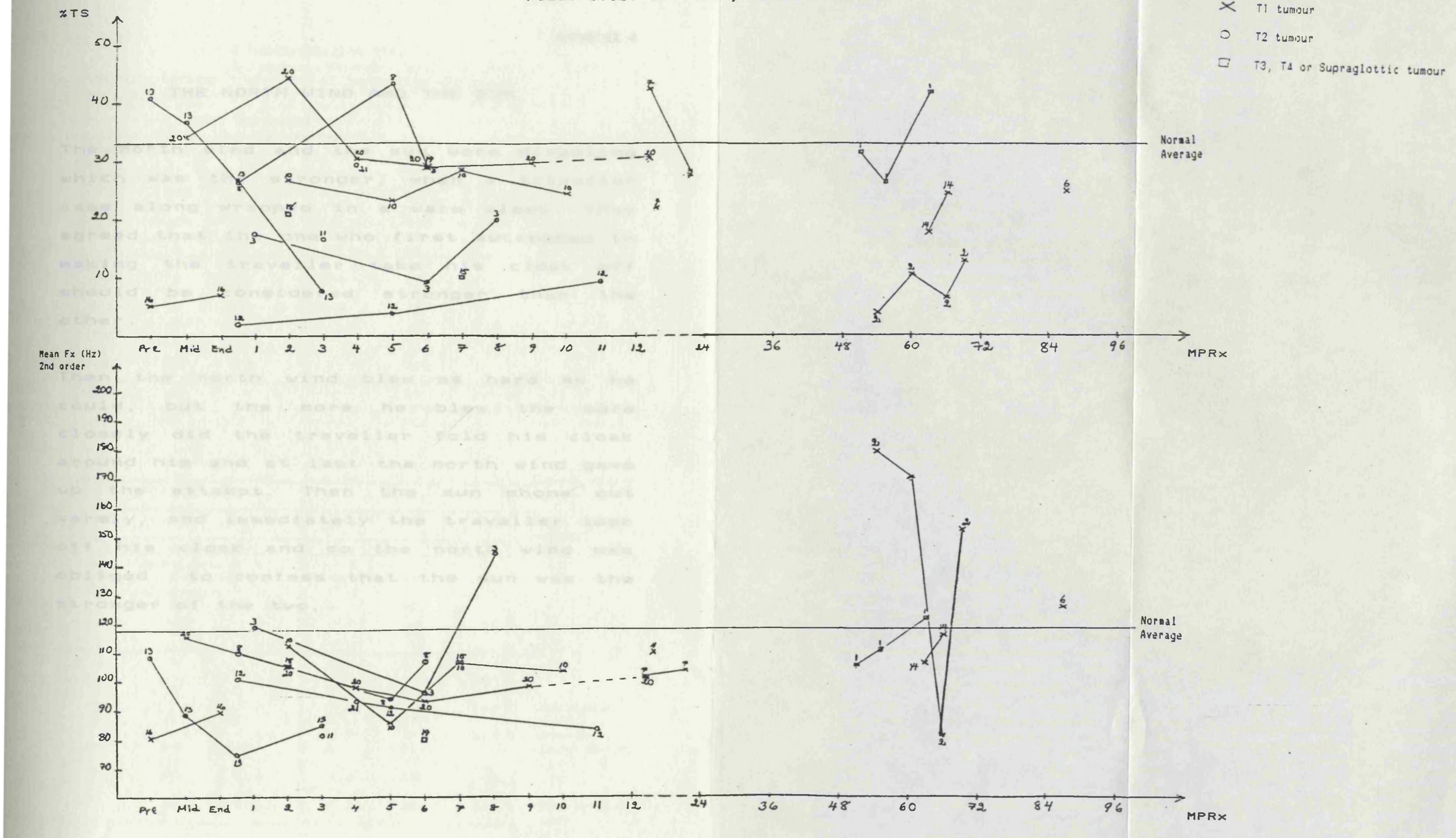
PILOT STUDY SUBJECTS, CONVERSATION



Erratum
Normal average
122 Hz

Dashed line indicates change in time scale from monthly to twelve monthly intervals.

PILOT STUDY SUBJECTS, READING ALOUD



Dashed line indicates change in time scale from monthly
to twelve monthly intervals.

THE NORTH WIND AND THE SUN.

The north wind and the sun were disputing which was the stronger, when a traveller came along wrapped in a warm cloak. They agreed that the one who first succeeded in making the traveller take his cloak off should be considered stronger than the other.

Then the north wind blew as hard as he could, but the more he blew the more closely did the traveller fold his cloak around him and at last the north wind gave up the attempt. Then the sun shone out warmly, and immediately the traveller took off his cloak and so the north wind was obliged to confess that the sun was the stronger of the two.

APPENDIX 5

TABLE 1

VPA ratings of T1 subjects - treated with 3F/week.

		PHONATION TYPE				
Subject No	MPRx	Harshness	Whisper	Creak	Laryngeal Tension	Other
2	53	5	5	5-6	5	Very 'dry'
	59	5	5	6	6	
	65	5	5	6	6	Very 'dry' 'creaky'
	70	4-5	5	6	6	
	71	4-5	5	6	6	
	89	5	5	6	6	
	112	-	3-4	5-6	5	Dry, creaky
4	16	-	1	3	3	
	107	-	2	4	4	Strong RP Never any prob. w. voice.
	119	2	3	2-3	2	
5	96	1	2-3	3-5	3-5	
6	86	2	4	5	5	
7	16	4	5	3	4	
	22	2-3	2-3	3	2	
	100	4	5	6	4	Sounds v. dry, limited range
10	2	2	3	4	3	
	5	2	2	5	3	
	7	1	2	4	4	
	10	1	2-3	3-4	2-4	
	61	-	2	2-4	3	
	94	-	2	4	3	

ctd. Table 1

		PHONATION TYPE				
Subject No	**MPRx**	**Harshness**	**Whisper**	**Creak**	**Laryngeal Tension**	**Other**
14	61	4	3	5	5	
	63	2	3	5	4	
	120	5-6	4	-	4	
	149	*5-6*	*5*	-	*3-5*	*Intermitt. aphonia*
	156	*5*	*6*	-	*5-6*	*Extremely dry*
16	*pre*	*3*	*4*	*4*	*3*	
	mid	*4*	*4*	*2*	*4*	
	end	*5*	*4*	*4*	*4*	
	2	-	*4*	*4-5*	*2*	
	3	-	*3*	*4-5*	*4*	
	61	*2*	*4*	*5-6*	*3*	*On diuretics 'dry' throat*
	63	*3*	*3*	*3*	*2*	
	88	2-3	4	4	3	Occas. too low Fx
	94	-	4	5	4	Poor breath support
20	pre-biopsy	6	4	4I	5	
	post-biopsy	5	3	-	3	
	mid	1	4	2	Lax 2	
	end	3	5	4	Lax 3	
	2	4	4	5	Lax 2	
	4	3	5	-	Lax 3	Monotonous soft
	6	4	5	2-3	Lax 2	Sounds v. 'dry' harsh
	9	4	4	1	2	
	14	3	5	-	Lax 4	Monotonous soft

ctd. Table 1

		PHONATION TYPE				
Subject No	**MPRx**	**Harshness**	**Whisper**	**Creak**	**Laryngeal Tension**	**Other**
ctd.						
20	92	5	4-5	-	4	
22	*5*	-	*4-5*	*5*	*4-5*	*Very creaky*
	6	-	*4-5*	*2-3*	*3*	*Very low Fx*
	9	*2*	*3-4*	*3*	*3*	
	12	2	4	3	3	
	13	2	4	3	Varying Lax 2 - Tense 3	
	17	-	4-5	3-4	-"-	
	57	-	-	5	4-5	
23	pre	3	5	-	5	
	1	4	4	-	4	
	3	4	3	3	2-3	
	5	3	3	4	5	
38	168	6	6	6	6	Smoking 15-20/day, Intermitt. aphonia
T1	178	5	5-6	-	5-6	- " - Hunched posture

TABLE 2

VPA Ratings of T1 subjects treated in 5F/week.

		PHONATION TYPE				
Subject No	MPRx	Harshness	Whisper	Creak	Laryngeal Tension	Other
25	*9*	*3-4*	*2*	*5*	*3*	
	11	*4-5*	*4-5*	*5I*	*4*	
	12	*3*	*3*	*4*	*4*	
	14	*1*	*1*	*3I*	*1*	
	16	5	5-6	4	3	
	20	5-6	5	4-5	4	Sounds dry
	23	6	5	-	Variable Tense 3- Lax 2	Smoking 2-3 small cigars
	38	6	5	-	4	Extremely dry 10-20/day
	50	6	5	-	3	Smoking a pipe
26	2	4	4-5	4	4	
	4	2	3-4	4	4	
	42	4	4	6	5	
	48	4-5	2	4	4	
	61	6	3	4	4	
27	18	-	2-3	2	Lax 1	
	39	2	2-3	3	3-4	
	56	-	4	4-5	4-5	Laryngeal mvmt. ++
	63	-	2	5-6	5	Speaks very fast.

ctd. Table 2

Subject No	**MPRx**	**PHONATION TYPE** **Harshness**	**Whisper**	**Creak**	**Laryngeal Tension**	**Other**
28	pre	4-5	5	4	4-5	
	2	5	4-5	5	5	Extremely low Fx
	3	3	2	3	3	
	4	3	3	3-5	2	
	34	4	3-4	4	4	
	46	4	3	5	5-6	
	58	-	4	5	3	
30	6	3-4	4-5	-	4	Smoking 5/day
	9	4-6	5	-	4-5	5-10/day
	10	4	4-5	-	4-5	
	12	5	4-5	-	5	
	16	4-5	5	-	5	10/day
	29	6	6	-	5	20/day
	36	6	5	-	5	10-15/day
	40	5	4	-	5	Voice better 10-15/day
31	pre	3	4-5	-	3	
	2	-	-	4-5	2-3	Tucks chin in
	6	-	-	3-4	3	
	10	-	2	4-5	5	
	17	-	2-3	5	4	
	26	-	2	5	5	
	37	-	1	4	3	Developed RA in all joints.

ctd. Table 2

		PHONATION TYPE				
Subject No	NPRx	Harshness	Whisper	Creak	Laryngeal Tension	Other
32	5	-	-	-	-	
	10	4-5	4	3	4-5	
	37	-	2	5-6	4-5	
	48	-	2	3-4	4-5	
36	37	-	5	3	Variable Tense 3 Lax 2	Voice resonant, but restrained
	44	-	5	4	3	
	48	-	4	4	3	
	63	3	4	3-4	2	
37	1	-	3-4	5-6	5-6	
	5	-	5	6	5-6	Has a cold
	11	-	4	6	6	Consistent creak ? Ventricular bands
39	3	6I	6	-	6	Intermitt. modal voice
	12	4-5	5	-	4-5	Modal voice
40	*9*	*4*	*5*	-	*3-4*	*Too loud*
	10	*4*	*4*	-	*2*	
	16	4	4	-	2	

TABLE 3

VPA ratings of T2 subjects

		PHONATION TYPE				
Subject No	**MPRx**	**Harshness**	**Whisper**	**Creak**	**Laryngeal Tension**	**Other**
3	1	3	4	3	Lax 2	Pitch breaks
	6	4	3	5	4	
	8	5	4	4	5	
8	2/52	4	5	1	4	
	5	3	4	2	2	
	6	1-2	4	2I	1	
	71	1	2	3	1	
	84	2	5	-	Lax 2	Very 'dry' Mucus ++
11	3	3	4	5-6	5	
12	2/52	4-5	5-6	4	3-4	Very low Fx
	2	6	5	-	5	Ventricular band voice
	5	6	6	-	6	Lower and harsher.
	11	5	4	6	6	
	13	6	6	-	4	Ventricular Less effort
	15	5	3	6	6	
	62	-	5	6	5	
	100	5	5	-	4-5	Consistent Ventricular

ctd. Table 3

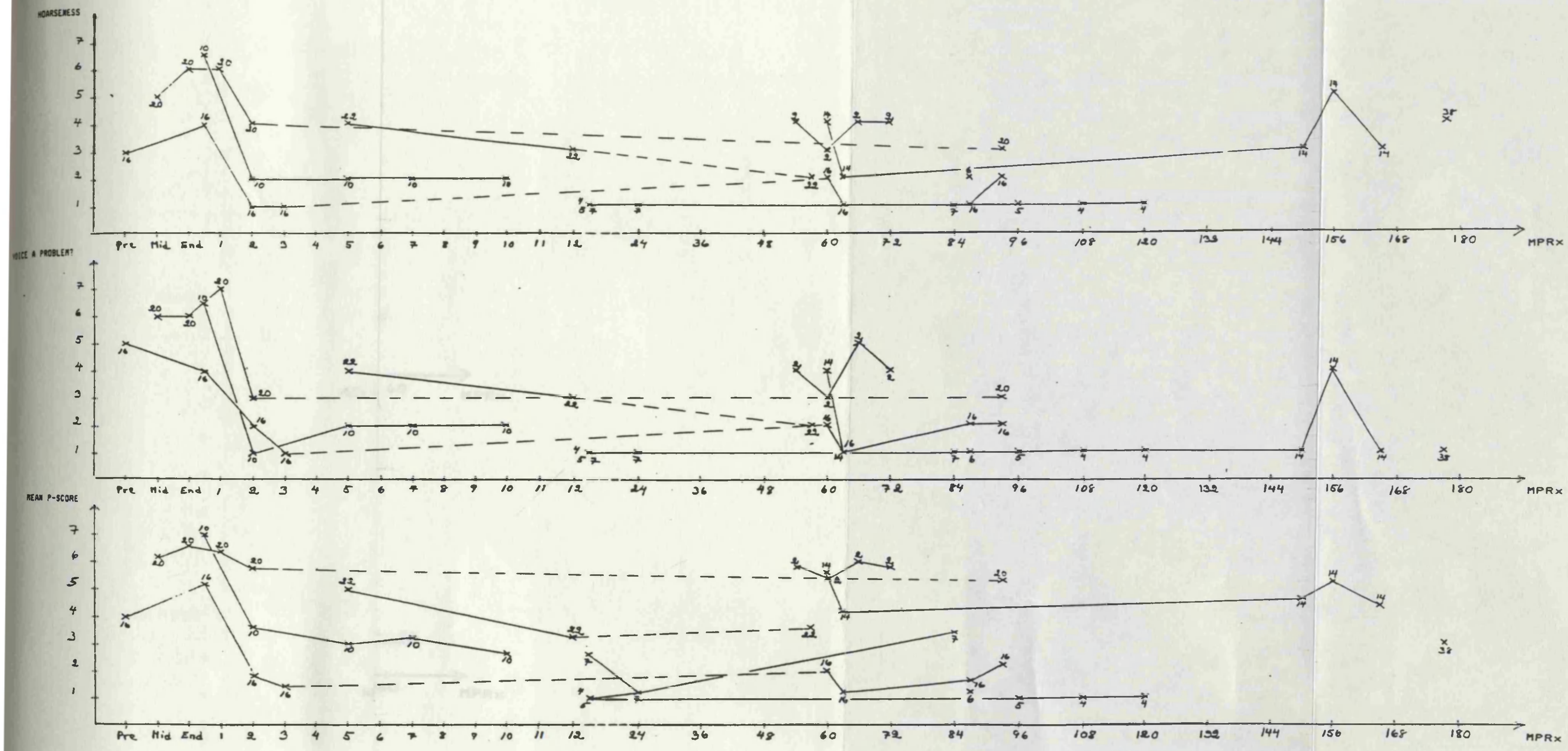
		PHONATION TYPE				
		Harshness	**Whisper**	**Creak**	**Laryngeal Tension**	**Other**
Subject No	**MPRx**					
13	pre	6	5	-	4	
	mid	6	5	2	4	Intermitt. aphonic
	2/52	6	6	4	5	Smoking 20/day
	2	6	6	-	5	No modal voice. 10-20/day
	9	5	6	-	5	Smoking 5-10/day
17	3	4	5	6	4	
21	4	3-4	3-4	4	4	
	55	4	3	4	4	
	63	4	4	2-3	3	
	64	3	5	4-5	3	
	94	3	5	2	2-3	
	101	4	3	3	1	Hyponasal rhinitis+ wide range
24	11	5	5	-	5	Diplophonic
	12	4	6	-	3	Voice less effortful
	23	6	6	-	6	Intermitt. modal voice
	25	6	6	-	6	
	44	6	6	-	6	Inspiratory stridor, Intermitt. aphonia, effort ++.

ctd. Table 3

		PHONATION TYPE				
		Harshness	Whisper	Creak	Laryngeal Tension	Other
Subject No	**MPRx**					
29	2	5	3	4-5	5	Feeling chesty, dehydrate
	3	5	3	5	5	
	6	4-5	3-4	5	4-5	Smoking 30/day
	11	-	2	6	6	Drinking
	34	-	-	6	6	Extremely creaky v.
	42	-	2	6	5-6	
33	3	-	4-5	4-5	5-6	
	7	2	4	5	3-4	'Dry'. Ne recurren
	13	4-5	4	3-4	2-5	Tucks chi in when speaking
	22	-	4	4-5	4	
	25	3	4	5	lax 3	Exaggerat pitch and intensity variation
	28	5	5	4-5	5	
	31	3	4-5	-	3	Occ.aphon poor br.s
	35	4	3	4	4	
34	67	4-5	4-5	-	4	Sl. Deaf
35	pre	5	5	6	5-6	Italian accent ++
	2	6	5	-	5	
	3	6	4-5	-	5-6	Too loud
	5	5-6	4-5	-	5	Too loud
	42	6	4	-	6	Too loud ?diplophon

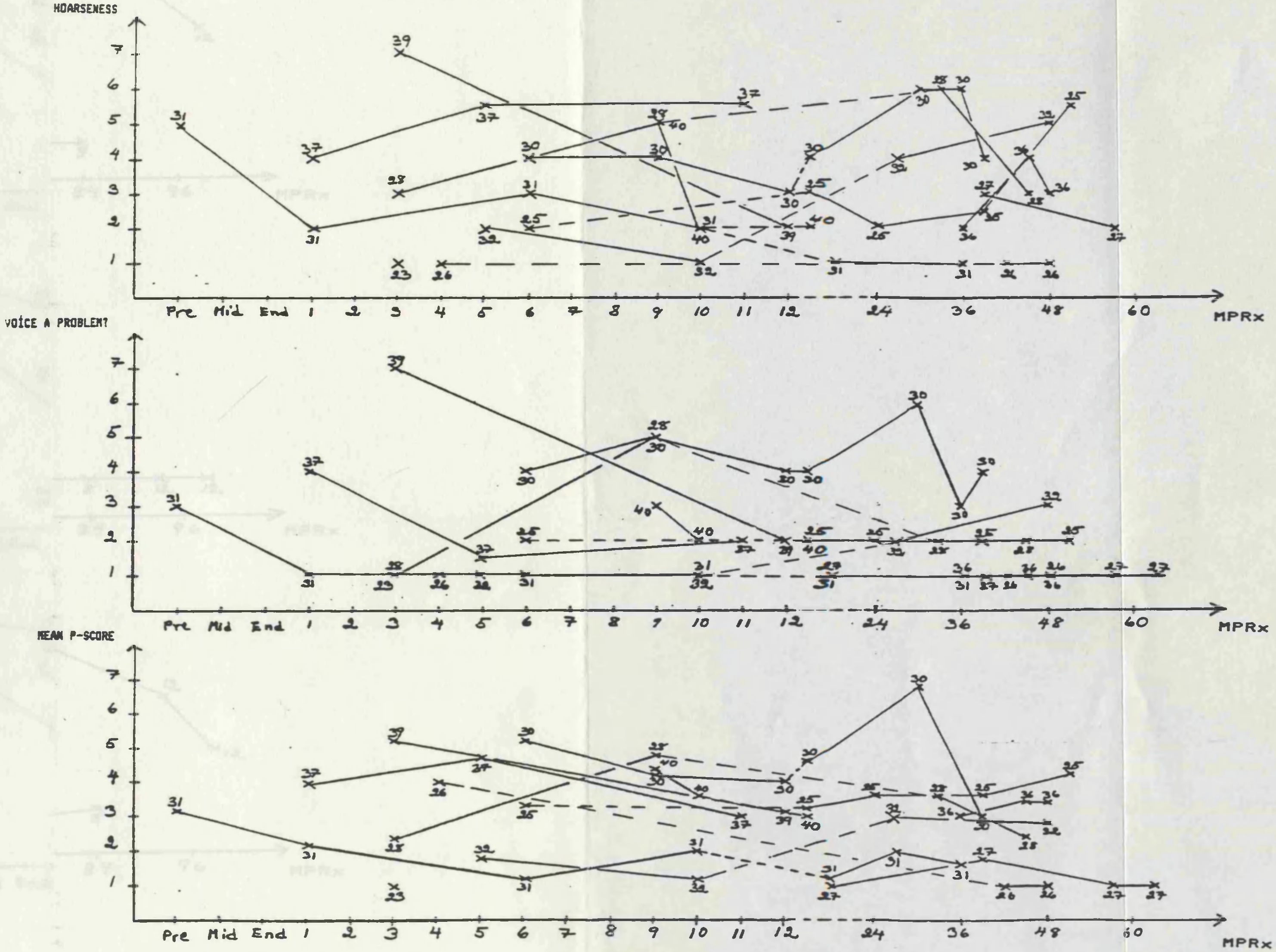
APPENDIX 6 A

T1 SUBJECTS 3F/WK, SELF - RATINGS ON QUESTIONNAIRE.



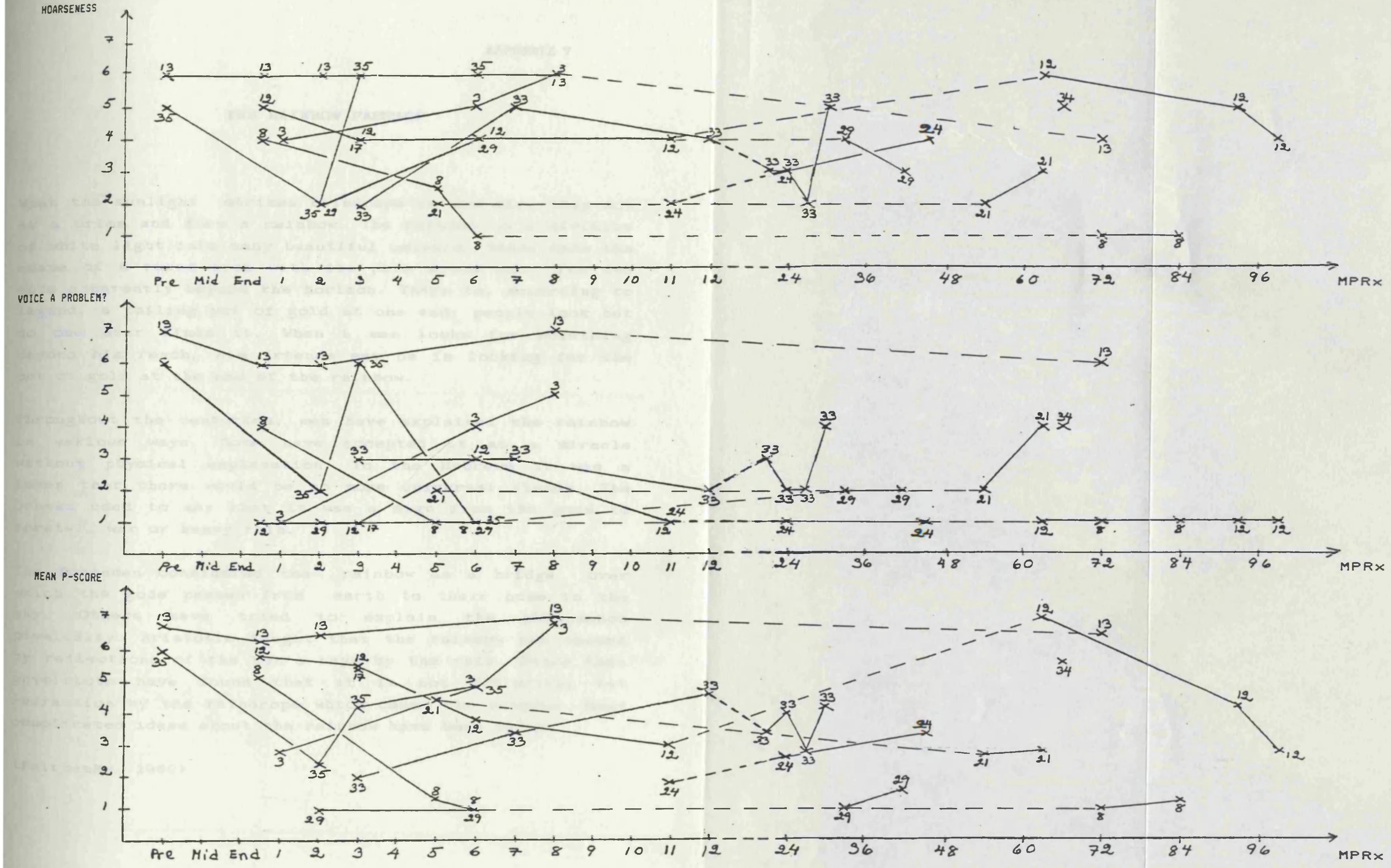
Dashed line indicates change in time scale from monthly to twelve monthly intervals.

T1 SUBJECTS 5F/WK, SELF - RATINGS ON QUESTIONNAIRE.



Dashed line indicates change in time scale from monthly to twelve monthly intervals.

T2 SUBJECTS, SELF - RATINGS ON QUESTIONNAIRE.



Dashed line indicates change in time scale from monthly
to twelve monthly intervals.

THE RAINBOW PASSAGE

When the sunlight strikes raindrops in the air, they act as a prism and form a rainbow. The rainbow is a division of white light into many beautiful colours. These take the shape of a round arch with its path above and its two ends apparently beyond the horizon. There is, according to legend, a boiling pot of gold at one end: people look but no one ever finds it. When a man looks for something beyond his reach, his friends say he is looking for the pot of gold at the end of the rainbow.

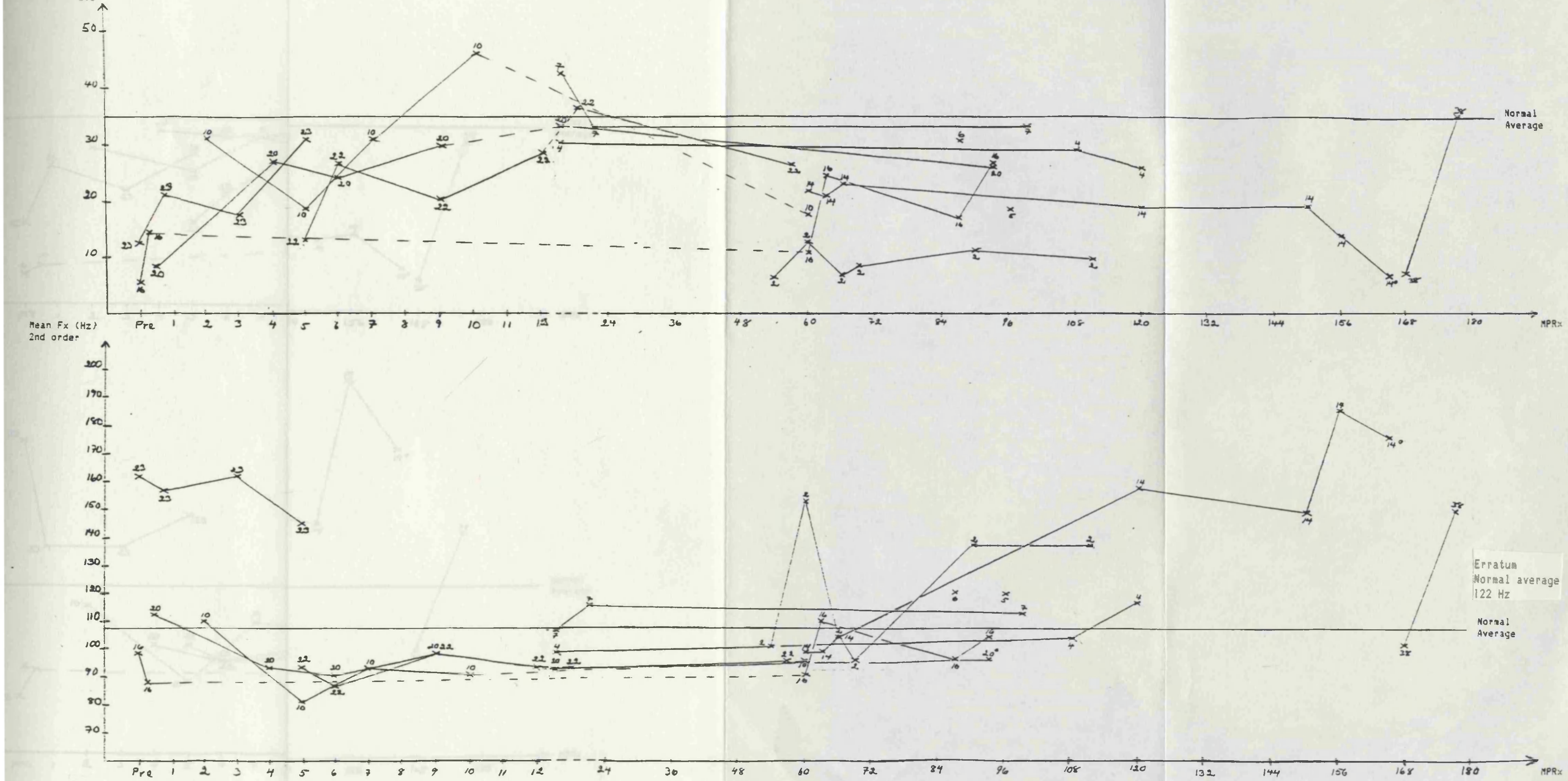
Throughout the centuries, men have explained the rainbow in various ways. Some have accepted it as a miracle without physical explanation. To the Hebrews it was a token that there would be no more universal floods. The Greeks used to say that it was a sign from the gods to foretell war or heavy rain.

The Norsemen considered the rainbow as a bridge over which the gods passed from earth to their home in the sky. Others have tried to explain the phenomenon physically. Aristotle taught that the rainbow was caused by reflections of the sun's rays by the rain. Since then physicists have found that it is not reflection but refraction by the raindrops which cause the rainbow. Many complicated ideas about the rainbow have been formed.

(Fairbanks, 1960)

APPENDIX 8 A

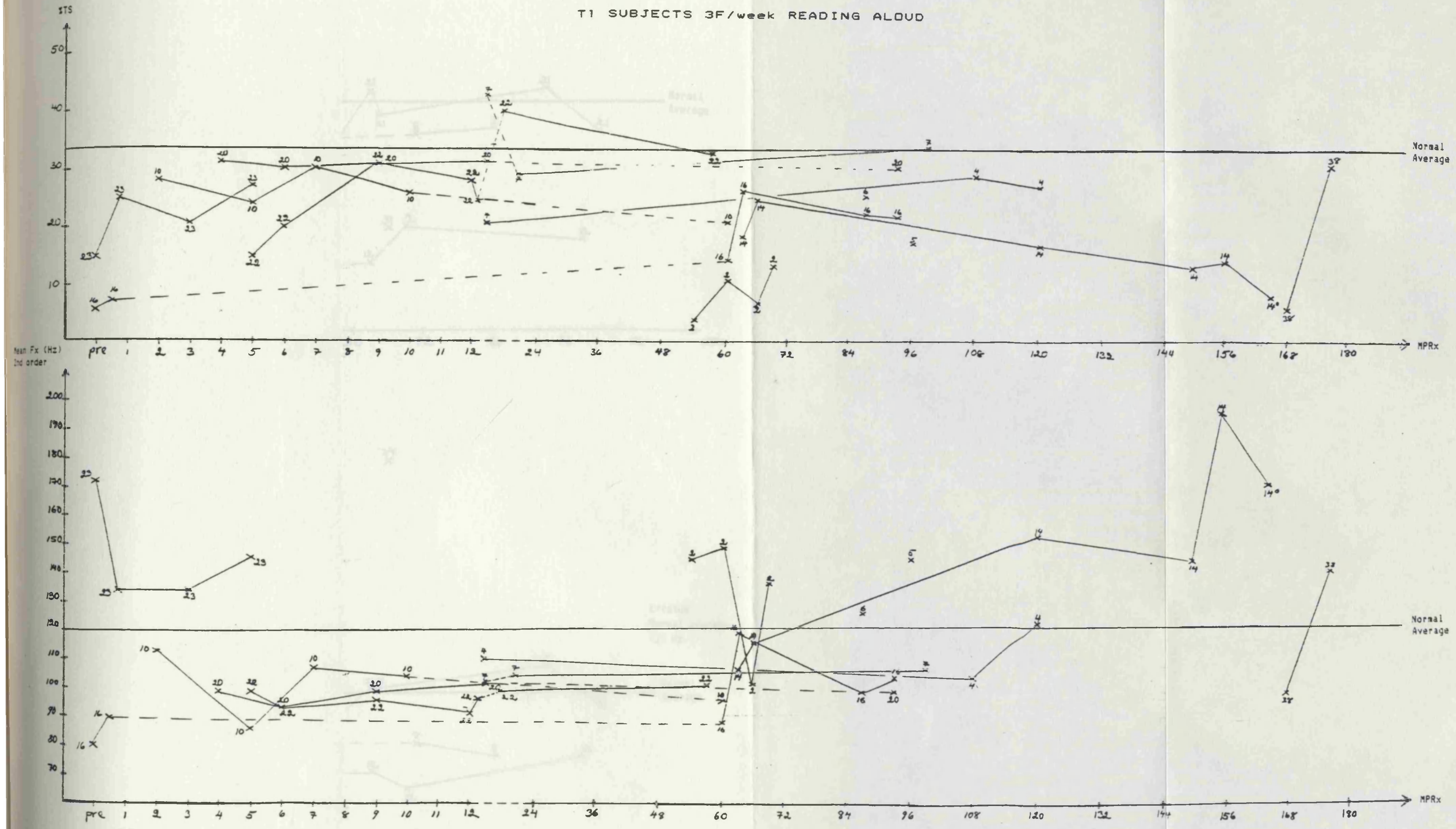
T1 SUBJECTS 3F/week CONVERSATION



Erratum
Normal average
122 Hz

Dashed line indicates change in time scale from monthly
to twelve monthly intervals.

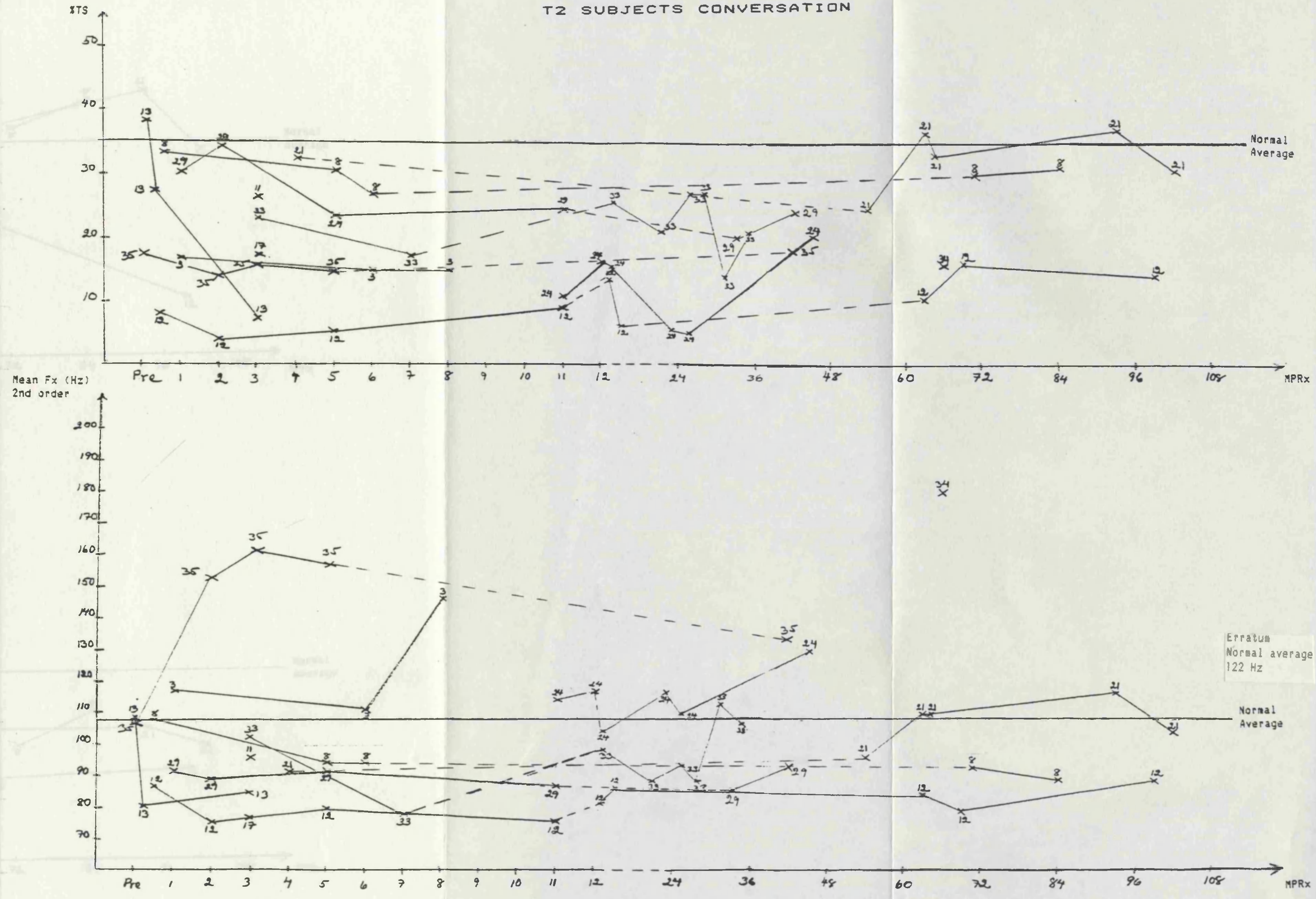
T1 SUBJECTS 3F/week READING ALOUD



Dashed line indicates change in time scale from monthly
to twelve monthly intervals.

APPENDIX 9 A

T2 SUBJECTS CONVERSATION

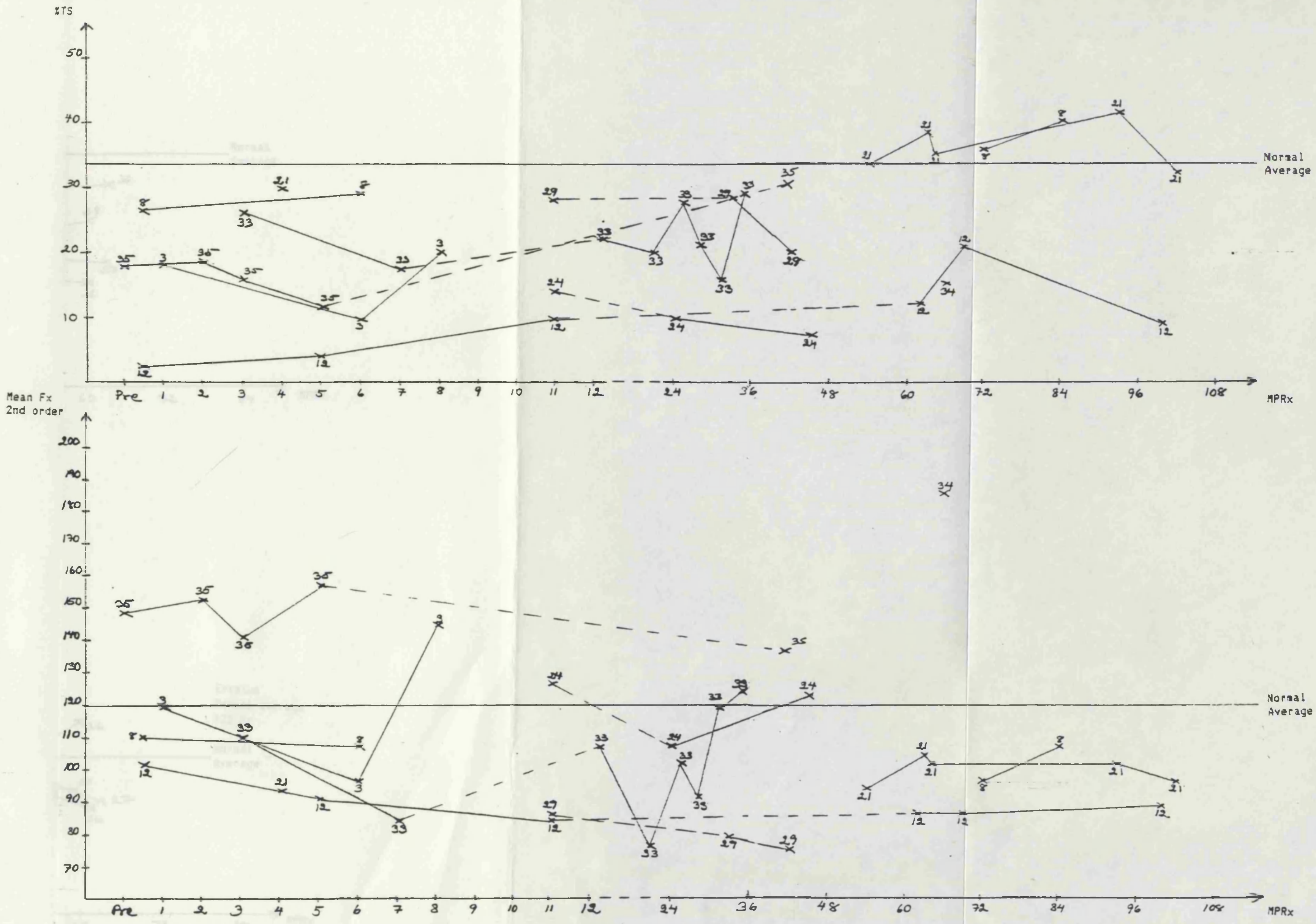


Erratum
Normal average
122 Hz

Dashed line indicates change in time scale from monthly
to twelve monthly intervals.

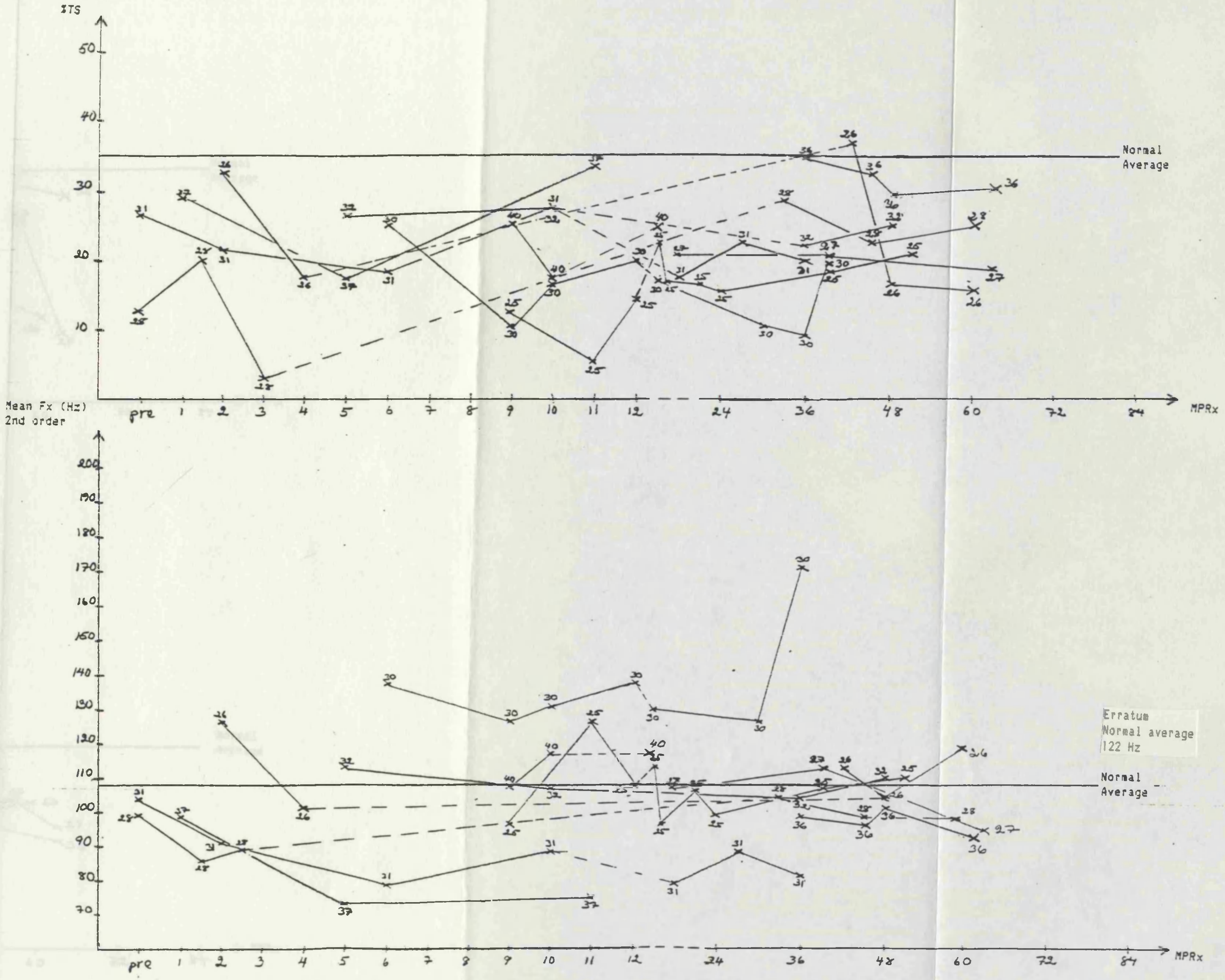
APPENDIX 9 B

T2 SUBJECTS READING ALOUD



Dashed line indicates change in time scale from monthly to twelve monthly intervals.

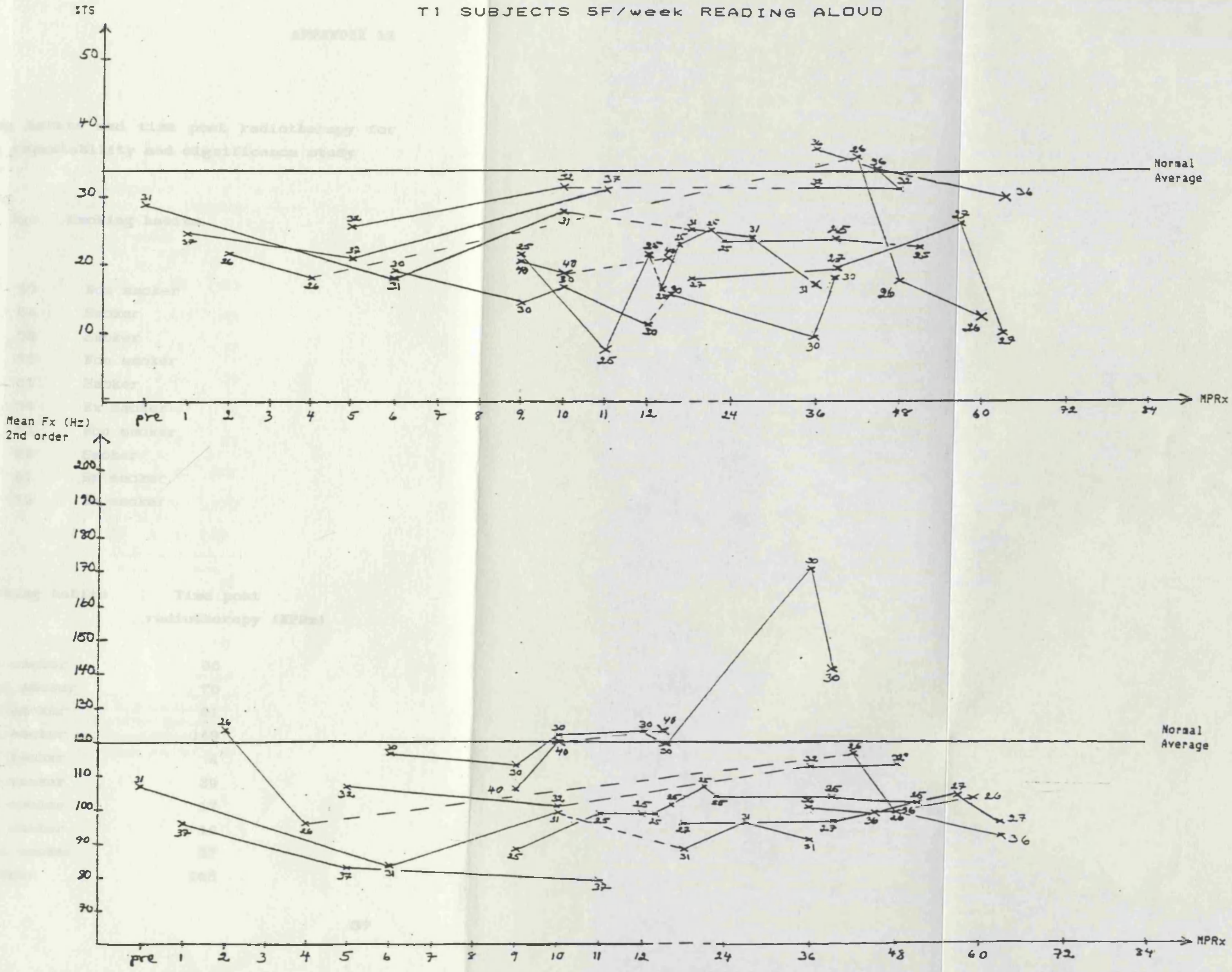
T1 SUBJECTS SF/week CONVERSATION



Erratum
Normal average
122 Hz

Dashed line indicates change in time scale from monthly to twelve monthly intervals.

T1 SUBJECTS SF/week READING ALOUD



Dashed line indicates change in time scale from monthly to twelve monthly intervals.

Identity 'tags', age, smoking habits and time post radiotherapy for two groups of speakers in the repeatability and significance study

Normal speaker	Age	Smoking habits
DGB	55	Non smoker
JSt	58	Smoker
PBr	53	Smoker
RA	75	Non smoker
FC	61	Smoker
H	59	Ex smoker
FW	53	Non smoker
TVM	50	Smoker
DTH	61	Ex smoker
RDD	52	Ex smoker

T1 subject	Age	Smoking habits	Time post radiotherapy (MPRx)
6	67	Ex smoker	86
7	70	Non smoker	76
10	70	Ex smoker	61
14	70	Ex smoker	149
26	60	Ex smoker	4
27	70	Ex smoker	39
31	49	Ex smoker	17
32	63	Ex smoker	10
36	65	Non smoker	37
38	71	Smoker	168

TABLE 1

TPS 1 VS ROM 1

NORMALS, Reading

Subject	1st order Mean (Hz)				2nd order Mean (Hz)			
	TPS 1	ROM 1	diff.	Average	TPS 1	ROM 1	diff.	Average
X	93,5	91,0	2,5	92,3	96,1	93,0	3,1	94,6
JSt	88,5	86,0	2,5	87,3	96,1	93,0	3,1	94,6
PBr	113,2	113,0	0,2	113,1	119,6	119,0	0,6	119,3
RA	153,1	149,0	4,1	151,1	157,3	153,0	4,3	155,2
FC	119,6	116,0	3,6	117,8	122,9	119,0	3,9	121,0
H	129,9	126,0	3,9	128,0	133,5	129,0	4,5	131,3
FW	86,1	83,0	3,1	84,6	88,5	88,0	0,5	88,3
TVM	137,2	133,0	4,2	135,1	141,0	137,0	4,0	139,0
DTH	116,4	113,0	3,4	114,7	116,4	113,0	3,4	114,7
RDD	126,4	123,0	3,4	124,7	126,4	123,0	3,4	124,7
Σx			30,9	1148,6			30,8	1182,7
d			3,1	114,9			3,1	118,3
Σx^2			107,9				112,9	
S_{diff}			1,2				1,4	

Table showing the agreement between two software systems, TPS and ROM, measuring individual speakers' fundamental frequency values on the same reading passage on the same occasion.

TABLE 2

TPS 1 VS ROM 1

NORMALS, Reading

%TS into 2nd order

Subject	TPS 1	ROM 1	diff.	Average
X	23,7	25,5	-1,8	24,6
JSt	33,2	32,0	1,2	32,6
PBr	20,8	22,1	-1,3	21,5
RA	28,6	30,4	-1,8	29,5
FC	43,1	43,2	-0,1	43,2
H	42,0	44,0	-2,0	43,0
FW	27,6	28,0	-0,4	27,8
TVM	45,8	45,6	0,2	45,7
DTH	38,5	38,4	0,1	38,5
ROD	29,1	30,1	-1,0	29,6
Σx			-6,9	336,0
d			-0,7	33,6
Σx^2			14,79	
S_{diff}			1,1	

Table showing the agreement between two software systems, TPS and ROM, measuring individual speakers' %TS values on the same reading passage on the same occasion.

APPENDIX 12

TABLE 3

TPS 1 VS ROM 1

T1 subjects, Reading

Subj.	1st order Mean (Hz)				2nd order Mean (Hz)			
	TPS 1	ROM 1	diff.	Average	TPS 1	ROM 1	diff.	Average
6	122,9	119,0	3,9	121,0	126,4	126,0	0,4	126,2
7	101,5	98,0	3,5	99,8	107,2	104,0	3,2	105,6
10	93,5	91,0	2,5	92,3	96,1	91,0	5,1	93,6
14	144,9	137,0	7,9	141,0	144,9	144,0	0,9	144,5
26	101,5	96,0	5,5	98,8	98,7	96,0	2,7	97,4
27	98,7	98,0	0,7	98,4	96,1	91,0	5,1	93,6
31	88,5	86,0	2,5	87,3	90,9	88,0	2,9	89,5
32	98,7	98,0	0,7	98,4	104,3	101,0	3,3	102,7
36	98,7	96,0	2,7	97,4	101,5	98,0	3,5	99,8
38	107,2	101,0	6,2	104,1	98,7	96,0	2,7	97,4
Σx			36,1	1038,5			27,8	1050,3
d			3,6	103,9			2,8	105,0
Σx^2			179,3				92,96	
S_{diff}			2,3				1,3	

Table showing the agreement between two software systems, TPS and ROM, measuring individual speakers' fundamental frequency values on the same reading passage on the same occasion.

TABLE 4

TPS 1 VS ROM 1

T1 subjects, Reading

%TS into 2nd order

Subject	TPS 1	ROM 1	diff.	Average
6	24,1	25,2	-1,1	24,7
7	33,4	33,2	0,2	33,3
10	18,0	17,8	0,2	17,9
14	9,5	9,5	0,0	9,5
26	20,9	21,3	-0,4	21,3
27	21,2	22,7	-1,5	22,0
31	22,8	23,1	-0,3	23,0
32	35,6	36,7	-1,1	36,2
36	38,4	38,6	-0,2	38,5
38	6,4	7,6	-1,2	7,0
Σx			-5,5	233,4
d			-0,6	23,3
Σx^2			6,7	
S_{diff}			0,6	

Table showing the agreement between two software systems, TPS and ROM, measuring individual speakers' %TS values on the same reading passage on the same occasion.

TABLE 5

TPS 2 VS ROM 2

NORMALS, Reading

Subject	1st order Mean (Hz)				2nd order Mean (Hz)			
	TPS 2	ROM 2	diff.	Average	TPS 2	ROM 2	diff.	Average
X	90,9	88,0	2,9	89,5	93,5	91,0	2,5	92,3
JSt	90,9	88,0	2,9	89,5	96,1	93,0	3,1	94,6
PBr	116,4	113,0	3,4	114,7	122,9	119,0	3,9	121,0
RA	153,1	149,0	4,1	151,1	157,3	153,0	4,3	155,2
FC	119,6	116,0	3,6	117,8	122,9	119,0	3,9	121,0
H	129,9	126,0	3,9	128,0	133,5	129,0	4,5	131,3
FW	86,1	83,0	3,1	84,6	88,5	86,0	2,5	87,3
TVM	137,2	133,0	4,2	135,1	141,0	137,0	4,0	139,0
DTH	113,2	110,0	3,2	111,6	116,4	113,0	3,4	114,7
RDD	126,4	123,0	3,4	124,7	126,4	123,0	3,4	124,7
Σx			34,7	1146,6			35,5	1181,1
d			3,5	114,7			3,5	118,1
Σx^2			122,4				130,4	
S_{diff}			0,5				0,7	

Table showing the agreement between two software systems, TPS and ROM, measuring individual speakers' fundamental frequency values on the same reading passage on the same occasion.

APPENDIX 12

TABLE 6

TPS 2 VS ROM 2

NORMALS, Reading

%TS into 2nd order

Subject	TPS 2	ROM 2	diff.	Average
X	25,79	27,47	-1,68	26,6
JSt	30,39	31,30	-0,91	30,9
PBr	25,00	25,13	-0,13	25,1
RA	28,61	28,64	-0,03	28,6
FC	40,96	41,62	-0,66	41,3
H	42,27	42,50	-0,23	42,4
FW	27,90	26,60	1,3	27,3
TVM	44,21	43,68	0,53	44,0
OTH	38,45	37,59	0,86	38,0
RDD	31,51	31,11	0,40	31,3
Σx			-0,55	335,5
d			-0,06	33,6
Σx^2			7,03	
S_{diff}			0,84	

Table showing the agreement between two software systems, TPS and ROM, measuring individual speakers' %TS values on the same reading passage on the same occasion.

TABLE 7

TPS 2 VS ROM 2

T1 subjects, Reading

Subj.	1st order Mean (Hz)				2nd order Mean (Hz)			
	TPS 2	ROM 2	diff.	Average	TPS 2	ROM 2	diff.	Average
6	126,4	123,0	3,4	124,7	126,4	126,0	0,4	126,2
7	101,5	98,0	3,5	99,8	104,3	104,0	0,3	104,2
10	93,5	91,0	2,5	92,3	96,1	91,0	5,1	93,6
14	141,0	137,0	4,0	139,0	144,9	141,0	3,9	143,0
26	101,5	96,0	5,5	98,8	96,1	96,0	0,1	96,1
27	101,5	98,0	3,5	99,8	96,1	93,0	3,1	94,6
31	93,5	91,0	2,5	92,3	88,5	86,0	2,5	87,3
32	98,7	98,0	0,7	98,4	104,3	101,0	3,3	102,7
36	98,7	96,0	2,7	97,4	101,5	98,0	3,5	99,8
38	107,2	101,0	6,2	104,1	98,7	96,0	2,7	97,4
Σx			33,5	1047,1			28,1	1044,9
d			3,4	104,7			2,8	104,5
Σx^2			119,6				114,3	
S_{diff}			0,9				2,0	

Table showing the agreement between two software systems, TPS and ROM, measuring individual speakers' fundamental frequency values on the same reading passage on the same occasion.

TABLE 8

TPS 2 VS ROM 2

T1 subjects, Reading

%TS into 2nd order

Subject	TPS 2	ROM 2	diff.	Average
6	22,9	23,8	-0,9	23,4
7	31,5	32,5	-1,0	32,0
10	18,6	18,9	-0,3	18,8
14	9,6	9,5	0,1	9,6
26	20,4	21,3	-0,9	20,9
27	24,4	24,4	0,0	24,4
31	19,6	19,6	0,0	19,6
32	31,7	33,5	-1,8	32,6
36	33,9	34,5	-0,6	34,2
38	5,9	5,1	0,8	5,5
Σx			-4,6	220,8
d			-0,5	22,1
Σx^2			6,96	
S_{diff}			0,7	

Table showing the agreement between two software systems, TPS and ROM, measuring individual speakers' %TS values on the same reading passage on the same occasion.

APPENDIX 12

TABLE 9

TPS 1 VS TPS 2

NORMALS, Reading

Subject	1st order Mean (Hz)				2nd order Mean (Hz)			
	TPS 1	TPS 2	diff.	Average	TPS 1	TPS 2	diff.	Average
X	93,5	90,9	2,6	92,2	96,1	93,5	2,6	94,8
JSt	88,5	90,9	-2,4	89,7	96,1	96,1	0,0	96,1
PBr	113,2	116,4	-3,2	114,8	119,6	122,9	-3,3	121,3
RA	153,1	153,1	0,0	153,1	157,3	157,3	0,0	157,3
FC	119,6	119,6	0,0	119,6	122,9	122,9	0,0	122,9
H	129,9	129,9	0,0	129,9	133,5	133,5	0,0	133,5
FW	86,1	86,1	0,0	86,1	88,5	88,5	0,0	88,5
TVM	137,2	137,2	0,0	137,2	141,0	141,0	0,0	141,0
DTH	116,4	113,2	3,2	114,8	116,4	116,4	0,0	116,4
RDD	126,4	126,4	0,0	126,4	126,4	126,4	0,0	126,4
Σx			0,2	1163,8			-0,7	1198,2
d			0,02	116,4			-0,07	119,8
Σx^2			33,00				17,65	
S_{diff}			1,9				1,4	

Table showing the repeatability of measurements using the TPS system on two different occasions on the same reading passage.

APPENDIX 12

TABLE 10 a

TPS 1 VS TPS 2

NORMALS, Reading

Subject	1st order Mode (Hz)				2nd order Mode (Hz)			
	TPS 1	TPS 2	diff.	Average	TPS 1	TPS 2	diff.	Average
X	83,8	83,8	0,0	83,8	83,8	83,8	0,0	83,8
JSt	88,5	88,5	0,0	88,5	88,5	88,5	0,0	88,5
PBr	101,5	101,5	0,0	101,5	122,9	119,6	3,3	121,3
RA	161,7	161,7	0,0	161,7	161,7	161,7	0,0	161,7
FC	119,6	119,6	0,0	119,6	133,5	119,6	13,9	126,6
H	133,5	133,5	0,0	133,5	133,5	133,5	0,0	133,5
FW	77,2	77,2	0,0	77,2	77,2	77,2	0,0	77,2
TVM	144,9	144,9	0,0	144,9	144,9	144,9	0,0	144,9
DTH	116,4	116,4	0,0	116,4	116,4	116,4	0,0	116,4
RDD	126,4	129,9	-3,5	128,2	126,4	126,4	0,0	126,4
Σx			-3,5	1155,3			17,2	1180,2
d			-0,4	115,5			1,7	118,0
Σx^2			12,3				204,1	
S_{diff}			1,1				4,4	

Table showing the repeatability of Fx Mode measurements using the TPS system on two different occasions on the same reading passage.

TABLE 10 b

TPS 1 VS TPS 2

NORMALS, Reading

	2nd order 90 % range minimum (Hz)				2nd order90 % range maximum (Hz)			
	TPS 1	TPS 2	diff.	Average	TPS 1	TPS 2	diff.	Average
X	73,0	73,0	0,0	73,0	144,8	141,0	3,8	142,9
JSt	75,1	73,0	2,1	74,1	133,5	133,5	0,0	133,5
PBr	86,1	88,5	-2,4	87,3	195,8	201,3	-5,5	198,6
RA	116,4	113,2	3,2	114,8	230,9	224,7	6,2	227,8
FC	104,3	104,3	0,0	104,3	157,3	153,1	4,2	155,2
H	107,2	107,2	0,0	107,2	166,2	170,8	4,6	168,5
FW	73,0	71,1	1,9	72,1	129,8	129,8	0,0	129,8
TVM	113,2	113,2	0,0	113,2	170,8	170,8	0,0	170,8
DTH	93,5	93,5	0,0	93,5	157,3	148,8	8,5	153,1
RDD	98,6	98,6	0,0	98,6	170,8	166,2	4,6	168,5
Σx			4,8	938,0			26,4	1648,6
d			0,48	93,8			2,64	164,9
Σx^2			24,02				215,34	
S_{diff}			1,55				4,02	

Table showing the repeatability of measurements using the TPS system on two different occasions on the same reading passage

APPENDIX 12

TABLE 11

TPS 1 VS TPS 2

NORMALS, Reading

% TS into 2nd order

	TPS 1	TPS 2	diff.	Average
X	23,7	25,79	-2,09	24,8
JSt	33,2	30,39	2,81	31,8
PBr	20,8	25,0	-4,2	22,9
RA	28,6	28,6	0,0	28,6
FC	43,1	40,96	2,14	42,0
H	42,0	42,27	-0,27	42,1
FW	27,6	27,9	-0,3	27,8
TVM	45,8	44,21	1,59	45,0
DTH	38,5	38,45	-0,05	38,5
RDD	29,1	31,51	-2,41	30,3
Σx			-1,69	334,3
d			-0,17	33,4
Σx^2			43,97	
S_{diff}			2,2	

Table showing the repeatability of measurements using the TPS system on two different occasions on the same reading passage.

APPENDIX 12

TABLE 12

TPS 1 VS TPS 2

TI subjects, Reading

Subject	1st order Mean (Hz)				2nd order Mean (Hz)			
	TPS 1	TPS 2	diff.	Average	TPS 1	TPS 2	diff.	Average
6	122,9	126,4	-3,5	124,7	126,4	126,4	0,0	126,4
7	101,5	101,5	0,0	101,5	107,2	104,3	2,9	105,8
10	93,5	93,5	0,0	93,5	96,1	96,1	0,0	96,1
14	144,9	141,0	3,9	143,0	144,9	144,9	0,0	144,9
26	101,5	101,5	0,0	101,5	98,7	96,1	2,6	97,4
27	98,7	101,5	-2,8	100,1	96,1	96,1	0,0	96,1
31	88,5	93,5	-5,0	91,0	90,9	88,5	2,4	89,7
32	98,7	98,7	0,0	98,7	104,3	101,5	2,8	102,9
36	98,7	98,7	0,0	98,7	101,5	101,5	0,0	101,5
38	107,2	107,2	0,0	107,2	98,7	98,7	0,0	98,7
Σx			-7,4	1059,9			10,70	1059,50
d			-0,74	105,9			1,07	106,0
Σx^2			60,30				28,77	
S_{diff}			2,5				1,39	

Table showing the repeatability of measurements using the TPS system on two different occasions on the same reading passage.

APPENDIX 12

TABLE 13 a

TPS 1 VS TPS 2

T1 subjects, Reading

Subject	1st order Mode (Hz)				2nd order Mode (Hz)			
	TPS 1	TPS 2	diff.	Average	TPS 1	TPS 2	diff.	Average
6	129,9	129,9	0,0	129,9	129,9	133,5	-3,6	131,7
7	110,2	110,2	0,0	110,2	110,2	110,2	0,0	110,2
10	90,9	88,5	2,4	89,7	90,9	90,9	0,0	90,9
14	157,3	157,3	0,0	157,3	153,1	157,3	-4,2	155,2
26	90,9	90,9	0,0	90,9	110,2	107,2	3,0	108,7
27	98,7	98,7	0,0	98,7	98,7	98,7	0,0	98,7
31	86,1	83,8	2,3	85,0	86,1	83,8	2,3	85,0
32	96,1	93,5	2,6	94,8	96,1	96,1	0,0	96,1
36	93,5	93,5	0,0	93,5	93,5	93,5	0,0	93,5
38	175,6	88,5	87,1	132,1	81,5	88,5	-7,0	85,0
Σx			94,4	1082			-9,5	1055,0
d			9,4	108,2			-0,95	105,5
Σx^2			7604,2				93,9	
S_{diff}			27,3				3,1	

Table showing the repeatability of Fx Mode measurements using the TPS system on two different occasions on the same reading passage.

TABLE 13 b

TPS 1 VS TPS 2

T1 subjects, Reading

Subject	2nd order 90 % range minimum (Hz)				2nd order90 % range maximum (Hz)			
	TPS 1	TPS 2	diff.	Average	TPS 1	TPS 2	diff.	Average
6	113,2	113,2	0,0	113,2	148,8	148,8	0,0	148,8
7	86,1	86,1	0,0	86,1	144,8	137,2	7,6	141,0
10	77,1	77,1	0,0	77,1	119,6	119,6	0,0	119,6
14	83,8	90,9	-7,1	87,4	185,3	195,8	-10,5	190,6
26	77,1	75,1	2,0	76,1	122,9	122,9	0,0	122,9
27	81,5	81,5	0,0	81,5	116,4	119,6	-3,2	118,0
31	75,1	73,0	2,1	74,1	116,4	116,4	0,0	116,4
32	81,5	79,3	2,2	80,4	133,5	133,5	0,0	133,5
36	81,5	81,5	0,0	81,5	141,0	144,8	-3,8	142,9
38	73,0	69,1	3,9	71,1	175,6	175,6	0,0	175,6
Σx			3,1	828,4			-9,9	1409,3
d			0,31	82,8			-0,99	140,9
Σx^2			78,87				192,69	
S_{diff}			2,9				4,6	

Table showing the repeatability of measurements using the TPS system on two different occasions on the same reading passage.

APPENDIX 12

TABLE 14

TPS 1 VS TPS 2

T1 subjects, Reading

% TS into 2nd order

Subject	TPS 1	TPS 2	diff.	Average
6	24,09	22,9	1,2	23,5
7	33,4	31,5	1,9	32,5
10	18,0	18,6	-0,6	18,3
14	9,5	9,6	-0,1	9,6
26	20,9	20,4	0,5	20,7
27	21,2	24,4	-3,2	22,8
31	22,8	19,6	3,2	21,2
32	35,6	31,7	3,9	33,7
36	38,4	33,9	4,5	36,2
38	6,4	5,9	0,5	6,2
Σx			11,8	224,7
d			1,18	22,5
Σx^2			61,86	
S_{diff}			2,3	

Table showing the repeatability of measurements using the TPS system on two different occasions on the same reading passage.

APPENDIX 12

TABLE 15

ROM 1 VS ROM 2

NORMALS, Reading

Subject	1st order Mean (Hz)				2nd order Mean (Hz)			
	ROM 1	ROM 2	diff.	Average	ROM 1	ROM 2	diff.	Average
X	91,0	88,0	3,0	89,5	93,0	91,0	2,0	92,0
JSt	86,0	88,0	-2,0	87,0	93,0	93,0	0,0	93,0
PBr	113,0	113,0	0,0	113,0	119,0	119,0	0,0	119,0
RA	149,0	149,0	0,0	149,0	153,0	153,0	0,0	153,0
FC	116,0	116,0	0,0	116,0	119,0	119,0	0,0	119,0
H	126,0	126,0	0,0	126,0	129,0	129,0	0,0	129,0
FW	83,0	83,0	0,0	83,0	88,0	86,0	2,0	87,0
TVM	133,0	133,0	0,0	133,0	137,0	137,0	0,0	137,0
DTH	113,0	110,0	3,0	111,5	113,0	113,0	0,0	113,0
RDD	123,0	123,0	0,0	123,0	123,0	123,0	0,0	123,0
Σx			4,0	1131,0			4,0	1165,0
d			0,4	113,1			0,4	116,5
Σx^2			22,0				8,0	
S_{diff}			1,5				0,8	

Table showing the repeatability of measurements using the ROM system on two occasions, ROM 1 and ROM 2, on the same reading passage.

APPENDIX 12

TABLE 16

ROM 1 VS ROM 2

NORMALS, Reading

Subject	1st order Mode (Hz)				2nd order Mode (Hz)			
	ROM 1	ROM 2	diff.	Average	ROM 1	ROM 2	diff.	Average
X	81,0	81,0	0,0	81,0	86,0	81,0	5,0	83,5
JSt	81,0	86,0	-5,0	83,5	86,0	86,0	0,0	86,0
PBr	96,0	98,0	-2,0	97,0	113,0	119,0	-6,0	116,0
RA	161,0	157,0	4,0	159,0	161,0	157,0	4,0	159,0
FC	113,0	116,0	-3,0	114,5	129,0	129,0	0,0	129,0
H	129,0	129,0	0,0	129,0	129,0	129,0	0,0	129,0
FW	75,0	75,0	0,0	75,0	75,0	75,0	0,0	75,0
TVM	141,0	141,0	0,0	141,0	141,0	129,0	12,0	135,0
DTH	113,0	113,0	0,0	113,0	113,0	113,0	0,0	113,0
RDD	123,0	123,0	0,0	123,0	119,0	123,0	-4,0	121,0
Σx			-6,0	1116,0			11,0	1146,5
d			-0,6	111,6			1,1	114,7
Σx^2			54,0				237,0	
S_{diff}			2,4				5,0	

Table showing the repeatability of measurements using the ROM system on two occasions, ROM 1 and ROM 2, on the same reading passage.

APPENDIX 12

TABLE 17

ROM 1 VS ROM 2

NORMALS, Reading

%TS into 2nd order

Subject	ROM 1	ROM 2	diff.	Average
X	25,5	27,5	-1,9	26,5
JSt	32,0	31,3	0,7	31,7
PBr	22,1	25,1	-3,0	23,6
RA	30,4	28,6	1,8	29,5
FC	43,2	41,6	1,6	42,4
H	44,0	42,5	1,5	43,3
FW	28,8	26,6	1,4	27,3
TVM	45,6	43,7	1,9	44,6
DTH	38,4	37,6	0,8	38,0
ROD	30,1	31,1	1,0	30,6
Σx			5,7	337,5
d			0,6	33,8
Σx^2			28,7	
S_{diff}			1,7	

Table showing the repeatability of measurements using the ROM system on two occasions, ROM 1 and ROM 2, on the same reading passage.

APPENDIX 12

TABLE 18

ROM 1 VS ROM 2

Tl subjects, Reading

Subject	1st order Mean (Hz)				2nd order Mean (Hz)			
	ROM 1	ROM 2	diff.	Average	ROM 1	ROM 2	diff.	Average
6	119,0	123,0	-4,0	121,0	126,0	126,0	0,0	126,0
7	98,0	98,0	0,0	98,0	104,0	104,0	0,0	104,0
10	91,0	91,0	0,0	91,0	91,0	91,0	0,0	91,0
14	137,0	137,0	0,0	137,0	144,0	141,0	3,0	142,5
26	96,0	98,0	-2,0	97,0	96,0	96,0	0,0	96,0
27	98,0	98,0	0,0	98,0	93,0	93,0	0,0	93,0
31	86,0	91,0	-5,0	88,5	88,0	86,0	2,0	87,0
32	98,0	96,0	2,0	87,0	101,0	98,0	3,0	99,5
36	96,0	96,0	0,0	96,0	98,0	98,0	0,0	98,0
38	101,0	104,0	-3,0	102,5	96,0	93,0	3,0	94,5
Σx			-12,0	1206,0			11,0	1034,1
d			-1,2	120,6			1,1	103,4
Σx^2			58,0				31,0	
S_{diff}			2,2				1,5	

Table showing the repeatability of measurements using the ROM system on two occasions, ROM 1 and ROM 2, on the same reading passage.

TABLE 19

ROM 1 VS ROM 2

T1 subjects, Reading

Subject	1st order Mode (Hz)				2nd order Mode (Hz)			
	ROM 1	ROM 2	diff.	Average	ROM 1	ROM 2	diff.	Average
6	126,0	129,0	-3,0	127,5	129,0	129,0	0,0	129,0
7	107,0	104,0	3,0	105,5	107,0	107,0	0,0	107,0
10	88,0	86,0	2,0	87,0	88,0	88,0	0,0	88,0
14	149,0	149,0	0,0	149,0	149,0	149,0	0,0	149,0
26	88,0	88,0	0,0	88,0	104,0	104,0	0,0	104,0
27	93,0	96,0	-3,0	94,5	96,0	96,0	0,0	96,0
31	83,0	83,0	0,0	83,0	83,0	83,0	0,0	83,0
32	93,0	93,0	0,0	93,0	93,0	93,0	0,0	93,0
36	91,0	88,0	3,0	89,5	91,0	88,0	3,0	89,5
38	170,0	86,0	84,0	128,0	86,0	86,0	0,0	86,0
Σx			89,0	1045,0			3,0	1024,5
d			8,9	104,5			0,3	102,5
Σx^2			7096,0				9,0	
S_{diff}			26,5				1,0	

Table showing the repeatability of measurements using the ROM system on two occasions, ROM 1 and ROM 2, on the same reading passage.

APPENDIX 12

TABLE 20

ROM 1 VS ROM 2

T1 subjects, Reading

%TS into 2nd order

Subject	ROM 1	ROM 2	diff.	Average
6	25,2	23,8	1,4	24,5
7	33,2	32,5	0,7	32,9
10	17,8	18,6	-0,8	18,2
14	9,5	10,3	-0,8	9,9
26	21,3	20,4	0,9	20,9
27	22,7	24,4	-1,7	23,6
31	23,1	19,6	3,5	21,4
32	36,7	33,5	3,2	35,1
36	38,6	34,5	4,1	36,6
38	7,6	5,1	2,5	6,4
Σx			13,0	229,5
d			1,3	23,0
Σx^2			52,98	
S_{diff}			2,0	

Table showing the repeatability of measurements using the ROM system on two occasions, ROM 1 and ROM 2, on the same reading passage.

TABLE 21

TPS1 vs PCLx1

NORMALS, Reading

	1st order Mean (Hz)				2nd order Mean (Hz)			
Subject	TPS1	PCLx1	diff.	Average	TPS1	PCLx1	diff	Average
DGB	141,0	129,9	11,1	135,5	141,0	129,9	11,1	135,5
JSt	88,5	88,5	0	88,5	96,1	91,0	5,1	93,6
PBr	113,2	119,6	-6,4	116,4	119,6	123,0	-3,4	121,3
RA	153,1	149,0	4,1	151,1	157,3	153,1	4,2	155,2
FC	119,6	119,6	0	119,6	122,9	119,6	3,3	121,3
H	129,9	123,0	6,9	126,5	133,5	126,4	7,1	130,0
FW	86,1	104,3	-18,2	95,2	88,5	91,0	-2,5	89,8
TVM	137,2	129,9	7,3	133,5	141,0	133,5	7,5	137,3
DTH	116,4	110,2	6,2	113,3	116,4	113,3	3,1	114,9
RDD	126,4	123,0	3,4	124,7	126,4	129,9	-3,5	128,2
Σx			14,4	1204,2			32,0	1226,7
d			1,44	120,4			3,2	122,7
Σx^2			663,12				324,08	
SD_{diff}			8,5				5,0	

Table showing agreement between individual speakers' fundamental frequency measurements using two different software systems, TPS1 vs PCLx1, on the same reading passage on the same occasion.

TABLE 22

TPS1 vs PCLx1

NORMALS, Reading

	1st Mode (Hz)				2nd Mode (Hz)			
Subject	TPS1	PCLx1	diff,	Average	TPS1	PCLx1	diff	Average
DGB	129,9	126,4	3,5	128,2	129,9	126,4	3,5	128,2
JSt	88,5	79,3	9,2	83,9	88,5	79,3	9,2	83,9
PBr	101,5	113,3	-11,8	107,4	122,9	113,3	9,6	118,1
RA	161,7	157,3	4,4	159,5	161,7	157,3	4,4	159,5
FC	119,6	113,3	6,3	116,5	133,5	126,4	7,1	130,0
H	133,5	126,4	7,1	130,0	133,5	126,4	7,1	130,0
FW	77,2	73,1	4,1	75,2	77,2	88,5	-11,3	82,9
TVM	144,9	133,5	11,4	139,2	144,9	133,5	11,4	139,2
DTH	116,4	110,2	6,2	113,3	116,4	110,2	6,2	113,3
RDD	126,4	119,6	6,8	123,0	126,4	116,4	10,0	121,4
Σx			47,2	1176,05			572	1206,5
d			4,7	117,6			5,7	120,7
Σx^2			577,04				705,32	
S_{diff}			3,8				6,5	

Example of the agreement between two different software systems TPS and PCLx measuring individual speakers' 1st and 2nd order Modes on the same reading passage on the same occasion,

TABLE 23

TPS1 vs PCLx1

NORMALS, Reading

	2nd order 90 % range Minimum (Hz)				2nd order 90 % Maximum (Hz)			
	TPS1	PCLx1	diff.	Average	TPS1	PCLx1	diff	Average
Subject								
DGB	116,4	113,3	3,1	114,9	180,3	170,8	9,5	175,6
JSt	75,1	73,1	2,0	74,1	133,5	129,9	3,6	131,7
PBr	86,1	93,5	-7,4	89,8	195,8	195,9	0,1	195,9
RA	116,4	116,4	0	116,4	230,9	230,9	0	230,9
FC	104,3	101,5	2,8	102,9	157,3	157,3	0	157,3
H	107,2	107,2	0	107,2	170,8	166,2	4,6	168,5
FW	73,0	73,1	0,1	73,1	129,8	126,4	3,4	128,1
TVM	113,2	113,3	0,1	113,3	170,8	170,8	0	170,8
DTH	93,5	91,0	2,5	92,3	157,3	166,2	-8,9	161,8
RDD	98,6	110,2	-11,6	104,4	170,8	170,8	0	170,8
Σx			-8,4	988,2			12,3	1691,3
d			-0,84	98,8			1,23	169,1
Σx^2			217,04				215,15	
SD_{diff}			4,8				4,7	

Table showing agreement between individual speakers' fundamental frequency measurements using two different software systems, TPS1 vs PCLx1, on the same reading passage on the same occasion.

TABLE 24

TPS1 vs PCLx1

NORMALS, Reading

% TS into 2nd order

Subject	TPS1	PCLx1	diff.	Average
DGB	36.7	48.6	-11.9	42.7
JSt	33.2	36.6	-3.4	34.9
PBr	20.8	32.2	-11.4	26.5
RA	29.6	48.5	-18.9	39.1
FC	43.1	42.5	0.6	42.8
H	42.3	49.8	-7.5	46.1
FW	27.6	19.9	7.7	23.8
TVM	45.8	44.9	0.9	45.4
DTH	38.5	38.0	0.5	38.3
RDD	29.1	44.8	-15.7	37.0
Σx			-59.1	376.6
d			-5.9	37.7
Σx^2			1003.79	
SD_{diff}			8.5	

Table showing agreement between individual speakers' fundamental frequency measurements using two different software systems, TPS1 vs PCLx1, on the same reading passage on the same occasion.

TABLE 25

TPS1 vs PCLx

T1 subjects, Reading

	1st order Mean (Hz)				2nd order Mean (Hz)			
Subject	TPS1	PCLx1	diff.	Average	TPS1	PCLx1	diff	Average
6	122,9	119,6	3,3	121,3	126,4	123,0	3,4	124,7
7	101,5	98,8	2,7	100,2	107,2	101,5	5,7	104,4
10	-	-	-	-	-	-	-	-
14	144,9	149,0	-4,1	147,0	144,9	141,0	3,9	143,0
26	101,5	101,5	0,0	101,5	98,7	93,5	5,2	96,1
27	98,7	88,5	10,2	93,6	96,1	91,0	5,1	93,6
31	88,5	83,8	4,7	86,2	90,9	83,8	7,1	87,4
32	98,7	93,5	5,2	96,1	104,3	96,1	8,2	100,2
36	98,7	96,1	2,6	97,4	101,5	96,1	5,4	98,8
38	107,2	123,0	-15,8	115,1	98,7	104,3	-5,6	101,5
Σx			8,8	958,2			38,4	949,5
d			0,98	106,5			4,3	105,5
Σx^2			444,6				290,5	
S_{diff}			7,4				4,0	

Table showing the agreement between two different software systems TPS and PCLx measuring individual speakers' 1st and 2nd order Means on the same reading passage on different occasions.

APPENDIX 12

TABLE 26

TPS1 vs PCLx

T1 subjects, READING

	1st order Mode (Hz)				2nd order Mode (Hz)			
	TPS1	PCLx1	diff.	Average	TPS1	PCLx1	diff	Average
Subject								
6	129,9	116,4	13,5	123,2	129,9	119,6	10,3	124,8
7	110,2	101,5	8,7	105,9	110,2	101,5	8,7	105,9
10	-	-	-	-	-	-	-	-
14	157,3	153,1	4,2	155,2	153,1	153,1	0	153,1
26	90,9	86,1	4,8	88,5	110,2	86,1	24,1	98,2
27	98,7	93,5	5,2	96,1	98,7	93,5	5,2	96,1
31	86,1	79,3	6,8	82,7	86,1	81,5	4,6	83,8
32	96,1	88,5	7,6	92,3	96,1	91,0	6,1	94,1
36	93,5	88,5	5,0	91,0	93,5	88,5	5,0	91,0
38	175,6	83,8	91,8	129,7	81,5	83,8	-2,3	82,7
Σx			147,6	964,5			61,7	929,7
d			16,4	107,2			6,9	103,3
Σx^2			8881,9				879,1	
SD_{diff}			28,4				7,6	

Example of the agreement between two different software systems TPS and PCLx measuring individual speakers' 1st and 2nd order Modes on the same reading passage on the same occasion.

TABLE 27

TPS1 vs PCLx

T1 subjects, Reading

	2nd order 90% range Minimum (Hz)				2nd order 90 % range Maximum (Hz)			
Subject	TPS1	PCLx1	diff.	Average	TPS1	PCLx1	diff	Average
6	113,2	113,3	-0,1	113,3	148,8	149,0	-0,2	148,9
7	86,1	83,8	2,3	85,0	144,8	137,2	7,6	141,0
10	-	-	-	-	-	-	-	-
14	83,8	119,6	-35,8	101,7	185,3	170,8	14,5	178,1
26	77,1	81,5	-4,4	79,3	122,9	123,0	-0,1	123,0
27	81,5	83,8	-2,3	82,7	116,4	113,3	3,1	114,9
31	75,1	75,1	0,0	75,1	116,4	110,2	6,2	113,3
32	81,5	79,3	2,2	80,4	133,5	129,9	3,6	131,7
36	81,5	81,5	0,0	81,5	141,0	137,2	3,8	139,1
38	73,0	75,1	-2,1	74,1	175,6	175,6	0,0	175,6
Σx			-40,2	772,9			38,5	1265,5
d			-4,0	85,9			3,9	126,6
Σx^2			1320,8				343,5	
S_{diff}			11,9				4,7	

Table showing the agreement between two different software systems TPS and PCLx measuring individual speakers' 2nd order 90 % range Minima and Maxima on the same reading passage on different occasions.

APPENDIX 12

TABLE 28

TPS1 vs PCLx

T1 subjects, READING

%TS into 2nd order

Subject	TPS1	PCLx1	diff.	Average
6	24.1	27.2	-3.1	25.7
7	33.4	34.3	-0.9	33.9
10	-	-	-	-
14	9.5	13.9	-4.4	11.7
26	20.9	21.9	-1.0	21.4
27	21.2	24.4	-3.2	22.8
31	22.8	35.5	-12.7	29.2 *
32	35.6	36.2	-0.6	35.9
36	38.4	34.4	4.0	36.4
38	6.4	6.5	-0.1	6.5
Σx			-22.0	223.3
d			-2.4	24.8
Σx^2			218.6	
S_{diff}			4.5	

Table showing the agreement between two different software systems TPS and PCLx measuring individual speakers' % TS values on the same reading passage on different occasions.

* (only two paragraphs analysed on the first occasion, 3 on the second; i.e. sample larger)

TABLE 29

TPS1 vs PCLx1

NORMALS, Speech

	1st order Mean (Hz)				2nd order Mean (Hz)			
Subject	TPS1	PCLx1	diff.	Average	TPS1	PCLx1	diff	Average
DGB	141,0	137,2	3,8	139,1	141,0	133,5	7,5	137,3
JSt	86,1	88,5	-2,4	87,3	90,9	86,1	4,8	88,5
PBr	116,4	116,4	0,0	116,4	116,4	116,4	0,0	116,4
RA	144,9	137,2	7,7	141,1	148,9	141,0	7,9	145,0
FC	116,4	110,2	6,2	113,3	119,6	116,4	3,2	118,0
H	116,4	113,3	3,1	114,9	122,9	116,4	6,5	119,7
FW	83,8	119,6	-35,8	101,7	86,1	93,5	-7,4	89,8
TVM	119,6	113,3	6,3	116,5	122,9	116,4	6,5	119,7
DTH	141,0	133,5	7,5	137,3	141,0	137,2	3,8	139,1
RDD	126,4	126,4	0,0	126,4	129,9	129,9	0,0	129,9
Σ x			-3,6	1193,9			25,3	1203,3
x			-0,4	119,4			2,5	120,3
Σx^2			1505,1				305,6	
S_{diff}			12,9				5,2	

Table showing the agreement between two different software systems, TPS and PCLx, measuring individual speakers' 1st and 2nd order Means on approximately the same speech sample on different occasions.

APPENDIX 12

TABLE 30

TPS1 vs PCLx1

NORMALS, Speech

	1st Mode (Hz)				2nd Mode (Hz)			
Subject	TPS1	PCLx1	diff.	Average	TPS1	PCLx1	diff	Average
DGB	129,9	123,0	6,9	126,5	129,9	123,0	6,9	126,5
JSt	98,7	96,1	2,6	94,8	98,7	101,5	-2,8	100,1
PBr	96,1	98,8	-2,7	97,5	104,3	98,8	5,5	101,6
RA	137,2	144,9	-7,7	141,1	153,1	153,1	0	153,1
FC	107,2	93,5	13,7	100,4	107,2	101,5	5,7	104,4
H	113,2	104,3	8,9	108,8	113,2	104,3	8,9	108,8
FW	75,1	88,5	-13,4	81,8	77,2	88,5	-11,3	82,9
TVM	116,4	113,3	3,1	114,9	116,4	116,4	0	116,4
DTH	129,9	123,0	6,9	126,5	129,9	123,0	6,9	126,5
RDD	133,5	133,5	0	133,5	133,5	133,5	0	133,5
Σ x			18,3	1125,8			19,8	1153,8
d			1,8	112,6			2,0	115,4
Σx^2			624,6				372,9	
S_{diff}			8,1				6,1	

Example of the agreement between two different software systems TPS and PCLx measuring individual speakers' Speaking 1st and 2nd order Modes on the same occasion.

APPENDIX 12

TABLE 31

TPS1 vs PCLx1

NORMALS, Speech

	2nd order 90 % range Minimum				2nd order 90 % range Maximum			
Subject	TPS1	PCLx1	diff.	Average	TPS1	PCLx1	diff	Average
DGB	113,2	113,3	-0,1	113,3	195,8	195,9	-0,1	195,9
JSt	65,5	67,3	-1,8	66,4	119,6	116,4	3,2	118,0
PBr	88,5	91,0	-2,5	89,8	206,9	195,9	11,0	201,4
RA	107,2	107,2	0,0	107,2	224,7	212,7	12,0	218,7
FC	93,5	91,0	2,5	92,3	201,3	195,9	5,4	198,6
H	98,6	96,1	2,5	97,4	166,2	166,2	0,0	166,2
FW	69,1	71,1	-2,0	70,1	175,6	175,6	0,0	175,6
TVM	104,3	101,5	2,8	102,9	157,3	153,1	4,2	155,2
DTH	96,1	93,5	2,6	94,8	287,5	287,5	0,0	287,5
RDD	104,3	110,2	-5,9	107,3	175,6	180,4	-4,8	178,0
Σx			-1,9	941,4			30,9	1895,1
x			-0,2	94,1			3,01	189,5
Σx^2			75,4				345,1	
S_{diff}			2,9				5,3	

Table showing the agreement between two different software systems, TPS and PCLx, measuring individual speakers' 90 % range Minima and Maxima on approximately the same speech sample on different occasions.

TABLE 32

TPS1 vs PCLx1

NORMALS, Speech

	% TS into 2nd order				Total sample 1st order	
Subject	TPS1	PCLx1	diff.	Average	TPS1	PCLx1
DGB	39,8	46,4	-6,6	43,1	7509	7909
JSt	26,7	27,8	-1,1	27,3	6217	6793
PBr	24,2	36,0	-11,8	30,1	6874	7133
RA	36,9	46,4	-9,5	41,7	6307	8164
FC	44,2	39,1	5,1	41,7	6155	6513
H	38,6	39,0	-0,4	38,8	6458	7668
FW	22,8	11,7	11,1	17,3	5816	7835
TVM	43,3	39,3	4,0	41,3	6154	6974
DTH	36,8	35,6	1,2	36,2	6267	6951
RDD	33,9	48,3	-14,4	41,1	6076	7413
Σx			-22,4	315,5	63833	73353
x			-2,2	31,6	6383,3	7335,3
Σx^2			648,4			
S_{diff}			8,2			

Table showing the agreement between two different software systems, TPS and PCLx, measuring individual speakers' %TS carried into 2nd order on approximately the same speech sample on different occasions.

Also showing the Total sample of Tx intervals analysed by each program.

TABLE 33

TPS1 vs PCLx

T1 subjects, Speech

	1st order Mean (Hz)				2nd order Mean (Hz)			
Subject	TPS1	PCLx1	diff.	Average	TPS1	PCLx1	diff	Average
6	119,6	113,3	6,3	116,5	119,6	116,4	3,3	118,0
7	107,2	107,2	0,0	107,2	113,2	107,2	6,0	110,2
10	90,9	91,0	-0,1	91,0	98,7	91,0	7,7	94,9*
14	141,0	141,0	0,0	141,0	148,9	141,0	7,9	145,0
26	107,2	104,3	2,9	105,8	104,3	96,1	8,2	100,2
27	107,2	98,8	8,4	103,0	110,2	101,5	8,7	105,9
31	86,1	88,5	-2,4	87,3	79,3	79,3	0,0	79,3
32	98,7	104,3	-5,6	101,5	104,3	104,3	0,0	104,3
36	98,7	93,5	5,2	96,1	101,5	96,1	5,4	98,8
38	104,3	119,6	-15,3	112,0	101,5	107,2	-5,7	104,4*
Σx			-0,6	1061,3			41,4	1060,9
d			-0,06	106,1			4,1	106,1
Σx^2			416,9				372,5	
S_{diff}			6,8				4,7	

Table showing the agreement between two different software systems, TPS and PCLx, measuring individual speakers' 1st and 2nd order Means on approximately the same speech sample on different occasions.

* The recorded speech sample was too short to allow analysis of 6000 samples.

TABLE 34

TPS1 vs PCLx1

T1 subjects, SPEECH

	1st order Mode (Hz)				2nd order Mode (Hz)			
	TPS1	PCLx1	diff.	Average	TPS1	PCLx1	diff	Average
Subject								
6	116,4	113,3	3,1	114,9	113,2	119,6	-6,4	116,4
7	104,3	104,3	0	114,3	110,2	104,3	5,9	107,3
10	81,5	83,8	-2,3	82,7	96,1	86,1	10,0	91,1
14	148,9	144,9	4,0	146,9	148,9	141,0	7,9	145,0
26	104,3	88,5	15,8	96,4	110,2	91,0	19,2	100,6
27	101,5	93,5	8,0	97,5	101,5	93,5	8,0	97,5
31	77,2	77,2	0	77,2	77,2	77,2	0	77,2
32	107,2	101,5	5,7	104,4	110,2	101,5	8,7	105,9
36	96,1	91,0	5,1	93,6	96,1	91,0	5,1	93,6
38	81,5	81,5	0	81,5	79,3	170,8	-91,5	125,1
Σ x			39,4	999,4			-32,8	1059,7
d			3,9	99,9			-3,3	106,0
Σx^2			403,04				9144,8	
S_{diff}			5,3				31,7	

Example of the agreement between two different software systems TPS and PCLx measuring individual speakers' Speaking 1st and 2nd order Modes on the same occasion.

TABLE 35

TPS1 vs PCLx

T1 subjects, Speech

	2nd order 90% range Minimum (Hz)				2nd order 90 % range Maximum (Hz)			
Subject	TPS1	PCLx1	diff.	Average	TPS1	PCLx1	diff	Average
6	107,2	104,3	2,9	105,8	144,8	144,9	-0,1	144,9
7	88,5	88,5	0,0	88,5	157,3	157,3	0,0	157,3
10	73,0	77,2	-4,2	75,1	153,1	149,0	4,1	151,1*
14	119,6	126,4	6,8	123,0	195,8	180,4	15,4	188,1
26	81,5	81,5	0,0	81,5	137,2	137,2	0,0	137,2
27	88,5	86,1	2,4	87,3	148,8	144,9	3,9	146,9
31	67,3	69,2	1,9	68,3	110,2	107,2	3,0	108,7
32	71,1	71,1	0,0	71,1	185,3	185,5	0,2	185,4
36	83,8	81,5	2,3	82,7	157,3	149,0	8,3	153,2
38	65,5	58,7	6,8	62,1	180,3	175,6	4,7	178,0*
Σx			18,9	845,4			39,5	1550,8
d			1,9	84,5			3,95	155,1
Σx^2			91,6				369,2	
S_{diff}			2,5				4,9	

Table showing the agreement between two different software systems TPS and PCLx measuring individual speakers' 2nd order 90 % range Minima and Maxima on approximately the same speech sample on different occasions.

* The recorded speech sample was too short to allow analysis of 6000 samples.

TABLE 36

TPS1 vs PCLx

T1 subjects, Speech

	%TS into 2nd order				Total sample 1st order	
Subject	TPS1	PCLx1	diff.	Average	TPS1	PCLx1
6	29,1	30,8	-1,7	30,0	6175	8864
7	34,6	31,2	3,4	32,9	6810	7495
10	20,8	33,8	-13,0	27,3	4045*	3563*
14	16,9	24,2	-7,3	20,6	6372	6928
26	23,5	18,9	4,6	21,2	6105	7982
27	27,4	29,8	-2,4	28,6	6018	6841
31	18,0	22,5	-4,5	20,3	6061	6910
32	30,5	29,3	1,2	29,9	6318	7459
36	34,2	37,0	-2,8	35,6	6018	6875
38	6,4	4,6	1,8	5,5	2194*	2445*
Σx			-20,7	251,8	56116	65362
d			- 2,1	25,2	5611,6	6536,2
Σx^2			296,4			
S_{diff}			5,3			

Table showing the agreement between two different software systems, TPS and PCLx, measuring individual speakers' % TS values on approximately the same speech sample on different occasions.

Also showing the Total sample of Tx intervals analysed on the two occasions.

* The recorded speech sample was too short to allow analysis of 6000 samples.

TABLE 37

PCLx1 vs PCLx2

NORMALS, Reading

Subject	1st order Mean (Hz) PCLx1	PCLx2	diff.	Average	2nd order Mean (Hz) PCLx1	PCLx2	diff	Average
DGB	129,9	129,9	0	129,9	129,9	129,9	0	129,9
JSt	88,5	88,5	0	88,5	91,0	91,0	0	91,0
PBr	119,6	119,6	0	119,6	123,0	123,0	0	123,0
RA	149,0	149,0	0	149,0	153,1	153,1	0	153,1
FC	119,6	116,4	3,2	118,0	119,6	119,6	0	119,6
H	123,0	123,0	0	123,0	126,4	126,4	0	126,4
FW	104,3	104,3	0	104,3	91,0	91,0	0	91,0
TVM	129,9	129,9	0	129,9	133,5	133,5	0	133,5
DTH	110,2	113,3	-3,1	111,75	113,3	113,3	0	113,3
RDD	123,0	123,0	0	123,0	129,9	126,4	3,5	128,2
Σx			0,1	1196,95			3,5	1209,0
d			0,01	119,7			0,35	120,9
Σx^2			19,85				12,25	
S_{diff}			1,49				1,11	

Example of individual speakers' Mean Fx values arrived at using the same software (PCLx) on repeated occasions for calculation of Fx parameters on the same reading passage.

APPENDIX 12

TABLE 38

PCLx1 vs PCLx2

Normals, READING

	1st Mode (Hz)				2nd Mode (Hz)			
Subject	PCLx1	PCLx2	diff.	Average	PCLx1	PCLx2	diff	Average
DGB	126,4	126,4	0	126,4	126,4	126,4	0	126,4
JSt	79,3	83,8	-4,5	81,6	79,3	83,8	-4,5	81,6
PBr	113,3	113,3	0	113,3	113,3	113,3	0	113,3
RA	157,3	157,3	0	157,3	157,3	157,3	0	157,3
FC	113,3	113,3	0	113,3	126,4	126,4	0	126,4
H	126,4	126,4	0	126,4	126,4	126,4	0	126,4
FW	73,1	83,8	-10,7	78,5	88,5	86,1	2,4	87,3
TVM	133,5	133,5	0	133,5	133,5	133,5	0	133,5
DTH	110,2	110,2	0	110,2	110,2	110,2	0	110,2
RDD	119,6	116,4	3,2	118,0	116,4	116,4	0	116,4
Σx			-12,0	1158,4			-2,1	1178,75
d			-1,2	115,8			-0,2	117,9
Σx^2			144,98				26,01	
S_{diff}			3,8				1,7	

Table showing the repeatability of individual speakers' Reading 1st and 2nd order Modes arrived at using the same software (PCLx) on repeated occasions on the same reading passage.

APPENDIX 12

TABLE 39

PCLx1 vs PCLx2

NORMALS, Reading

	2nd order 90% range Minimum (Hz)				2nd order 90% range Maximum (Hz)			
Subject	PCLx1	PCLx2	diff.	Average	PCLx1	PCLx2	diff	Average
DGB	113,3	113,3	0	113,3	170,8	175,6	-4,8	173,2
JSt	73,1	75,1	-2,0	74,1	129,9	129,9	0	129,9
PBr	93,5	93,5	0	93,5	195,9	206,9	-11,0	201,4
RA	116,4	119,6	-3,2	118,0	230,9	230,9	0	230,9
FC	101,5	104,3	-2,8	102,9	157,3	157,3	0	157,3
H	107,2	107,2	0	107,2	166,2	166,2	0	166,2
FW	73,1	73,1	0	73,1	126,4	126,4	0	126,4
TVM	113,3	113,3	0	113,3	170,8	170,8	0	170,8
DTH	91,0	91,0	0	91,0	166,2	180,4	-14,2	173,3
RDD	110,2	107,2	3,0	108,7	170,8	170,8	0	170,8
Σx			5,0	995,1			-30,0	1700,2
d			0,5	99,5			-3,0	170,0
Σx^2			31,08				345,68	
s_{diff}			1,78				5,33	

Example of individual speakers' Mean Fx values arrived at using the same software (PCLx) on repeated occasions for calculation of Fx parameters on the same reading passage.

APPENDIX 12

TABLE 40

PCLx1 vs PCLx2

NORMALS, Reading

	% TS into 2nd order				% Irregularity			
Subject	PCLx1	PCLx2	diff.	Average	PCLx1	PCLx2	diff	Average
DGB	48,6	51,0	-2,4	49,8	7,7	4,7	3,0	6,2
JSt	36,6	38,8	-2,2	37,7	17,0	11,5	5,5	14,25
PBr	32,2	32,0	0,2	32,1	14,0	14,9	-0,9	14,45
RA	48,5	49,2	-0,7	48,85	6,8	6,4	0,4	6,6
FC	42,5	40,5	2,0	41,5	13,0	16,0	-3,0	14,5
H	49,8	48,1	1,7	48,95	7,5	9,6	-2,1	8,55
FW	19,9	20,3	-0,4	20,1	42,0	40,9	1,1	41,45
TVM	44,9	42,5	2,4	43,7	11,0	14,5	-3,5	12,75
DTH	38,0	37,0	1,0	37,5	14,3	16,0	-1,7	15,15
RDD	44,8	45,3	-0,5	45,05	5,4	5,4	0	5,4
Σx			-1,1	405,25			-1,2	139,3
d			0,11	40,5			-0,12	13,9
Σx^2			25,19				69,98	
S_{diff}			1,67				2,79	

Example of individual speakers' %TS values and % Irregularity arrived at using the same software (PCLx) on repeated occasions on the same reading passage.

TABLE 41

PCLx1 vs PCLx2

T1 subjects, READING

	1st order Mean (Hz)				2nd order Mean (Hz)			
Subject	PCLx1	PCLx2	diff.	Average	PCLx1	PCLx2	diff	Average
6	119,6	119,6	0	119,6	123,0	123,0	0	123,0
7	98,8	98,8	0	98,8	101,5	101,5	0	101,5
10	98,8	101,5	-2,7	100,2	101,5	104,3	-2,8	102,9
14	149,0	144,9	4,1	147,0	141,0	141,0	0	141,0
26	101,5	101,5	0	101,5	93,5	93,5	0	93,5
27	88,5	88,5	0	88,5	91,0	91,0	0	91,0
31	83,8	86,1	-2,3	85,0	83,8	86,1	-2,3	85,0
32	93,5	91,0	2,5	92,3	96,1	96,1	0	96,1
36	96,1	96,1	0	96,1	96,1	96,1	0	96,1
38	123,0	119,6	3,4	121,3	104,3	104,3	0	104,3
Σx			0,9	1050,3			-5,1	1034,4
d			0,09	105,0			-0,5	103,4
Σx^2			47,02				13,1	
SD_{diff}			2,3				1,08	

Table showing the repeatability of individual speakers' 1st and 2nd order Means arrived at using the same software PCLx, on the same reading passage on two different occasions

APPENDIX 12

TABLE 42

PCLx1 vs PCLx2

T1 subjects, READING

	1st order Mode (Hz)				2nd order Mode (Hz)			
Subject	PCLx1	PCLx2	diff.	Average	PCLx1	PCLx2	diff	Average
6	116,4	119,6	-3,2	118,0	119,6	119,6	0	119,6
7	101,5	101,5	0	101,5	101,5	101,5	0	101,5
10	91,0	91,0	0	91,0	104,3	104,3	0	104,3
14	153,1	153,1	0	153,1	153,1	153,1	0	153,1
26	86,1	86,1	0	86,1	86,1	86,1	0	86,1
27	93,5	93,5	0	93,5	93,5	93,5	0	93,5
31	79,3	79,3	0	79,3	81,5	83,8	-2,3	82,7
32	88,5	88,5	0	88,5	91,0	91,0	0	91,0
36	88,5	88,5	0	88,5	88,5	88,5	0	88,5
38	83,8	83,8	0	83,8	83,8	83,8	0	83,8
Σ x			-3,2	983,3			-2,3	1004,05
d			-0,3	98,3			-0,2	100,4
Σx^2			10,24				5,29	
S_{diff}			1,0				0,7	

Table showing the repeatability of individual speakers' Reading 1st and 2nd order Modes arrived at using the same software (PCLx) on repeated occasions on the same reading passage.

TABLE 43

PCLx1 vs PCLx2

T1 subjects, READING

	2nd order 90% range Minimum (Hz)				2nd order 90% range Maximum (Hz)			
Subject	PCLx1	PCLx2	diff.	Average	PCLx1	PCLx2	diff	Average
6	113,3	110,2	3,1	111,8	149,0	144,9	4,1	147,0
7	83,8	83,8	0	83,8	137,2	137,2	0	137,2
10	86,1	86,1	0	86,1	141,0	141,0	0	141,0
14	119,6	123,0	-3,4	121,3	170,8	170,8	0	170,8
26	81,5	83,8	-2,3	82,7	123,0	123,0	0	123,0
27	83,8	81,5	2,3	82,7	113,3	116,4	-3,1	114,9
31	75,1	75,1	0	75,1	110,2	113,3	-3,1	111,8
32	79,3	79,3	0	79,3	129,9	129,9	0	129,9
36	81,5	79,3	2,2	80,4	137,2	141,0	-3,8	139,1
38	75,1	75,1	0	75,1	175,6	175,6	0	175,6
Σx			1,9	878,2			-5,9	1390,2
d			0,2	87,8			-0,6	139,0
Σx^2			36,59				50,47	
SD_{diff}			2,01				2,3	

Table showing the repeatability of individual speakers' 2nd order 90 % range Minima and Maxima arrived at using the same software (PCLx) on repeated occasions on the same reading passage.

TABLE 44

PCLx1 vs PCLx2

T1 subjects, READING

	% TS into 2nd order				% Irregularity			
Subject	PCLx1	PCLx2	diff.	Average	PCLx1	PCLx2	diff	Average
6	27,2	26,3	0,9	26,8	25,9	28,7	-2,8	27,3
7	34,3	34,4	-0,1	34,4	17,4	17,8	-0,4	17,6
10	37,8	37,9	-0,1	37,9	10,4	10,9	-0,5	10,7
14	13,9	15,04	-1,1	14,5	50,7	48,6	2,1	49,7
26	21,9	21,2	0,7	21,6	35,8	36,5	0,7	36,6
27	24,4	23,2	1,2	23,8	29,0	29,6	-0,6	29,3
31	35,5	35,5	0	35,5	12,8	12,1	0,7	12,5
32	36,2	35,5	0,7	35,9	12,7	13,2	-0,5	12,95
36	34,4	36,0	-1,6	35,2	16,5	15,0	1,5	15,8
38	6,5	6,5	0	6,5	78,0	78,0	0	78,0
Σx			0,58	271,8			0,2	290,2
d			0,06	27,2			0,02	29,0
Σx^2			7,45				16,5	
SD_{diff}			0,9				1,4	

Table showing the repeatability of individual speakers' % TS and % Irregularity arrived at using the same software PCLx, on the same reading passage on two different occasions .

APPENDIX 12

TABLE 45

PCLx1 vs PCLx2

NORMALS Speech

Subject	1st order Mean (Hz) PCLx1	PCLx2	diff.	Average	2nd order Mean (Hz) PCLx1	PCLx2	diff	Average
DGB	137,2	141,0	-3,8	139,1	133,5	141,0	-7,5	137,3
JSt	88,5	88,5	0	88,5	86,1	88,5	-2,4	87,3
PBr	116,4	116,4	0	116,4	116,4	119,6	-3,2	118,0
RA	137,2	137,2	0	137,2	141,0	144,9	-3,9	143,0
FC	110,2	110,2	0	110,2	116,4	116,4	0	116,4
H	113,3	113,3	0	113,3	116,4	116,4	0	116,4
FW	119,6	129,9	-10,3	124,8	93,5	96,1	-2,6	94,8
TVM	113,3	116,4	-3,1	114,9	116,4	116,4	0	116,4
DTH	133,5	133,5	0	133,5	137,2	137,2	0	137,2
RDD	126,4	126,4	0	126,4	129,9	129,9	0	129,9
Σx			-17,2	1204,3			-19,6	1196,6
d			-1,72	120,4			-1,96	119,7
Σx^2			130,14				94,22	
S_{diff}			3,34				2,49	

Example of individual speakers' Mean Fx values arrived at using the same software (PCLx) on repeated occasions for calculation of Fx parameters on the same reading passage.

APPENDIX 12

TABLE 46

PCLx1 vs PCLx2

Normals, SPEECH

	1st Mode (Hz)				2nd Mode (Hz)			
Subject	PCLx1	PCLx2	diff.	Average	PCLx1	PCLx2	diff	Average
DGB	123,0	123,0	0	123,0	123,0	123,0	0	123,0
JSt	93,5	93,5	0	93,5	93,5	93,5	0	93,5
PBr	96,1	96,1	0	96,1	101,5	101,5	0	101,5
RA	144,9	144,9	0	144,9	153,1	153,1	0	153,1
FC	93,5	93,5	0	93,5	101,5	101,5	0	101,5
H	104,3	101,5	2,8	102,9	104,3	96,1	8,2	100,2
FW	88,5	88,5	0	88,5	88,5	86,1	2,4	87,3
TVM	113,3	113,3	0	113,3	116,4	113,3	3,1	114,9
DTH	123,0	119,6	3,4	121,3	123,0	119,6	3,4	121,3
RDD	133,5	126,4	7,1	130,0	133,5	137,2	-3,7	135,4
Σx			13,3	1107,0			13,4	1131,7
d			1,3	110,7			1,3	113,2
Σx^2			69,81				107,86	
S_{diff}			2,4				3,2	

Table showing the repeatability of individual speakers' Speaking 1st and 2nd order Modes arrived at using the same software (PCLx) on repeated occasions.

TABLE 47

PCLx1 vs PCLx2

NORMALS, Speech

	2nd order 90% range Minimum (Hz)				2nd order 90% range Maximum (Hz)			
Subject	PCLx1	PCLx2	diff.	Average	PCLx1	PCLx2	diff	Average
DGB	113,3	116,4	-3,1	114,9	196,0	224,7	-28,8	210,3
JSt	67,3	65,5	1,8	66,4	116,4	126,4	-10,0	121,4
PBr	91,0	93,5	-2,5	92,3	195,9	201,3	- 5,4	198,6
RA	107,2	104,3	2,9	105,8	212,7	218,6	- 5,9	215,7
FC	91,0	91,0	0	91,0	195,9	195,9	0	195,9
H	96,1	96,1	0	96,1	166,2	170,8	-4,6	168,5
FW	71,1	71,1	0	71,1	175,6	185,5	-9,9	180,6
TVM	101,5	101,5	0	101,5	153,1	153,1	0	153,1
DTH	93,5	93,5	0	93,5	287,5	295,5	-8,0	291,5
RDD	110,2	107,2	3,0	108,7	180,4	175,6	4,8	178,0
Σx			2,1	941,15			-67,8	1913,5
d			0,21	94,12			- 6,78	191,3
Σx^2			36,51				1199,62	
S_{diff}			2,00				9,07	

Example of individual speakers' Mean Fx values arrived at using the same software (PCLx) on repeated occasions for calculation of Fx parameters on the same reading passage.

TABLE 48

PCLx1 vs PCLx2

NORMALS, Speech

	% TS into 2nd order				% Irregularity			
Subject	PCLx1	PCLx2	diff.	Average	PCLx1	PCLx2	diff	Average
DGB	46,4	47,6	-1,2	47,0	10,0	7,3	2,7	8,7
JSt	27,8	30,0	-2,2	28,9	24,6	19,6	5,0	22,1
PBr	36,0	37,0	-1,0	36,5	14,2	13,3	0,9	13,8
RA	46,4	45,7	0,7	46,1	7,9	7,2	0,7	7,6
FC	39,1	36,5	2,6	37,8	11,8	15,5	-3,7	13,7
H	39,0	35,5	3,5	37,3	14,3	18,7	-4,4	16,5
FW	11,7	9,7	-2,0	10,7	64,9	67,9	-3,0	66,4
TVM	39,3	40,0	-0,7	39,7	17,2	16,4	0,8	16,8
DTH	35,6	36,2	-0,6	35,9	13,5	13,7	-0,2	13,6
RDD	48,3	45,8	2,5	47,1	4,6	5,2	-0,6	4,9
Σx			1,58	366,81			-1,8	184,1
d			0,16	36,68			-0,18	18,41
Σx^2			37,92				77,04	
S_{diff}			2,05				2,9	

Example of individual speakers' %TS values and % Irregularity arrived at using the same software (PCLx) on repeated occasions on the same reading passage.

APPENDIX 12

TABLE 49

PCLx1 vs PCLx2

T1 subjects, SPEECH

	1st order Mean (Hz)				2nd order Mean (Hz)			
	PCLx1	PCLx2	diff.	Average	PCLx1	PCLx2	diff	Average
Subject								
6	113,3	116,4	-3,1	114,9	116,4	116,4	0	116,4
7	107,2	104,3	2,9	105,8	107,2	107,2	0	107,2
10	91,0	91,0	0	91,0	91,0	91,0	0	91,0
14	141,0	141,0	0	141,0	141,0	141,0	0	141,0
26	104,3	107,2	-2,9	105,8	96,1	98,8	-2,7	97,5
27	98,8	96,1	2,7	97,5	101,5	101,5	0	101,5
31	88,5	86,1	2,4	87,3	79,3	79,3	0	79,3
32	104,3	101,5	2,8	102,9	104,3	101,5	2,8	102,9
36	93,5	93,5	0	93,5	96,1	96,1	0	96,1
38	119,6	119,6	0	119,6	107,2	107,2	0	107,2
Σx			4,8	932,2			0,1	1040,1
d			0,48	93,2			0,01	104,0
Σx^2			46,72				15,13	
SD_{diff}			2,2				1,3	

Table showing the repeatability of individual speakers' 1st and 2nd order Speech Mean Fx values arrived at using the same software (PCLx) on repeated occasions.

APPENDIX 12

TABLE 50

PCLx1 vs PCLx2

T1 subjects, SPEECH

	1st order Mode (Hz)				2nd order Mode (Hz)			
Subject	PCLx1	PCLx2	diff.	Average	PCLx1	PCLx2	diff	Average
6	113,3	110,2	3,1	111,8	119,6	119,6	0	119,6
7	104,3	98,8	5,5	101,6	104,3	104,3	0	104,3
10	83,8	81,5	2,3	82,7	86,1	86,1	0	86,1
14	144,9	144,9	0	144,9	141,0	141,0	0	141,0
26	88,5	88,5	0	88,5	91,0	96,1	-5,1	93,6
27	93,5	96,1	-2,6	94,8	93,5	98,8	-5,3	96,2
31	77,2	77,2	0	77,2	77,2	79,3	-2,1	78,3
32	101,5	101,5	0	101,5	101,5	101,5	0	101,5
36	91,0	91,0	0	91,0	91,0	91,0	0	91,0
38	81,5	81,5	0	81,5	170,8	170,8	0	170,8
Σx			8,3	975,5			-12,5	1082,4
d			0,8	97,6			-1,3	108,2
Σx^2			51,91				58,51	
S_{diff}			2,2				2,2	

Table showing the repeatability of individual speakers' Speaking 1st and 2nd order Modes arrived at using the same software (PCLx) on repeated occasions.

TABLE 51

PCLx1 vs PCLx2

T1 subjects, SPEECH

	2nd order 90% range Minimum (Hz)				2nd order 90% range Maximum (Hz)			
	PCLx1	PCLx2	diff.	Average	PCLx1	PCLx2	diff	Average
Subject								
6	104,3	104,3	0	104,3	144,9	144,9	0	144,9
7	88,5	88,5	0	88,5	157,3	153,1	4,2	155,2
10	77,2	77,2	0	77,2	149,0	133,5	15,5	141,3
14	126,4	126,4	0	126,4	180,4	180,4	0	180,4
26	81,5	81,5	0	81,5	137,2	137,2	0	137,2
27	86,1	86,1	0	86,1	144,9	141,0	3,9	143,0
31	69,2	69,2	0	69,2	107,2	107,2	0	107,2
32	71,1	71,1	0	71,1	185,5	175,6	9,9	180,6
36	81,5	81,5	0	81,5	149,0	149,0	0	149,0
38	58,7	63,7	-5,0	61,2	175,6	175,6	0	175,6
Σx			-5,0	847,0			33,5	1514,3
d			-0,5	84,7			3,4	151,4
Σx^2			25,0				371,1	
SD_{diff}			1,6				5,4	

Table showing the repeatability of individual speakers' 2nd order 90% Speech range Minima and Maxima arrived at using the same software (PCLx) on repeated occasions.

TABLE 52

PCLx1 vs PCLx2

T1 subjects, SPEECH

	% TS into 2nd order				% Irregularity			
	PCLx1	PCLx2	diff.	Average	PCLx1	PCLx2	diff	Average
Subject								
6	30,8	29,3	1,5	30,1	20,8	24,4	-3,6	22,6
7	31,2	31,5	-0,5	31,4	27,6	27,4	0,2	27,5
10	33,8	32,8	1,0	33,3	17,6	18,9	-1,3	18,3
14	24,2	23,1	1,1	23,7	37,0	40,2	-3,2	38,6
26	18,9	17,9	1,0	18,4	42,3	44,6	-2,3	43,5
27	29,8	28,3	1,5	29,1	21,6	25,3	-3,7	23,5
31	22,5	22,1	0,4	22,3	27,4	27,6	-0,2	27,5
32	29,3	30,7	-1,4	30,0	24,3	22,7	1,6	23,5
36	37,0	37,04	-0,04	37,0	14,2	13,9	0,3	14,1
38	4,6	4,7	-0,1	4,7	78,2	77,9	0,3	78,1
Σx			4,66	259,77			-11,9	317,0
d			0,47	25,98			-1,2	31,7
Σx^2			9,93				52,69	
SD_{diff}			0,9				2,1	

Table showing the repeatability of individual speakers' Speech %TS values and % Irregularity arrived at using the same software (PCLx) on repeated occasions.

TABLE 53

TPS 1 VS TPS 2

NORMALS, Speech

Subject	1st order Mean (Hz)				2nd order Mean (Hz)			
	TPS 1	TPS 2	diff.	Average	TPS 1	TPS 2	diff.	Average
X	110,2	-	-	-	113,2	-	-	-
JSt	88,5	86,1	2,4	87,3	93,5	90,9	3,5	92,2
PBr	116,4	116,4	0,0	116,4	119,6	116,4	3,2	118,0
RA	144,9	144,9	0,0	144,9	153,1	148,9	4,2	151,0
FC	119,6	116,4	3,2	118,0	122,9	119,6	3,3	121,3
H	122,9	116,4	6,5	119,7	126,4	122,9	3,5	124,7
FW	116,4	83,8	32,9	100,3	90,9	86,1	4,8	88,5
TVM	122,9	119,6	3,3	121,3	126,4	122,9	3,5	124,7
DTH	137,2	141,0	-3,8	139,1	144,9	141,0	3,9	143,0
RDD	129,9	126,4	3,5	128,2	137,2	129,9	7,3	133,6
Σx			48,0	1075,2			37,2	1097,0
d			5,3	119,5			4,1	121,9
Σx^2			1178,2				167,1	
S_{diff}			1,1				1,3	

Raw data of individual speakers' Mean Fx values arrived at using the same softwareTPS, for analysis of approximately the same speech sample on two different occasions 1 and 2. Mean differences (d), Standard deviation of these differences (SD_{diff}).

TABLE 54

TPS 1 VS TPS 2

NORMALS, Speech

Subject	2nd order 90% range (Hz) Minimum (Hz)				2nd order 90% range Maximum (Hz)			
	TPS 1	TPS 2	diff.	Average	TPS 1	TPS 2	diff.	Average
X	81,5	-	-	-	212,7	-	-	-
JSt	69,1	65,5	3,6	67,3	133,5	119,6	13,9	126,6
PBr	93,5	88,5	5,0	91,0	195,8	206,9	-11,1	201,4
RA	107,2	107,2	0,0	107,2	212,7	224,7	-12,0	218,7
FC	93,5	93,5	0,0	93,5	201,3	201,3	0,0	201,3
H	98,6	98,6	0,0	98,6	180,3	166,2	14,1	173,3
FW	75,1	69,1	6,0	72,1	180,3	175,6	4,7	178,0
TVM	104,3	104,3	0,0	104,3	161,7	157,3	4,4	159,5
DTH	93,5	96,1	-2,6	94,8	303,7	287,5	16,2	295,6
RDD	110,2	104,3	5,9	107,3	185,3	175,6	9,7	180,5
Σx			17,9	836,1			39,9	1734,7
d			1,99	92,9			4,4	192,7
Σx^2			115,5				1057,2	
S_{diff}			3,2				10,5	

Raw data of individual speakers' Mean Fx values arrived at using the same software TPS, for analysis of approximately the same speech sample on two different occasions 1 and 2. Mean differences (d), Standard deviation of these differences (SD_{diff})

TABLE 55

TPS 1 VS TPS 2

NORMALS, Speech

%TS into 2nd order

Subject	TPS 1	TPS 2	diff.	Average
X	28,9	-	-	-
JSt	27,7	26,7	1,0	27,2
PBr	38,7	24,2	14,5	31,5
RA	45,8	36,9	8,9	41,4
FC	41,5	44,2	-2,7	42,9
H	39,0	38,6	0,4	38,8
FW	18,6	22,8	-4,2	20,7
TVM	40,1	43,3	-3,2	41,7
DTH	34,7	36,8	-2,1	35,8
RDD	39,9	33,9	6,0	36,9
Σx			18,6	316,9
d			2,1	35,2
Σx^2			366,2	
S_{diff}			6,4	

Raw data of individual speakers' % Total Sample carried into second order, arrived at using the same software TPS, for analysis of approximately the same speech sample on two different occasions 1 and 2, Mean differences (d), Standard deviation of these differences (SD_{diff})

APPENDIX 12

TABLE 56

TPS 1 VS TPS 2

Tl subjects, Speech

Subject	1st order Mean (Hz)				2nd order Mean (Hz)			
	TPS 1	TPS 2	diff.	Average	TPS 1	TPS 2	diff.	Average
6	119,6	119,6	0,0	119,6	119,6	119,6	0,0	119,6
7	107,2	107,2	0,0	107,2	113,2	113,2	0,0	113,2
10	93,5	90,9	2,6	92,2	96,1	98,7	-2,6	97,4
14	141,0	141,0	0,0	141,0	148,9	148,9	0,0	148,9
26	113,2	107,2	6,0	110,2	101,5	104,3	-2,8	102,9
27	122,9	107,2	15,7	115,1	113,2	110,2	3,0	111,7
31	86,1	86,1	0,0	86,1	79,3	79,3	0,0	79,3
32	104,3	98,7	5,6	101,5	107,2	104,3	2,9	105,8
36	96,1	98,7	-2,6	97,4	98,7	101,5	-2,8	100,1
38	104,3	104,3	0,0	104,3	101,5	101,5	0,0	101,5
Σx			27,3	1074,6			-2,3	1080,4
d			2,7	107,5			-0,2	108,0
Σx^2			327,4				30,9	
S_{diff}			3,0				1,8	

Raw data of individual speakers' Mean Fx values arrived at using the same softwareTPS, for analysis of approximately the same speech sample on two different occasions 1 and 2. Mean differences (d), Standard deviation of these differences (SD_{diff}).

TABLE 57

TPS 1 VS TPS 2

T1 subjects, Speech

Subject	2nd order 90% range (Hz) Minimum (Hz)				2nd order 90% range Maximum (Hz)			
	TPS 1	TPS 2	diff.	Average	TPS 1	TPS 2	diff.	Average
6	107,2	107,2	0,0	107,2	144,8	144,8	0,0	144,8
7	90,9	88,5	2,4	89,7	153,1	157,3	-4,2	155,2
10	71,1	73,0	-1,9	72,1	153,1	153,1	0,0	153,1
14	119,6	119,6	0,0	119,6	190,6	195,8	-5,2	193,2
26	77,1	81,5	-4,4	79,3	137,2	137,2	0,0	137,2
27	88,5	88,5	0,0	88,5	161,7	148,8	12,9	155,3
31	67,3	67,3	0,0	67,3	107,2	110,2	-3,0	108,7
32	71,1	71,1	0,0	71,1	180,3	185,3	-5,0	182,8
36	81,5	83,8	-2,3	82,7	144,8	157,3	-12,5	151,1
38	60,3	65,5	-5,2	62,9	175,6	180,3	-4,7	178,0
Σx			-11,4	840,4			-21,7	1559,4
d			-1,1	84,0			-2,2	155,9
Σx^2			61,7				423,4	
S_{diff}			2,3				6,5	

Raw data of individual speakers' 2nd order 90% range Fx values arrived at using the same software TPS, for analysis of approximately the same speech sample on two different occasions 1 and 2. Mean differences (d), Standard deviation of these differences (SD_{diff})

APPENDIX 12

TABLE 58

TPS 1 VS TPS 2

T1 subjects, Speech

%TS into 2nd order

Subject	TPS 1	TPS 2	diff.	Average
6	30,1	29,1	1,5	29,9
7	33,3	34,6	-1,3	34,0
10	17,5	20,8	-3,3	19,2
14	19,0	16,9	2,1	18,0
26	17,5	23,5	-6,0	20,5
27	20,7	27,4	-6,7	24,1
31	17,5	18,0	-0,5	17,8
32	27,6	30,5	-2,9	29,1
36	34,8	34,2	0,6	34,5
38	6,8	6,4	0,4	6,6
Σx			-16,1	233,7
d			-1,6	23,4
Σx^2			109,3	
S_{diff}			3,0	

Raw data of individual speakers' % Total Sample carried into second order, arrived at using the same software TPS, for analysis of approximately the same speech sample on two different occasions 1 and 2. Mean differences (d), Standard deviation of these differences (SD_{diff}).

APPENDIX 13

Frequency tables for calculation of interrater agreement on 20 random voices.

Table 1 HARSHNESS - Raters AP vs PR

Observed cell frequences:

Rater AP

Scalar degree:		0	1	2	3	4	5	6	Row total
0		0	0	0	0	0	0	0	0
1	R a	0	1	1	0	0	0	0	2
2	t e	0	1	4	1	0	0	0	6
3	r	0	0	1	1	0	0	0	2
4	P R	1	1	0	2	4	0	0	8
5		0	0	0	1	1	0	0	2
6		0	0	0	0	0	0	0	0
Column Total		1	3	6	5	5	0	0	20

Table 2 WHISPER - Raters AP vs PR

Observed cell frequences:

Rater AP

Scalar degree:		0	1	2	3	4	5	6	Row total
0		0	0	0	0	0	0	0	0
1	R a	0	1	1	1	1	0	0	4
2	t e	2	2	5	0	0	0	0	9
3	r	0	0	2	1	0	0	0	3
4	P R	0	0	1	0	2	1	0	4
5		0	0	0	0	0	0	0	0
6		0	0	0	0	0	0	0	0
Column Total		2	3	9	2	3	1	0	20

APPENDIX 13

Table 3 CREAK - Raters: AP vs PR

Observed cell frequences:

Scalar degree:		0	1	2	3	4	5	6	Row total
				Rater AP					
0		2	0	0	0	0	0	0	2
1	Rater PR	6	1	1	0	0	0	0	8
2		1	0	1	0	0	0	0	2
3		2	0	0	0	0	0	0	2
4		0	1	0	0	1	2	0	4
5		0	0	0	0	0	2	0	2
6		0	0	0	0	0	0	0	0
Column Total		11	2	2	0	1	4	0	20

Table 4 LARYNGEAL TENSION - Raters: AP vs PR

Observed cell frequences:

Scalar degree:		0	1	2	3	4	5	6	Row total
				Rater AP					
0		3	0	0	0	0	0	0	3
1	Rater PR	1	2	1	0	0	0	0	4
2		0	7	1	1	0	0	0	9
3		0	0	0	2	1	0	0	3
4		0	0	0	1	0	0	0	1
5		0	0	0	0	0	0	0	0
6		0	0	0	0	0	0	0	0
Column Total		4	9	2	4	1	0	0	20

Table 5 PHARYNGEAL CONSTRICTION - Raters: AP vs PR

Observed cell frequences:

Scalar degree:		0	1	2	3	4	5	6	Row total
					Rater AP				
0		12	1	0	0	0	0	0	13
1	R a	3	2	0	0	0	0	0	5
2	t e	0	1	0	0	0	0	0	1
3	r	0	0	0	0	0	0	0	0
4	P R	0	1	0	0	0	0	0	1
5		0	0	0	0	0	0	0	0
6		0	0	0	0	0	0	0	0
Column Total		15	5	0	0	0	0	0	20

Table 6 HARSHNESS - Raters EC vs AP

Observed cell frequences:

Scalar degree:		0	1	2	3	4	5	6	Row total
					Rater AP				
0		1	0	0	0	0	0	0	1
1	R a	1	2	0	0	0	0	0	3
2	t e	1	3	1	0	0	1	0	6
3	r	1	1	2	0	0	1	0	5
4	E C	0	1	1	0	2	1	0	5
5		0	0	0	0	0	0	0	0
6		0	0	0	0	0	0	0	0
Column Total		4	7	4	0	2	3	0	20

APPENDIX 13

Table 7 WHISPER - Raters EC vs AP

Observed cell frequences:

Rater AP

Scalar degree:		0	1	2	3	4	5	6	Row total
0		1	0	0	0	0	0	0	1
1	Rater EC	0	1	1	1	0	0	0	3
2		1	2	8	0	1	0	0	12
3		0	0	0	1	0	0	0	1
4		0	0	0	0	1	1	0	2
5		0	0	0	0	1	0	0	1
6		0	0	0	0	0	0	0	0
Column Total		2	3	9	2	3	1	0	20

Table 8 CREAK - Raters EC vs AP

Observed cell frequences:

Rater AP

Scalar degree:		0	1	2	3	4	5	6	Row total
0		5	3	1	1	1	0	0	11
1	Rater EC	1	0	1	0	0	0	0	2
2		1	0	1	0	0	0	0	2
3		0	0	0	0	0	0	0	0
4		1	0	0	0	0	0	0	1
5		1	0	1	0	1	1	0	4
6		0	0	0	0	0	0	0	0
Column Total		9	3	4	1	2	1	0	20

Table 9 LARYNGEAL TENSION - Raters EC vs AP

Observed cell frequences:

Scalar degree:		0	1	2	3	4	5	6	Row total
				Rater AP					
0		0	0	3	1	0	0	0	4
1	R	0	2	5	0	2	0	0	9
2	a t	0	0	0	1	1	0	0	2
3	e r	0	0	0	0	4	0	0	4
4	E	0	0	0	0	0	1	0	1
5	C	0	0	0	0	0	0	0	0
6		0	0	0	0	0	0	0	0
Column Total		0	2	8	2	7	1	0	20

Table 10 PHARYNGEAL CONSTRICTION - Raters EC vs AP

Observed cell frequences:

Scalar degree:		0	1	2	3	4	5	6	Row total
				Rater AP					
0		3	5	3	2	2	0	0	15
1	R	0	1	0	0	4	0	0	5
2	a t	0	0	0	0	0	0	0	0
3	e r	0	0	0	0	0	0	0	0
4	E	0	0	0	0	0	0	0	0
5	C	0	0	0	0	0	0	0	0
6		0	0	0	0	0	0	0	0
Column Total		3	6	3	2	6	0	0	20

APPENDIX 13

Table 11 HARSHNESS - Raters EC vs PR

Observed cell frequences:

Scalar degree:		0	1	2	3	4	5	6	Row total
				Rater PR					
0		0	0	0	0	0	0	0	0
1	R a	0	2	0	0	0	0	0	2
2	t e	2	2	1	0	0	1	0	6
3	r	1	0	1	0	0	0	0	2
4	E C	1	3	1	0	1	2	0	8
5		0	0	1	0	1	0	0	2
6		0	0	0	0	0	0	0	0
Column Total		4	7	4	0	2	3	0	20

Table 12 WHISPER - Raters EC vs PR

Observed cell frequences:

Scalar degree:		0	1	2	3	4	5	6	Row total
				Rater PR					
0		0	0	2	1	0	0	0	3
1	R a	0	0	3	1	0	0	0	4
2	t e	0	2	3	0	4	0	0	9
3	r	0	0	0	0	2	1	0	3
4	E C	0	0	0	0	0	1	0	1
5		0	0	0	0	0	0	0	0
6		0	0	0	0	0	0	0	0
Column Total		0	2	8	2	6	2	0	20

Table 13 CREAK - Raters EC vs PR

Observed cell frequences: Row total

Scalar degree:		0	1	2	3	4	5	6	Row total
					Rater PR				
0		1	0	0	0	0	0	0	1
1	R a	3	3	1	1	0	0	0	8
2	t e	2	0	1	0	0	0	0	3
3	r	1	0	0	0	1	0	0	2
4	E C	2	0	1	0	1	0	0	4
5		0	0	1	0	0	1	0	2
6		0	0	0	0	0	0	0	0
Column Total		9	3	4	1	2	1	0	20

Table 14 LARYNGEAL TENSION - Raters EC vs PR

Observed cell frequences: Row total

Scalar degree:		0	1	2	3	4	5	6	Row total
					Rater PR				
0		0	0	2	1	0	0	0	3
1	R a	0	0	3	1	0	0	0	4
2	t e	0	2	3	0	4	0	0	9
3	r	0	0	0	0	2	1	0	3
4	E C	0	0	0	0	0	1	0	1
5		0	0	0	0	0	0	0	0
6		0	0	0	0	0	0	0	0
Column Total		0	2	8	2	6	2	0	20

Table 15 PHARYNGEAL CONSTRICTION - Raters EC vs PR

Observed cell frequences:

Scalar degree:		0	1	2	3	4	5	6	Row total
				Rater PR					
0		3	3	3	2	1	0	0	12
1	Rater EC	0	3	0	0	3	0	0	6
2		0	0	0	0	1	0	0	1
3		0	0	0	0	0	0	0	0
4		0	0	0	0	1	0	0	1
5		0	0	0	0	0	0	0	0
6		0	0	0	0	0	0	0	0
Column Total		3	6	3	2	6	0	0	20

APPENDIX 14

Frequency tables for calculation of interrater agreement on 37 irradiated voices.

Table 1 HARSHNESS - Raters AP vs PR

Observed cell frequences:

Scalar degree:	0	1	2	3	4	5	6	Row total
Rater PR			Rater AP					
0	1	0	0	0	0	0	0	1
1	0	0	1	1	0	0	0	2
2	0	2	6	4	0	0	0	12
3	0	0	5	4	2	1	0	12
4	0	0	0	2	5	0	0	7
5	0	0	1	1	1	0	0	3
6	0	0	0	0	0	0	0	0
Column Total	1	2	13	12	8	1	0	37

Table 2 WHISPER - Raters AP vs PR

Observed cell frequences:

Scalar degree:	0	1	2	3	4	5	6	Row total
Rater PR			Rater AP					
0	0	1	1	0	0	0	0	2
1	1	8	2	2	0	0	0	13
2	0	3	8	1	0	0	0	12
3	0	0	1	4	1	0	0	6
4	0	0	0	1	1	1	0	3
5	0	0	0	0	0	1	0	1
6	0	0	0	0	0	0	0	0
Column Total	1	12	12	8	2	2	0	37

APPENDIX 14

Table 3 CREAK - Raters: AP vs PR

Observed cell frequences:

Scalar degree:		0	1	2	3	4	5	6	Row total
				Rater AP					
0		4	1	0	0	0	0	0	5
1	Rater PR	5	2	1	0	1	0	0	9
2		1	1	2	2	0	0	0	6
3		0	0	2	0	5	0	0	7
4		0	0	1	0	4	3	0	8
5		0	0	0	0	0	1	1	2
6		0	0	0	0	0	0	0	0
Column Total		10	4	6	2	10	4	1	37

Table 4 LARYNGEAL TENSION - Raters: AP vs PR

Observed cell frequences:

Scalar degree:		0	1	2	3	4	5	6	Row total
				Rater AP					
0		3	2	0	0	0	0	0	5
1	Rater PR	0	5	4	0	1	0	0	10
2		0	7	5	2	0	0	0	14
3		0	0	1	2	4	0	0	7
4		0	0	0	1	0	0	0	1
5		0	0	0	0	0	0	0	0
6		0	0	0	0	0	0	0	0
Column Total		3	14	10	5	5	0	0	37

Table 5 HARSHNESS - Raters EC vs AP

Observed cell frequences:

Scalar degree:		0	1	2	3	4	5	6	Row total
					Rater AP				
0		0	0	1	0	0	0	0	1
1	R a	0	1	0	1	0	0	0	2
2	t e	0	1	2	6	3	1	0	13
3	r	2	1	4	4	1	0	0	12
4	E C	0	1	0	3	2	2	0	8
5		0	0	1	0	0	0	0	1
6		0	0	0	0	0	0	0	0
Column Total		2	4	8	14	6	3	0	37

Table 6 WHISPER - Raters EC vs AP

Observed cell frequences:

Scalar degree:		0	1	2	3	4	5	6	Row total
					Rater AP				
0		0	1	0	0	0	0	0	1
1	R a	0	1	7	4	0	0	0	12
2	t e	0	0	4	5	1	1	0	11
3	r	0	0	3	1	2	2	0	8
4	E C	0	0	1	0	1	1	0	3
5		0	0	0	0	0	1	1	2
6		0	0	0	0	0	0	0	0
Column Total		0	2	15	10	4	5	1	37

Table 7 CREAK - Raters EC vs AP

Observed cell frequences: Row total

Scalar degree:		0	1	2	3	4	5	6	Row total
		Rater AP							
0		5	0	3	1	1	0	0	10
1	Rater EC	1	0	3	0	0	0	0	4
2		0	0	2	3	0	1	0	6
3		0	0	0	1	1	0	0	2
4		0	1	0	3	5	1	0	10
5		1	0	0	1	2	0	0	4
6		0	0	0	0	0	1	0	1
Column Total		7	1	8	9	9	3	0	37

Table 8 LARYNGEAL TENSION - Raters EC vs AP

Observed cell frequences: Row total

Scalar degree:		0	1	2	3	4	5	6	Row total
		Rater AP							
0		0	0	2	1	0	0	0	3
1	Rater EC	0	2	7	6	0	0	0	15
2		0	0	1	3	4	1	0	9
3		0	0	0	0	1	3	1	4
4		0	0	0	3	2	0	0	5
5		0	0	0	0	0	0	0	0
6		0	0	0	0	0	0	0	0
Column Total		0	2	10	13	7	4	1	37

Table 9 HARSHNESS - Raters EC vs PR

Observed cell frequences:									Row total
Scalar degree:		0	1	2	3	4	5	6	
					Rater PR				
0		0	0	1	0	0	0	0	1
1	R a	0	1	0	1	0	0	0	2
2	t e	1	2	2	6	1	0	0	12
3	r	1	0	5	4	1	1	0	12
4	E C	0	1	0	2	3	1	0	7
5		0	0	0	1	1	1	0	3
6		0	0	0	0	0	0	0	0
Column Total		2	4	8	14	6	3	0	37

Table 10 WHISPER - Raters EC vs PR

Observed cell frequences:									Row total
Scalar degree:		0	1	2	3	4	5	6	
					Rater PR				
0		0	1	0	1	0	0	0	2
1	R a	0	1	8	4	0	0	0	13
2	t e	0	0	6	3	2	1	0	12
3	r	0	0	0	2	1	3	0	6
4	E C	0	0	0	0	2	1	0	3
5		0	0	0	0	0	0	1	1
6		0	0	0	0	0	0	0	0
Column Total		0	2	14	10	5	5	1	37

APPENDIX 14

Table 11 CREAK - Raters EC vs PR

Observed cell frequences:

Scalar degree:		0	1	2	3	4	5	6	Row total
				Rater PR					
0		4	0	1	0	0	0	0	5
1	Rater EC	2	0	4	1	2	0	0	9
2		0	0	3	2	1	0	0	6
3		0	0	0	3	2	2	0	7
4		0	1	0	3	4	0	0	8
5		1	0	0	0	0	1	0	2
6		0	0	0	0	0	0	0	0
Column Total		7	1	8	9	9	3	0	37

Table 12 LARYNGEAL TENSION - Raters EC vs PR

Observed cell frequences:

Scalar degree:		0	1	2	3	4	5	6	Row total
				Rater PR					
0		0	1	3	1	0	0	0	5
1	Rater EC	0	1	3	4	2	0	0	10
2		0	0	4	6	3	1	0	14
3		0	0	0	2	2	3	0	7
4		0	0	0	0	0	0	1	1
5		0	0	0	0	0	0	0	0
6		0	0	0	0	0	0	0	0
Column Total		0	2	10	13	7	4	1	37

APPENDIX 15

Frequency tables for calculation of intra-rater agreement, EC I VS EC II, 20 randomly selected voice samples.

Table 1 HARSHNESS - Raters EC I vs EC II

Observed cell frequences:

Scalar degree:		0	1	2	3	4	5	6	Row total
					Rater EC I				
0		1	1	0	0	0	0	0	2
1	R a	2	0	0	0	0	0	0	2
2	t e	1	5	1	0	0	0	0	7
3	r	0	0	1	0	0	0	0	1
4	E C	0	0	1	0	0	0	0	1
5	II	0	0	1	0	2	2	0	5
6		0	0	1	0	0	1	0	2
Column Total		4	6	5	0	2	3	0	20

Table 2 WHISPER - Raters EC I vs EC II

Observed cell frequences:

Scalar degree:		0	1	2	3	4	5	6	Row total
					Rater EC I				
0		0	0	0	0	0	0	0	0
1	R a	1	0	0	0	0	0	0	1
2	t e	1	2	0	1	0	0	0	4
3	r	0	0	5	0	0	0	0	5
4	E C	0	0	4	1	1	0	0	6
5	II	0	0	1	0	2	1	0	4
6		0	0	0	0	0	0	0	0
Column Total		2	2	10	2	3	1	0	20

Table 3 CREAK - Raters EC I vs EC II

Observed cell frequences: Row total

Scalar degree:		0	1	2	3	4	5	6	Row total
					Rater EC I				
0		4	0	1	0	0	0	0	5
1	Rater EC II	0	0	0	0	0	0	0	0
2		1	3	0	0	0	0	0	4
3		2	0	0	1	0	0	0	3
4		2	0	2	1	1	0	0	6
5		0	0	0	0	1	0	0	1
6		0	0	0	0	0	1	0	1
Column Total		9	3	3	2	2	1	0	20

Table 4 LARYNGEAL TENSION - Raters EC I vs EC II

Observed cell frequences: Row total

Scalar degree:		0	1	2	3	4	5	6	Row total
					Rater EC I				
0		0	0	0	0	0	0	0	0
1	Rater EC II	1	0	0	0	0	0	0	1
2		0	1	2	0	0	0	0	3
3		3	1	1	1	0	0	0	6
4		0	0	1	1	3	0	0	5
5		0	0	0	0	4	1	0	5
6		0	0	0	0	0	0	0	0
Column Total		4	2	4	2	7	1	0	20

Frequency tables for calculation of intra-rater agreement, EC I VS EC II, 37 irradiated voices.

Table 5 HARSHNESS - Raters EC I vs EC II

Observed cell frequences:

Scalar degree:		0	1	2	3	4	5	6	Row total
		Rater EC I							
0		1	0	0	3	0	0	0	4
1	Rater EC II	1	2	2	0	0	0	0	4
2		0	2	4	5	0	0	0	11
3		0	0	2	2	0	0	0	4
4		0	0	0	3	2	0	0	5
5		0	0	0	1	4	2	0	7
6		0	0	0	0	0	1	0	1
Column Total		2	4	8	14	6	3	0	37

Table 6 WHISPER - Raters EC I vs EC II

Observed cell frequences:

Scalar degree:		0	1	2	3	4	5	6	Row total
		Rater EC I							
0		0	0	0	0	0	0	0	0
1	Rater EC II	0	1	1	0	0	0	0	2
2		0	1	5	0	0	0	0	6
3		0	0	3	5	0	0	0	8
4		0	0	4	5	2	0	0	11
5		0	0	1	0	3	5	0	9
6		0	0	0	0	0	0	1	1
Column Total		0	2	14	10	5	5	1	37

Table 7 CREAK - Raters EC I vs EC II

Observed cell frequences: Scalar degree:		0	1	2	3	4	5	6	Row total
				Rater EC I					
0		6	0	3	0	0	0	0	9
1	Rater EC II	1	0	1	0	0	0	0	2
2		0	0	3	1	0	0	0	4
3		0	1	1	6	3	0	0	11
4		0	0	0	2	3	0	0	5
5		0	0	0	0	3	3	0	6
6		0	0	0	0	0	0	0	0
Column Total		7	1	8	9	9	3	0	37

Table 8 LARYNGEAL TENSION - Raters EC I vs EC II

Observed cell frequences: Scalar degree:		0	1	2	3	4	5	6	Row total
				Rater EC I					
0		0	0	0	0	0	0	0	0
1	Rater EC II	0	1	1	0	0	0	0	2
2		0	1	2	1	0	0	0	4
3		0	0	4	8	0	0	0	12
4		0	0	3	4	5	0	0	12
5		0	0	0	0	2	4	0	6
6		0	0	0	0	0	0	1	1
Column Total		0	2	10	13	7	4	1	37

TABLE 1

Pilot subjects 'revisited'

Subject Number:	SMOKING Before Radiotherapy	SMOKING on reassessment	Age on last assessment	Occupation
2	No	No	68	Retired Civil Servant
4	No	No	71	Retired translator
5	20/day	5/day	88	Retired ?
7	No	No	70	Retired Transport Engineer
8	25/day	No	62	Retired Travel agent
10	15-20/day	No	70	Retired company director
12	60-80/day	10-20/day	48	Unemployed chef
14	30-40/day	No	70	Retired civil servant
16	No	No	61	Catering Manager
20	No	No	65	Retired Architect
21	½oz/day	No	43	Furniture Restorer

Table showing the smoking habits before and after radiotherapy, ages and occupations of the Pilot subjects who were reasssessed for the Main study.

APPENDIX 16

TABLE 2

Main study subjects - 24 MPRx follow up +

Subject Number:	SMOKING Before Radiotherapy	SMOKING on Reassessment	Age on last assessment	Occupation
22	No	No	49	Company director
24	12-15/day	No	65	Printer
25	40/day	No	53	Lecturer
26	'Chain-smoked'	No	61	Retired labourer
27	12-14/day	No	63	Retired builder
28	20-30/day	No	58	Plasterer
29	40/day	15/day	53	Unemployed
30	30-40/day	5-10/day	70	Retired labourer
31	10-15/day	No	51	Ventilation engineer

ctd.

ctd. **TABLE 2**

Subject Number:	SMOKING Before Radiotherapy	SMOKING on Reassessment	Age on last assessment	Occupation
32	1 oz/day	No	66	Accountant
33	30/day	No	68	Retired diplomat
36	No	No	67	Retired accountant

Table showing the smoking habits before and after radiotherapy, ages and professions of the subjects who were referred to the Main study.

APPENDIX 16

TABLE 3

Voice therapy group.

Subject Number:	SMOKING Before Radiotherapy	SMOKING on Reassessment	Age on last assessment	Occupation
14	40/day	No	70	Retired Civil Servant
16	No	No	59	Catering manager
21	½oz/day	No	43	Furniture Restorere
22	No	No	49	Company director
24	12-15/day	No	65	Printer
25	40/day	No	53	Lecturer
28	20-30/day	No	58	Plasterer
30	30-40/day	5-10/day	70	Retired labourer
33	30/day	No	68	Retired diplomat

ctd. TABLE 3

Voice therapy group.

	SMOKING Before Radiotherapy	SMOKING on Reassessment	Age on last assessment	Occupation
Subject number:				
39	20-30/day	5-6/day	69	Retired HGV driver.
40	20/day	No	53	Editor

Table showing the smoking habits before and after radiotherapy, ages and professions of the subjects who received voice therapy in the Main study.

TABLE 4

Main study subjects who did not receive voice therapy.

Subject Number:	SMOKING Before Radiotherapy	SMOKING on reassessment	Age on last assessment	Occupation
23	20-30/day	2-3/day	61	Retired Laboratory Supervisor
26	'Chain-smoker'	No	61	Unemployed Labourer
27	12-14/day	No	72	Retired builder
29	40-50/day	15/day	53	Unemployed seaman
31	10-20/day	No	51	Ventilation Engineer
32	1 oz/day pipe	No	66	Accountant
34	20/day	No	78	Retired
35	No info.	No	67	Retired Valet
36	No	No	67	Retired Accountant
37	20-30/day	12-15/day	55	Businessman
38	40/day	15-20/day	73	Retired Bookbinder

Table showing subjects referred to the Main Study who did not receive voice therapy.