VOCAL REHABILITATION AFTER TOTAL LARYNGECTOMY

ALISON ROSALIND PERRY

This thesis is submitted in partial fulfillment of the CNAA requirements for the degree of Doctor of Philosophy.

Sponsoring Establishment: Leicester Polytechnic.

Collaborating Establishment: Charing Cross Hospital, London.

October 1989.

VOCAL REHABILITATION AFTER TOTAL LARYNGECTOMY

ALISON ROSALIND PERRY

OCTOBER 1989

ACKNOWLEDGEMENTS

I would like to acknowledge the help of the following people:

Professor Pam Grunwell, Head of Studies, Leicester School of Speech Pathology, for her supervision and guidance throughout the preparation of this text, and to
Dr. Chris Code, for his helpful co-supervision and comments on this work.

From Charing Cross Hospital, London:
Mr. A.D. Cheesman, FRCS, Consultant Otolaryngologist,
Dr. J. McIvor, FRCR, Cosultant Radiologist, and
Professor A. Guz, FRCP, Head of the Department of Medicine.

Without their tremendous support, encouragement and practical help, the work could not have been completed.

My colleagues in Riverside Speech therapy service, notably
Roberta Green, District Speech Therapist;
Isobel Bassett and Alison Bridges, Locum speech therapists;
Sylvie Barnett and my long-suffering colleagues in the Department of Speech Therapy at Charing Cross Hospital;

Yvonne Edels for her helpful comments, constant inspiration and enthusiasm for the subject of Laryngectomy, and for her preparedness to discuss problems which arose during the preparation of this thesis.

The Department of Medical Electronics:
Nick Abrahams, Tony Brewer and, especially, David McIntyre, who gave technical assistance with the recordings of the subjects studied and the analysis thereof.

Mr. Michael Lynch from The Department of Electronics, Cambridge University, for assistance in acoustic analysis.

The Department of Medical Illustrations at Charing Cross Hospital for help in the preparation of figures and tables in this thesis, as well as for providing the illustrations to the text.

Ms S.Godbolt and the librarians at Charing Cross Hospital Medical School for their unfailing help and expertise.

Finally,
Simon Furse for his help in word-processing, computing and for his belief in this project and in me.

This work was supported in total by a North West Thames Locally Organised Research Fund and I remain grateful to the Regional Health Authority for their generous grant which enabled me to complete this work.

NOTE

CONVENTIONS USED IN THIS THESIS

1. I have observed the convention of referring to all laryngectomees as "he" where necessary in view of the ratio of male : female patients as being approximately 5:1, rather than continuously using the awkward format of "he/she".

2. Due to the difference in American spelling and English spelling of esophagus/oesophagus respectively, I have observed the original spelling where quotations are given (viz - "esophagus" if an American text is quoted; "oesophagus" if an English/European quote), but otherwise consistently used the English spelling.

ABSTRACT

VOCAL REHABILITATION AFTER TOTAL LARYNGECTOMY

This thesis considers the options for speech rehabilitation for patients who have undergone total laryngectomy - artificial larynges, oesophageal speech and/or surgical speech restoration - and examines the premise that the main reason for failure to develop speech following total laryngectomy is due to anatomical/physiological considerations.

A data collection protocol has been devised, in which simultaneous acoustic and aerodynamic measurements are recorded on all laryngectomised subjects during a videofluoroscopic examination.
Sixty-eight laryngectomees were examined in total, 13 female and 55 male, over a two-year period at a monthly-held data collection clinic.

The results of the study are consistent with the experimental hypothesis. Analysis of acoustic and aerodynamic measurements support the hypothesis of a continuum of tonicity in the reconstructed pharynx, from hypotonicity (where the pharyngo-oesophageal segment is too flaccid to vibrate and produce sound) through tonicity (where good speech results) to hypertonicity and spasm (depending on the degree of excess tone in the reconstructed pharyngo-oesophageal segment).

The objective assessment developed in this study accounts for those subjects who fail to achieve oesophageal speech, rather than simply charting oesophageal speech progress as earlier assessments have done. This assessment helps to determine the path of subsequent clinical management for the laryngectomee who wishes to acquire speech, whether this is oesophageal speech or speech after undergoing a further surgical/ prosthetic procedure.

This thesis suggests the development of a proactive and principled model for laryngectomy treatment and indicates future research and directions to pursue in order to improve the prospects for speech rehabilitation in laryngectomised subjects.

CONTENTS

CONTENTS

CONTENTS

CHAPTER 1 : TOTAL LARYNGECTOMY : CAUSES AND CONSEQUENCES.

The early diagnosed squamous cell carcinoma which is restricted to the vocal fold, causing symptoms of hoarseness when the tumour is small, has a good prognosis for cure with radiotherapy. These tumours are staged as T1 or T2 carcinomas (I.U.C.C. classification, 1987) and have 85 - 95% chance of cure with radiotherapy (DXT).

Unfortunately, not all patients are diagnosed at this stage. Graham (1983) wrote...

> "If unlucky, he (the patient) will try to ignore his hoarseness in the hope that it will go away; or visit his doctor, be prescribed a cough medicine, go home and take it regularly for four months and be admitted to hospital as an emergency on the point of suffocation with his larynx full of a huge, stinking tumour."(p.1)

For those patients with extended tumours staged as T3 or T4 carcinomas, surgery, in the form of a total laryngectomy, is the best therapy. The indication for additional post- operative DXT depends upon the post operative findings and the histological examination of the removed specimen. A full discussion of "Radiotherapy or Surgery" is given by Berry (1983).

1.1 Total Laryngectomy

Total laryngectomy entails complete removal of the larynx, including the hyoid bone superiorly and the upper tracheal rings inferiorly. The remnant of the severed trachea is brought out at the front of the neck and sutured to the skin above the sternum. In this way, an open airway or tracheostoma (stoma) is created. The removal of the larynx results in the anterior dissection of the pharyngeal musculature and mucosa. The resulting defect in the anterior pharyngeal wall is closed by approximation of its edges. First the mucosa is closed and then the closed defect is reinforced by layers of pharyngeal and remaining strap muscles. Singer (1988) stated that ..

> "the actual preservation of the cricopharyngeus muscle and pharyngeal constrictors is not given much attention during laryngectomy, except the remnants of these muscles are used as a "third layer" for closure of the pharynx after laryngectomy. These muscles are not thought of as a reconstructed sphincter, but only as a buttress to the pharyngeal closure."(p.119).

Before the pharynx is closed, a nasogastric tube is passed through the nose, down the oesophagus to the stomach for post -operative nutrition.

The pharyngeal repair is formed meticulously to prevent fistula formation (Applebaum and Levine, 1977). The surgical steps in total laryngectomy are fully described by Cheesman (1983).

Simpson, Smith and Gordon (1972) investigated the influence of muscle reconstruction on the mechanism of oesophageal speech production. Four different types of surgical closure were used and correlated with subsequent radiological appearance and the success or failure of oesophageal speech production. There was a significantly higher number of good speakers when the surgical technique was designed to reconstruct a fully muscle-controlled pharynx which

> "maintained the through-link between the supra-hyoid, tongue and jaw muscles to reproduce the basic mechanism required for the new voice-production complex."

There was thus evidence to suggest that good muscle closure was important for the subsequent development of oesophageal speech.

After total laryngectomy, the food passage is completely separated from the airway.

1.2 Consequences of Laryngectomy.

Following laryngectomy, the chances of being cured from malignant disease are relatively high. Depending on the site of malignancy, its extent and whether or not local metastases have occurred, the mean five year survival rates vary between 50 and 85% (Mahieu, 1988). Consequently, most laryngectomees live for many years with the handicaps associated with the laryngectomised state.

Clearly, the goal of this type of oncological surgery must be the total eradication of the malignancy. Less radical surgical procedures which are performed to minimise the consequences of laryngectomy can compromise the oncological principle of radical tumour excision. Therefore, the indications for less radical surgical procedures must be thoroughly scrutinised (Mahieu, op cit.).

1.2.1 Loss of Laryngeal Functions.

The most obvious handicap of total laryngectomy is loss of vocal function. Although warned of this, the laryngectomee will often be both frustrated and distressed. The absence of sounds which normally accompany laughter and crying can be distressing and puzzling for the laryngectomee.

Glottic closure, such as occurs in coughing, abdominal press and lifting heavy weights is no longer possible. Also, because of the loss of Valsalva manoeuvre, which "fixes" the glottis during exertion, sexual problems may arise for male laryngectomees who experience consequent problems in sustaining a penile erection.

Although severe problems resulting from this lost ability of glottal closure are rarely observed (Jakobi and Muller, 1958; Diedrich and Youngstrom, 1966), they can be distressing and result in further emotional/psychological difficulties. Eventually, compensating mechanisms enable sufficient stabilised intrathoracic pressure to occur, even without glottal closure but the laryngectomee faces many emotional crises before full adjustment is reached (Sanchez-Salazar et al, 1972).

1.2.2 Loss of Olfactory Functions

Loss of nasal airflow affects the sense of smell and thus the sense of taste (Darvill, 1983). Methods have been described (Damste 1979) to re-establish nasal airflow to improve these senses but it can be distressing and puzzling to the laryngectomee. The ability to blow the nose is also impaired (Wirth, 1972; Schultz-Coulon 1984) and laryngectomees have to learn to consciously wipe the mucus from the nares.

1.2.3 Airway changes.

Total laryngectomy results in "the anatomy and physiology of the airway (being) dramatically altered" (Cheesman, op cit). As the upper respiratory tract, with its humidification function, is completely bypassed, the air or oxygen which the laryngectomee breathes must be humidified for him.

In the immediate post-operative stages, humidified air is delivered directly to the laryngectomy stoma by a tube. In time, the lower respiratory tract adapts and intermittent humidification by administering a fine spray of water directly into the trachea can be substituted.

Pitkin (1953) found coughing and mucus discharge to be one of the chief discomforts of the laryngectomee. Most laryngectomees learn to clear mucus from the lungs by rapid expiration followed by wiping the resultant secretions away from the stoma with a tissue, but the technique takes several weeks to learn.

In the initial period, the mucus is suctioned from the tracheostomy by applying a sterile suction tube. This may be necessary every few minutes for the first 24 - 48 hours (Cheesman, op cit). Eventually, the laryngectomee learns to wear a foam tracheostomy cover or bib, moistened to humidify the air which, if inhaled neither moistened nor warmed, may cause irritation of the tracheal mucosa. Shaving, bathing and showering habits must be consciously adapted (Damste, op cit.).

1.2.4 Swallowing.

In the immediate post-operative period, nutrition will be given as liquid down the nasogastric (n-g) tube. The laryngectomee is not allowed to swallow anything by mouth and even his saliva is initially suctioned away. This ensures no strain is placed on the surgical repair of the pharynx too early. Before removing the n-g tube, the integrity of the pharyngeal repair is tested by allowing the patient to drink a coloured solution. If no signs of fistula formation are found, the feeding (n-g) tube is removed and the patient is ready to cautiously build up a swallowing programme from liquids to semi - solids etc. This normally occurs on the eighth to tenth post - operative day.
Although disturbances of swallowing have been reported following laryngectomy, this seems to be a minor importance in the laryngectomee's nutrition. Most laryngectomees experience no severe swallowing problems (Sandberg, 1970; Bronwer, Snow and van Dam 1979) but pharyngeal stenosis, the development of pharyngeal pouches and swallowing disco-ordination can occur. Most laryngectomees seem to adapt well to their changed pharyngo-oesophageal anatomy and swallow functionally, if not in exactly the same way as pre-operatively.

Any late development of dysphagia is a potentially serious complaint and will need investigation e.g. using videofluoroscopy. Fortunately, a tight pharyngeal repair generally improves spontaneously and it is only occasionally necessary to dilate the pharynx by bouginage or to surgically reconstruct it.

1.2.5 Psychological Consequences.

The psychological consequences of this surgery are often underestimated, but a full discussion of this is beyond the scope of this thesis. Therefore, only the most important psychosocial consequences are mentioned here.

The laryngectomee's rehabilitation will be influenced by the anxiety associated with the diagnosis and treatment of the malignancy (Schall, 1938, Darvill, op cit.) For a long period of time, the patient and his family and friends will be troubled by whether or not the laryngectomee is really cured of his carcinoma.

The loss of the larynx affects a patient's sense of personal worth, his family and social relationships and his employment (Locke, 1966).

Wochnick (1972) in his study of the psychological state of 52 male laryngectomees found the following rank order of complaints post-operatively: disturbances of contact (rapport) and feelings of being handicapped; abnormal workings of the vegetative nervous system or of physical functions; deterioration of abilities and skills; disturbances of self-confidence; general irritability and restlessness; a basic mood of dissatisfaction; anxiety and joylessness; strong hypersensitivity.

A successful voice and speech rehabilitation is found to facilitate psychological adjustment (Wochnick, op cit., Dhillon et al, 1982) but it may not be enough to ensure total rehabilitation of the laryngectomee (Diedrich and Youngstrom, op cit; Watts, 1975; Pfrang 1986; Pfrang et al 1986).

It is difficult to ascertain whether psychological/social problems militate against successful voice rehabilitation or whether the lack of successful voice results in psychological /social problems.

Duguay (1979) listed psychological/social problems as being factors militating against learning to communicate post-operatively and yet Cheesman et al (1986), in their study of 50 "failed" oesophageal speakers, using videofluoroscopy, found anatomical/physical barriers to speech acquisition in all the subjects examined.

1.3 Speech Production / Speech Rehabilitation Procedures after Total Laryngectomy.

The most disabling consequence of laryngectomy is generally considered to be the loss of vocal function. Since 1873 when Billroth, in Vienna, achieved the first successful laryngectomy for cancer, methods have been sought to restore the speech facility of laryngectomees.

1.3.1 Normal Laryngeal Speech.

Basic speech pathology teaches us that there are four processes involved in the production of human speech, namely; initiation, vibration, resonance and articulation. (Edels, 1983).

For normal (laryngeal) speech, the initiator for speech production is the steady stream of air from the lungs on expiration. The vibrator is the larynx (i.e. vocal cords) which produce the basic vocal tone. The sound generation depends upon the interaction of aerodynamic force during expiration and the myoelastic forces of the vocal cords. The flexible mucosal lining of the vocal folds is brought into motion by expired air, thus creating a mucosal wave. This wave opens and closes the glottis periodically causing a periodic interruption of the expired column of air, which in turn results in sound. The resonators are the cavities above the larynx; pharynx, oral and nasal cavities, which amplify the sound.

The articulators - i.e. tongue, teeth, lip, hard and soft palate - act to shape and modify the basic amplified sound into recognisable phonemes (See Table 1).

1.3.2 Changed Anatomy After Laryngectomy.

During laryngectomy, it is the larynx - i.e. the vibrator - which is removed. This means that the laryngectomee is essentially aphonic. The primary goal in speech restoration is to provide a substitute for the original vibrator (vocal cords) by alternative means. Articulatory characteristics such as dialect are not impaired. Therefore, "speech restoration" is perhaps not correct in the strict sense; the laryngectomee has to develop a new energy source and sound source (i.e. initiator and vibrator) in order to acquire a substitute voice. The new sound source must be connected to the vocal tract so that it may be used for articulation and thus shaped into speech. Available substitutes are: artificial larynges; oesophageal speech; and/or surgical speech restoration.

1.3.3 Artificial Larynges

All speech aids are designed to produce sound that can be transmitted into the vocal tract. The vibratory source for sound may be a plastic reed, a flexible rubber membrane or a rigid membrane. These vibratory sources are activated either electronically or by pulmonary air pressure. There are two basic type of device; pneumolarynges (powered by pulmonary air) and electrolarynges (those powered by batteries).

Electrolarynges are further sub-divided into two categories (mouth type and neck type) to indicate the anatomical area at which the sound enters the vocal tract. The number and type of available artificial larynges is too numerous to discuss fully in this text, but these are well described by Gardener and Harris (1961), Salmon and Golstein (1978) and Salmon (1983). The principles of the different types of artificial larynges are illustrated by the following examples:

1.3.3.1 Pneumatic Artificial Larynx Devices.

The Tokyo artificial larynx consists of a steel cup which fits externally over the tracheostoma, a steel tube leading to and from a cylindrical unit that houses a stretched rubber membrane, and a plastic or rubber mouth tube (Figs.1, 2).

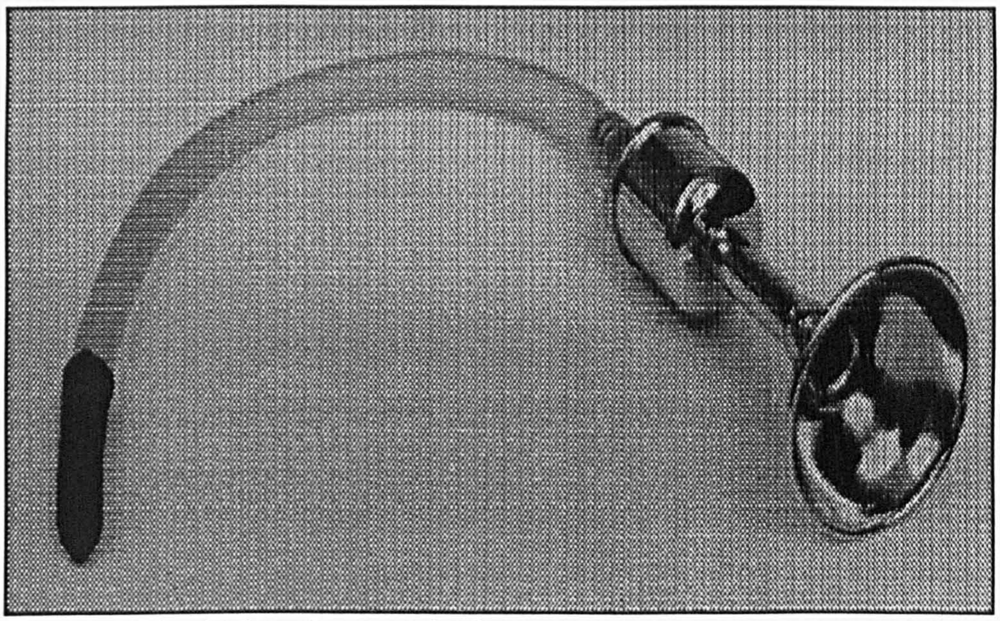


Figure 1.The Tokyo Pneumatic Larynx.

Expired lung air is used to vibrate the rubber membrane and the sound is directed into the oral cavity via the mouth tube. Pitch can be adjusted by altering the tension in the rubber membrane. An additional modification may be to drill a "breathing hole" in the steel tube (where the diaphragm is sited) so that the laryngectomee can simply use the drilled hole open for inspired air, then cover the hole with a finger to divert expired air into the mouth via the rubber diaphragm. In this way, he does not have to remove the tracheostoma cup for breathing (Mueller and Kupperman, 1973).

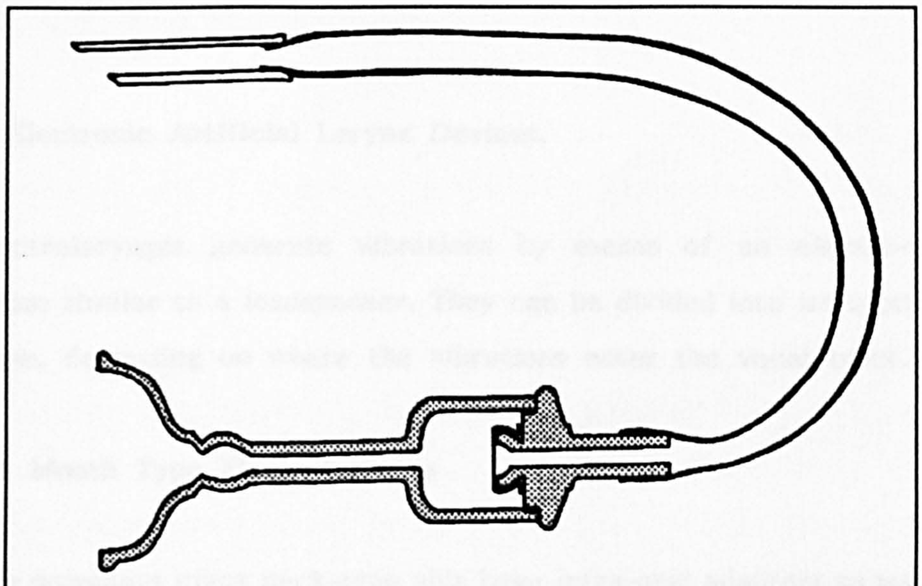


Figure 2.Internal dimensions of the Tokyo Pneumatic Larynx.

There are certain limitations to be recognised when considering a pneumatic device for a laryngectomee (Salmon, op cit):

1. Laryngectomees with respiratory disorders (e.g. emphysema) will have difficulty generating sufficient lung pressure to drive the diaphragm.

2. Coughing may be triggered or intensified by pressure from the stomal cup against the sensitive tissue around the stoma.

3. An adequate airtight tracheal coupling may be a problem if the laryngectomee has irregular neck contours.

4. Sound transmission via mouth tube may become frequently blocked with saliva - this also occurs with electronic mouth-type devices.

1.3.3.2 Electronic Artificial Larynx Devices.

The electrolarynges generate vibrations by means of an electro-vibratory mechanism similar to a loudspeaker. They can be divided into intra-oral and/or neck type, depending on where the vibrations enter the vocal tract.

1.3.3.2.1 Mouth Type Electrolarynges

Although nowadays many neck-type aids have intra-oral adaptors to make them more versatile, the Cooper-Rand is probably the most widely used intra-oral aid. It consists of a pulse generator connected by a connecting cord to a hand-held pulse generator and a plastic mouth tube (Fig. 3). Pitch and loudness are adjusted by controls at the top of the pulse generator. A "no hands" version of this aid is available for patients who are unable to activate an aid with their hands but who need an electrolarynx.

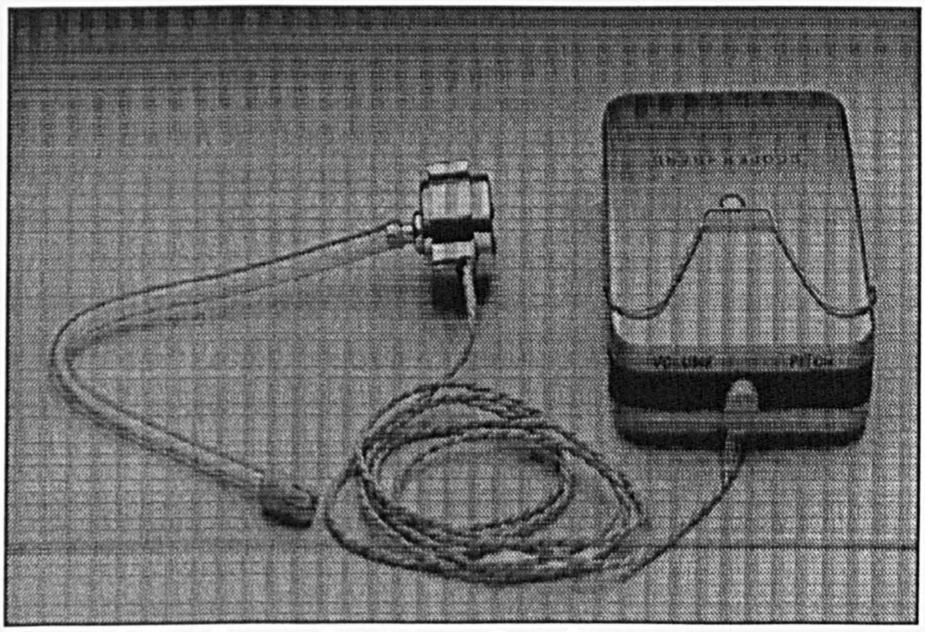


Figure 3.The Cooper-Rand Intra-Oral Electrolarynx.

1.3.3.2.2 Neck Type Electrolarynges

There are many varieties of neck-type vibrators which are available, but the most commonly used in the U.K. is the Siemens Servox. This instrument is manufactured in West Germany and is made of both plastic and metal (Fig.4). Vibration of the diaphragm is produced by movement of a piston driven by a battery-powered motor. The battery is rechargeable and an extra battery plus battery charger are included in the price of the aid. Since 1985, there has also been an intra-oral adaptor provided, which consists of a cap which fits over the vibratory head of the Servox and a tube leading from this into the oral cavity. This will convert the neck-type aid into an intra-oral one for early post-operative use. In this way, the laryngectomee may use a Servox intra-orally in the immediate days after surgery and then convert it into a neck-type aid, once the sutures are removed and the laryngectomees' neck has healed.

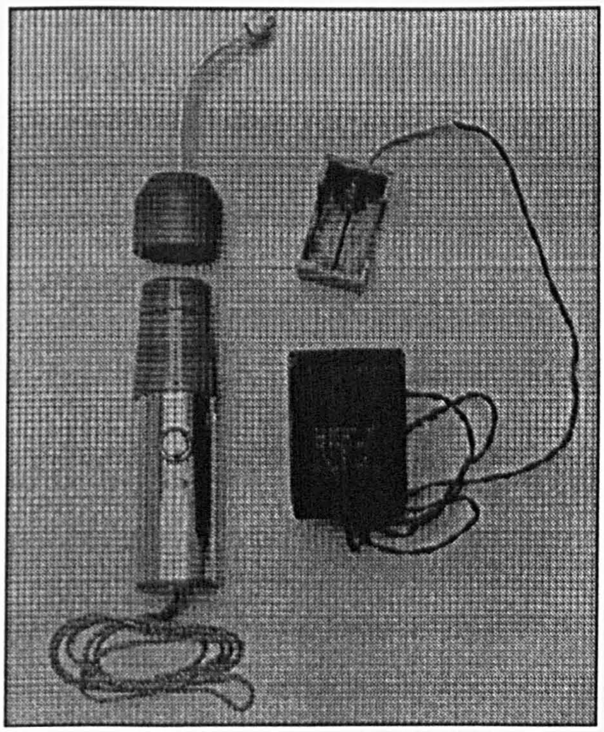


Figure 4. The Servox Neck-Type Electrolarynx.

The limitations of the neck-type aid are as follows:

1. The neck tissue may be fibrotic or oedematous which will give difficulty in sound transmission.

2. The neck tissue is necrotic following DXT so that there is little or no feeling in the neck. This will result in difficulty in accurate and consistent placement of the head of the instrument against the best resonator in the neck.

Many laryngectomees dislike artificial larynges as they find the tone monotonous as well as difficulty in the more subtle use of aids e.g. whispering or quiet asides become impossible. Additionally, the weight of an instrument (average = 9 oz., Salmon, op cit) is heavy to hold; and any hearing loss can limit the laryngectomee's self-monitoring skills, which results in an overall reduction in intelligibility.

1.3.4 Oesophageal Speech

As can be seen from Table 1, the initiator in oesophageal speech is the moving column of air from the oesophagus. For success to ensue with this means of speech production, the laryngectomee has to be able to trap air in the oral cavity, inject the air into the oesophagus and briefly retain it there, then regurgitate this air via the reconstructed pharyngo-oesophageal (P-E) segment to produce sound. It is the P-E segment which becomes the new vibratory source for voice. The laryngectomee has to exert voluntary control of this method of sound production in order to use it as a substitute voice. Thus, some laryngectomees are able to produce good, intelligible speech by using the acoustic properties of the vocal tract to modify this oesophageal sound.

PHYSICAL REQUIREMENTS	LARYNGEAL SPEECH	OESOPHAGEAL SPEECH	T-E (FISTULA) SPEECH
Initiator	Moving column of air from lungs.	Moving column of air from oesophagus.	Moving column of air from lungs.
Vibrator	Vocal cords.	Pharyngo-oesophageal (P-E) segment.	Pharyngo-oesophageal (P-E) segment.
Resonator	Vocal tract, i.e. pharynx, nose & mouth.	Vocal tract, i.e. pharynx, nose & mouth.	Vocal tract, i.e. pharynx nose & mouth.
Articulators	Tongue, teeth, lips and soft palate.	Tongue, teeth, lips and soft palate.	Tongue, teeth, lips and soft palate.

Adapted from Edels, Y (1983)

Table 1.Comparison of Laryngeal, Oesophageal & Tracheo-oesophageal Speech.

1.3.4.1 Pseudoglottis: The P-E Segment - Structure and Innervation.

This area of investigation has received much attention by researchers and there has been a great deal of controversy concerning the nature, location and physiology of the structures which produce oesophageal sound. The actual source is considered to be the so-called pseudoglottis, a structure situated at the junction of the pharyngo-oesophagus or P-E segment, at the same cervical vertebrae level where the larynx used to be sited (Damste, 1958; Moolenaar-Bijl, 1952). Most research supports the finding that the P-E segment is located between cervical vertebrae levels four and six, and Diedrich and Youngstrom assert that the P-E segment is located between the level of the fourth and seventh cervical vertebrae 90% of the time.

Recent electroglottographic studies (Lecluse, Tiwari and Snow, 1981; Schutte, 1980; Motta et al 1985) support the fact of the P-E segment being the vibratory source for oesophageal speech. This is also supported by radiographic studies (Benzen, 1976; Novak et al, 1982), stroboscopic studies of the P-E segment in oesophageal speech production (Hirose, 1983; Brandenberg, Bless and Salinsky, 1985) and ultrasound pseudoglottis imaging studies (Bockler et al, 1988).

Electromyographic studies (Shipp, 1970) to examine muscular control of the P-E segment during oesophageal speech production have been inconclusive (see section 2.4.). Shipp concluded that the pattern of muscular activity in the P-E segment was idiosyncratic and varied from subject to subject, although there were simultaneous bursts of EMG activity from both the inferior constrictor and the cricopharyngeus muscles. This was interpreted as being a passive stretch reaction to the bolus of air being injected. During the onset of phonation, there was increased activity from the cricopharyngeus and often some parallel activity in the inferior constrictor muscle that diminished rapidly upon completion of phonation.

The location of the P-E segment may be a factor influencing the quality of oesophageal speech produced. A highly situated P-E segment will usually result in a high-pitched voice, often mistaken for pharyngeal speech (Shanks, 1985, personal communication).

The P-E segment may be anything from 0.8 in. to 1.2 in. long and has been shown to contain three muscle groups: at the top, fibres from the inferior constrictor; in the middle, the lowest fibres of the inferior constrictor form a band called the cricopharyngeus muscle; and this blends inferiorly with a circular band of fibres from the mouth of the oesophagus (Edels, op cit., Blom, 1979).

Earlier researchers described the P-E segment as being primarily formed by the cricopharyngeal muscle (Negus, 1938; Robe et al 1956; Damste 1958, op cit; van den Berg et al, 1959; Seeman 1961; Levin, 1962) although this was refuted by Micheli-Pellegrini (1957) who suggested that the mouth of the oesophagus was the level at which sound was produced. It is now accepted by most clinically-working researchers that...

> "the perfectly timed opening and closing of the cricopharyngeus muscle AND ITS ADJACENT FIBRES.. affect the quality and usefulness of the esophageal sound" (Zwitman, 1979).

It was Diedrich and Youngstrom who referred to this area as the site of the pseudo-glottis and used the term pharyngo-oesophageal segment (abbreviated to P-E segment) to label it. Various other terminology includes pharyngo-oesophageal junction and pharyngo-oesophageal sphincter. The laryngectomee aims to gain voluntary control of the P-E segment to allow the intake and release of air for the production of sound.

Without the presence of a P-E segment, there is difficulty in producing speech. If a laryngectomee has a hypotonic P-E segment then the anterior and posterior walls of the oesophageal mucosa do not abut to produce sound. The extreme example of this is the patient who has had a post-cricoid carcinoma and thus lost the larynx, oesophagus and pharynx. This may result in a gastric pull-up to close the defect in the laryngo-pharynx and thus no vibrating segment for speech as no sphincter exists. The laryngectomee then has to produce or create a P-E segment by the use of digital pressure on the outside of the neck. This is fully discussed by Logemann (1983) and Bleach (in press.)

The innervation of the P-E segment was described by Inglefinger (op cit) as reported by Diedrich (1968):

"On anatomic grounds, the sphincter appears innervated by an overlapping supply of elements derived from the nucleus ambiguus and dorsal vagal nucleus, and carried by the X and to a lesser degree, from the IX and XI nerves. The resting tonicity of the sphincter and its opening are under influences carried by these nerves...Because of its high resting tone, the obvious response of the sphincter is to the inhibitory component of the peristaltic impulse. Although the sphincter opens under conditions other than swallowing the response mechanisms are unknown...........

It is likely, however that voluntary, unlike involuntary, relaxation of the upper esophageal sphincter is not easily elicitable by itself, but is one step in an arrangement of multiple performances." (p310).

Decroix and co-workers (1958) in their extensive study of the anatomical and physiological bases of oesophageal speech, provide some tentative evidence that there may be proprioceptive fibres present in the muscle tracts of the P-E segment. Proprioceptive innervation, demonstrated by the presence of neuromuscular spindles in the oesophageal wall, can explain the contraction of the lower part of the sphincter and permits the control and the precise degree of voluntary contraction needed in oesophageal speech. In the dissection of 50 pharyngo-larynges, Decroix et al demonstrated the consistency of the branches of the recurrent nerve intended for the "mouth of the oesophagus." They stressed the importance of preserving as much of the innervation of this area as possible.

1.3.4.2 The Oesophagus

At either end the oesophagus is protected by a sphincter. Superiorly, the P-E segment separates the hypopharynx from the oesophagus. Inferiorly, the cardiac sphincter separates the oesophagus from the stomach. Salmon (1979) states that the diameter within the upper oesophageal sphincter is a little over 0.5 ins. while the diameter prior to its union with the stomach is less than 1.0 ins. The P-E segment is anything from 0.8 to 1.2 ins. long while the cardiac sphincter is 2 ins. long (Edels, op cit.). The adult oesophagus, regardless of age or sex of the patient, is about 9 ins. long.

Unlike the trachea, which retains a patency in both normal and laryngectomised subjects, the oesophagus is collapsed at rest. It normally only opens under pressure from above to allow the passage of food in a peristaltic wave. The oesophagus at rest has a negative pressure registering between -4 and -7 mmHg below atmospheric pressure. This may decrease from -10 to -20 mmHg during pulmonary inspiration. Whereas, prior to surgery, the oesophagus played no part in voice, for the laryngectomee it must act as a new air reservoir, expulsion of its air acting as the initiator for vibration of the P-E segment. In order to trap and retain air within the oesophagus, the resistance of the P-E segment must be temporarily overcome.

Also, there must be sufficient pressure at both the P-E segment and cardiac sphincter to prevent air escaping until it is to be expelled for production of oesophageal speech. Wolfe et al (1971) showed hiatus hernia and/or gastro - oesophageal reflux resulting in loss of oesophageal reservoir through leakage of air into the stomach could account for failure to acquire oesophageal speech in a percentage of laryngectomees.

1.3.4.3 Air Intake

How does the P-E segment function? Damste and others hypothesised that the tonicity of the P-E segment may be overcome in two ways: either by voluntary relaxation of the P-E segment (inhalation); or by pressure, exerted from above, overcoming the resistance of the P-E segment and consequently forcing it open (injection).

Swallowing as an effective technique for air charging does not exist (Edels, op cit).

METHODS OF AIR INTAKE.

At rest, air in the laryngectomee's nasal, oral and pharyngeal cavities continues to circulate at atmospheric (i.e. positive) pressure. The P-E segment is closed and registers positive pressure, while the oesophagus is closed down and registers negative pressure.

Before sound can be generated, air must pass the P-E segment and enter the oesophagus, which will then register a positive pressure relative to that in the oral/pharyngeal cavities. This may be done in two ways:

INHALATION

This is also called aspiration (Seeman 1922) and insufflation.

As air intake is attempted, a patent (i.e. open) airway is established either between the lips and P-E segment, the nose and P-E segment, or the lips, nose and P-E segment. The tongue lies on the floor of mouth not to occlude the airway during the air intake phase of inhalation. The air reservoir in the oesophagus is filled by means of a quick respiratory inhalation. The negative intra-thoracic pressure, associated with inspiration, is transmitted to the oesophagus. This results in an even lower pressure (e.g. -15 mmHg) than that which normally exists (-4 to -7 mmHg). If the P-E segment relaxes and opens, air will flow from the mouth and nose via the pharynx into the oesophagus. As the pressure differences are equalised, the P-E segment closes, leaving air contained within the inflated oesophagus ready for voice production.

INJECTION

This was extensively described for the first time by Moolenaar - Bijl in 1952 (op cit). The air is forced past the P-E segment by increasing the pressure of air within the oral cavity. This can be achieved by shutting the escape routes for air (the oral and nasal ports) and then reducing the size of the air chamber (oropharyngeal cavity). When the enclosed air is sufficiently pressurised to overcome the resistance of the P-E segment, air will enter the oesophagus. Relaxation of the P-E segment assists this process. There are basically two methods of injection:

(a) Standard injection (or tongue-pump injection, glossal press or glossopharyngeal press) where the decrease in size of the oropharyngeal cavity is achieved by a piston or pumping type action of the tongue. Voice production using standard injection implies two distinct phases; active air intake phase, and air expulsion phase. There are two discrete activities occurring in temporal sequence. For both the inhaler and the injector, the oesophagus is inflated independent of any speech phoneme. There is a pre-speech manoeuvre resulting in the oesophagus being inflated prior to voice produced. The inhaler/injector will usually find it easier to produce voiced consonants and vowels to begin.

(b) Consonant injection. The action of the tongue injecting air into the oesophagus may also be associated with articulation of a voiceless consonant, particularly a plosive, fricative or sibilant. In this case, the method is called consonant injection (plosive injection) and may be differentiated from standard injection by the fact that the air intake and voice production phases either occur simultaneously (as with the use of /s/) or in inseparable association with a voiceless sound. Studies have shown (Edels, op cit) that the latency period (i.e. time taken from beginning of the movements to inject air to voice production) is less for consonant injectors than for standard injectors.

1.3.4.4 Air Expulsion following Air Intake.

Van den Berg, Moolenaar-Bijl and Damste (1958) as quoted by Mahieu (1988) gave a detailed description of the complex mechanism of air expulsion. Five mechanisms may come into action:

a. A voluntary active contraction of the striated muscles of the upper part of the oesophagus.

b. A direct compression of the region of the cardia by the diaphragm.

c. A passive increase of oesophageal pressure by the decrease of the volume of the lungs during expiration. At the end of a maximal expiration, this increase can attain very high values, as the tissue itself is then compressed.

d. A passive behaviour of the non-muscular components of the wall of the oesophagus.

e. A passive increase of the oesophageal pressure by expiratory movements of the thorax. This effect becomes extremely important with a filled oesophagus, during loud speech or with strong effort to speak, when a local distention of the oesophagus into the trachea gives rise to a high airflow resistance.

As described by Van den Berg, the two prime factors maintaining the pressure of air in the oesophagus, required for oesophageal speech production, are (d) and (e) above.

(d). The elastic properties of the walls of the oesophagus is an important factor when using injection as a method of air charging, where air is forced through the P-E segment into the upper oesophagus. The resulting pressure rise in the upper oesophagus, only present briefly, is higher for the injection than for the inhalation method of air charging. According to Damste (1958 op cit), this mechanism alone is enough to guarantee expulsion of air if the plosive injection method of air-charging is used at normal conversational loudness level.

(e). This mechanism is used with the inhalation method of air charging or when, after injection is used as an air charging method, higher sub-pseudoglottic pressures are required, such as for loud or high pitched phonation. This method is often employed by beginners who still have poor control over the P-E segment and who generally let a long time lapse between air intake and phonation onset. This mechanism also plays the major role in air expulsion if undesirable swallowing as a method of air charging has occurred.

Most authors (Diedrich and Youngstrom, op cit; Isshiki, 1969; Snidecor, 1969; Salmon, 1979, op cit; and Edels, op cit) advocate a combination of injection and inhalation as methods for air intake. No good oesophageal speaker exclusively uses only one method; an integration of both methods seems to automatically ensue, regardless of which mechanism is initially taught.

1.3.5 Failure to Develop Oesophageal Speech.

Despite some laryngectomees managing to develop speech with little effort and with no help from speech therapy (Priest, 1984), a large proportion fail to attain satisfactory proficiency. A review of the literature shows the number of laryngectomees who fail to develop oesophageal speech ranges from 14% (Hunt, 1964) to 76% (Schaefer and Johns 1982). There is tremendous variability in defining "failed" oesophageal speech and this fact, together with the possible reasons for failure, are fully discussed in the literature review (Section 2.2.) and will not be repeated here. Suffice it to state that the average estimates place the number of oesophageal speech failures at approximately one-third (Putney 1958; Lauder, 1969; Watts, op cit; Salmon, 1979, op cit; Kawasaki et al, 1983; Natvig, 1983). These results were derived from retrospective studies. Using prospective data, Gates and co-workers (1982) found that 74% of their examined laryngectomees failed to acquire oesophageal speech. When those who did not attend for therapy were excluded, the failure rate was still a relatively high 66%.

Even when successful, oesophageal speech has shortcomings as a means of communication. Due to the small oesophageal air reservoir (80 ccs) used as an initiator for speech, the laryngectomee finds it difficult to sustain length of phonation and loudness. Even superior oesophageal speakers only manage to sustain a vowel sound for 5 or 6 seconds duration (Berlin, 1963a) and needs to constantly re-inflate the oesophagus to air-charge prior to phonation.

Normal alaryngeal loudness is about 65-60 dB at a comfortable level, with a range of 0-95 dB (Davis, 1979). Oesophageal speech, by contrast, has an average level of 40-50 dB, so it is not surprising that many laryngectomees complain of not being heard against background noise. The range of loudness is normally only about 20 dB, the laryngectomee finding it impossible to either whisper or to shout loudly.

Rate of oesophageal speech is reduced when compared to laryngeal. Superior oesophageal speakers tend to be able to achieve 110 words per minute compared with a mean of 165 words per minute for non-laryngectomised speakers. Thus the best oesophageal speakers are still only able to achieve the lowest range of normal speakers (Shipp 1967, Hoops and Noll 1969).

Acoustically, voice quality depends in part on periodicity of phonation. A consistent vocal tone gives good voice quality whereas the less consistent the tone, the poorer the vocal quality. Due to its aperiodicity, the quality of voice in oesophageal speech is reduced when compared to laryngeal voice.

Finally, pitch (frequency) of oesophageal speech is considerably lower than that of laryngeal speech. The average voice fundamental frequency of male oesophageal speakers is judged as being approximately 65 Hz or half that of a normal male laryngeal speaker. Female range is 80-87 Hz. These studies are fully discussed in the next chapter (2.3.1).

Thus, oesophageal speech, even in superior speakers, has shortcomings as a means of communication when compared to laryngeal speakers. It is quieter, lower in pitch, slower in rate and of less duration than normal laryngeal speech. Many of these shortcomings are due to the changed initiator for speech production. Even when fully extended, the oesophagus will only hold 80 ccs of air, which is considerably reduced when compared with the 2 litres of lung air that is used as an initiator for laryngeal speech. It was in order to improve this situation, by re-connecting the lungs once more as an initiator for speech, that much of the efforts in surgical speech restoration after laryngectomy have been directed. It was felt that, once this problem had been addressed and solved, post-laryngectomy speech would be improved in all its parameters. The third option for communication following laryngectomy is surgical restoration of speech.

1.3.6 Surgical Speech Restoration.

Surgical manoeuvres of various kinds have been used within the severe limits imposed by adequate excision of the carcinoma and by the often considerable reluctance of irradiated tissues to heal, to facilitate the acquisition of laryngeal speech. Edwards (1983) indicates that the essential objective of a wide variety of surgical speech restoration (abbreviated to SSR in this text) is the reconstruction, after removal of the larynx, of the physiological situation where expired air from the lungs generates and controls the voice and facilitates its articulation into speech. "Fistula voice" or "fistula speech" describes this process. Edwards stated...

> "The potential of the fistula methods for restoring to the laryngectomee immediately alaryngeal speech of surpassing power, fluency, flexibility and natural quality, compared with oesophageal speech and other "artificial" methods, seems immense."

Various ingenious methods of SSR have been attempted by surgeons since the time when Billroth (1873) in Vienna achieved the first successful laryngectomy for cancer. For reasons of surgical safety, the anterior pharyngeal wall was left unrepaired and the patient had a large external opening into the pharynx above the tracheostoma; in effect an external vocal fistula. Billroth's assistant, Gussenbauer, devised a silver cannula or prosthesis to fit into the trachea below and the pharynx above. A metallic reed, activated by expired air, produced the vibrator for speech and an external hinged lid on the pharyngeal tube was designed to prevent food and liquid leaking into the prosthesis. The patient was reported to have spoken intelligibly with this device (Edwards, op cit).

Various improved prostheses were produced over the next decade, but this early era of SSR ended when surgeons succeeded in safely closing the pharyngeal defect. Closure of the pharynx allowed development of oesophageal speech and led to development of artificial larynges (see 1.2.3.).

Little was heard of SSR in the first half of the twentieth century, apart from a few scattered efforts. However, in 1935 Guttman in Chicago created a crude direct vocal fistula by passing a diathermy needle through the posterior tracheal wall into the oesophageal lumen. When expired air was directed through this fistula, excellent "fistula speech" was reported. This proved to be the forerunner of today's tracheo-oesophageal (T-E) puncture, shunt or fistula techniques of Singer and Blom (1980) and Panje (1981).

The modern era of SSR continued in the 1950's with experimentation by Briani in Italy (1952) and Conley in the USA (1959) of external vocal fistulas and internal tracheo-oesophageal fistulas respectively. Conley experimented with internal T-E fistulas constructed variously from mucous membrane, skin and free vein grafts and used with prostheses. There were serious complications in constructing stable vocal fistulas which did not become occluded or become stretched with severe leakage.

Asai's report (1966) from Japan of a high-entry internal tracheo-hypopharyngeal fistula was an important landmark in SSR. This method, of using a skin-lined shunt without a valved prosthesis, was extensively used and modified by Miller in the USA (Miller, 1967, 1968, 1971, 1974 and 1975). In this group of laryngectomees, where the fistula enters at a high pharyngeal level, the vibratory source for speech is most likely formed by the shunt orifice in the pharynx (Tarnowska, Jack and Mozolewski, 1982). It is also possible that many subjects may be using pharyngeal speech in that the pharyngeal mucosa and tongue-base contribute to the sound production.

Many surgeons could not repeat Miller's success with the Asai technique and have had serious difficulties in preventing occlusion of the skin tube fistula or stretching with severe leakage.

In the last decade, the procedures resulting in internal low tracheo-pharyngeal shunt as described by Staffieri (1980) and Amatsu (1978, 1980, 1985) have been widely applied and modified (Sisson et al, 1978, 1980; McConnel and Whitmore, 1980; Hall et al, 1985; Novak 1981 op cit; Tiwari et al 1982; Algaba et al 1983, 1984, 1985; Vuyk, 1985 and Algaba 1987). In this type of SSR the low entry tracheo-pharyngeal or tracheo-oesophageal fistula ends below the P-E segment. This is usually the site of vibration for oesophageal speech; therefore it seems logical that these shunt methods also employ the P-E segment as the main vibrator for speech.

Zwitman and Calcaterra (1973) were first to notice that the voice produced with a T-E fistula is similar to oesophageal speech production produced by the P-E segment. This suggests that the sound vibrator was the same for both methods and was confirmed radiologically by Zwitman's cinefluoroscopic analysis of T-E fistula speakers which revealed a similar P-E segment to that which was found in oesophageal speakers.

The different levels at which a fistula enters the vocal or digestive tract will have consequences for the different quality and pitch of speech perceived. Speech produced with a high-entry internal vocal fistula (i.e. tracheo-pharyngeal fistula) will have a breathy quality and is higher pitched than speech produced with a lower entry level vocal fistula (Nichols, 1969; Evans and Drummond, 1985).
The low-entry vocal fistula is to be preferred from a theoretical view because in these cases the resonator (vocal tract) is relatively unaltered and can be used over its entire length for acoustic modulation of the basic sound.

Serafini (1969, 1980) described a tracheohyoidpexy technique aimed at total restoration of laryngeal functions (respiration, deglutition and phonation). Most of his patients were able to phonate well post-operatively but permanent decannulation was achieved in very few. Other authors have reported good results with Serafini's technique or with modifications thereof (Arslan, 1972, Sisson et al 1978 op cit).

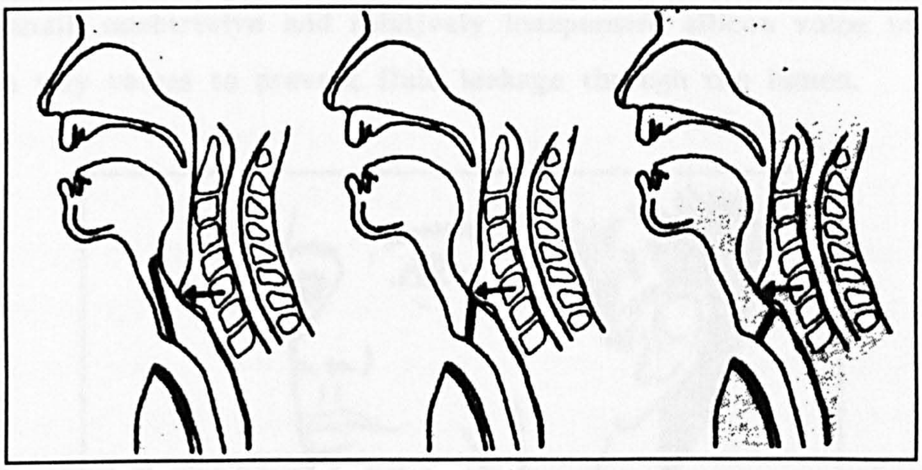


Figure 5. Positions of Vocal Fistulae.

1.3.6.1 Problems Associated with SSR.

All of the previously mentioned techniques of SSR result in a substitute speech production which is superior to that produced by mechanical devices (artificial larynges) and that produced by oesophageal speech. However, they present drawbacks, the most important of which are:

a. Leakage of food, fluids and saliva through the fistula from the oesophagus to the trachea and lungs, resulting in aspiration.

b. Obstruction, stenosis or spontaneous closure of the fistula, resulting in poor or absent speech production.

Several attempts have been made to overcome these problems. Amatsu et al (1986) used two oesophageal muscle flaps in an effort to gain some control over the opening and closing of the fistula. Ghosh (1977, 1985) created a neo-epiglottis to protect the fistula outlet in the oesophagus. Despite these efforts, good speech without aspiration was not always achieved-and remained the goal of surgeons involved in SSR.

1.3.6.2 Tracheo-Oesophageal (T-E) "Puncture" Techniques with Valved Prostheses.

Singer and Blom (op cit) and Panje (op cit) incorporated into a low T-E fistula small, unobtrusive and relatively inexpensive silicon voice prostheses with one way valves to prevent fluid leakage through the lumen.

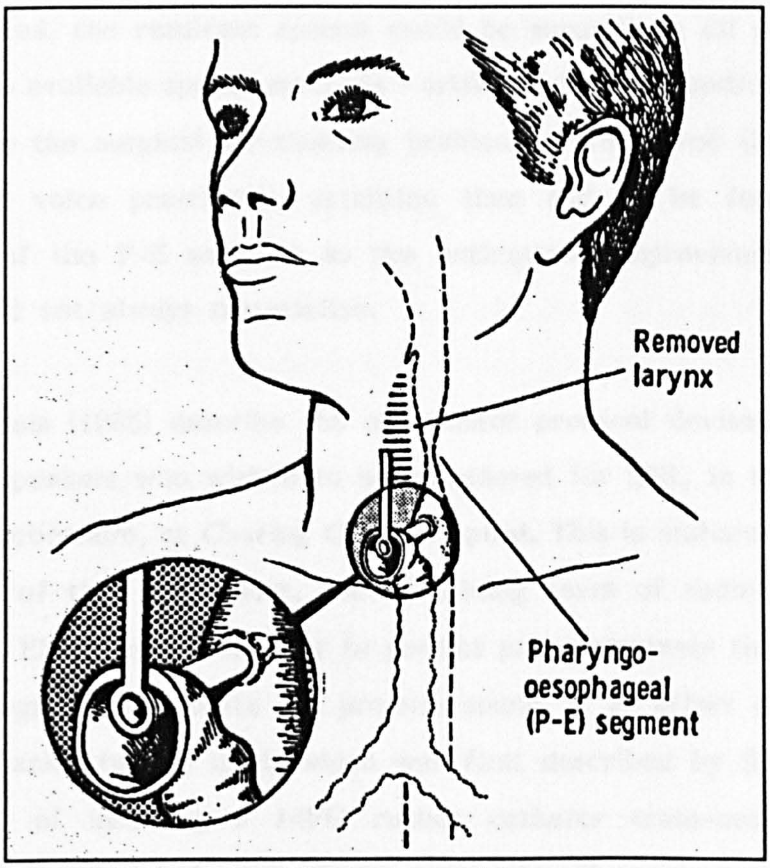


Figure 6.The Blom-Singer Procedure.

Advantages of these procedures include their simplicity, safety, reversibility of the procedure, freedom from aspiration and avoidance of long hospitalisation and heavy expense, as well as their suitability for many laryngectomees, including those who have had DXT. The puncture or endoscopic technique, foreshadowed by the crude punctures made by Guttman in the 1930's, uses a direct track made by a scalpel or sharpened trocar through the posterior wall of the trachea into the oesophageal lumen. The controlled midline puncture is stented for 36-48 hours and then a correctly sized and fitted valved silicon voice prosthesis is inserted into the surgical fistula. The results are often judged by naive listeners to be superior to oesophageal speech in terms of intelligibility and acceptability (Pinzola et al, 1988).

It was assumed that, once the surgical problems of aspiration and /or stenosis had been solved, the resultant speech would be superior in all parameters to other hitherto available speech methods - artificial larynges and/or oesophageal speech. While the surgical air-shunting problems were solved (by use of the puncture and voice prosthesis), attention then had to be focused on the competency of the P-E segment as the anticipated improvement in speech production did not always materialise.

Perry and Edels (1985) describe the assessment protocol devised for "failed" oesophageal speakers who wished to be considered for SSR, in the form of a Singer-Blom procedure, at Charing Cross Hospital. This is elaborated in section 3.1. As part of this assessment, the examining team of radiologist, speech therapist and ENT surgeon attempt to predict pre-operatively the competency of the P-E segment to vibrate and produce sound. In an effort to assess this, an air insufflation test is used, which was first described by Seeman (1967). This consists of inserting a 14FG rubber catheter trans-nasally, via the pharynx, into the top of the oesophagus, using videofluoroscopy to ensure that the radio opaque catheter tip is under the P-E segment. The examiner blows air "passively" into the oesophagus via the catheter as the subject attempts to phonate. This technique is fully described in the Methodology chapter (3.7).

In a large number of subjects examined the P-E segment would not release to allow air to flow freely for speech production; rather, the segment appeared radiologically to "lock" and tighten under air flow conditions. The subjects did not experience this phenomenon when swallowing, only when under air insufflation testing.

Singer (1988, op cit) describes the resumption of constrictor function in the laryngectomised population, although the function differs from normals. Early manometric investigations of the upper oesophageal sphincter described a reflexive elevation of tone during oesophageal insufflation (Clerf and Putney, 1942) and Singer proposes the same mechanism for pharyngeal constriction "hypertonicity" during oesophageal distention with insufflation. This is believed to occur in a large percentage of laryngectomees and may be a major factor in failure to develop oesophageal and tracheo-oesophageal speech (Singer and Blom, 1981; Johns and Cantrell, 1981; Cheesman et al, 1986, op cit; Mahieu et al, 1987).

Singer, Blom and Hamaker (1986) have experimented with pharyngeal plexus anaesthetic blocks for investigation of an initial observation of speech failure due to "spasm." After percutaneous ligocaine block of the pharyngeal plexus, and examination under videofluoroscopy,

> "..it was noted that marked relaxation occurred in the upper esophageal sphincter region with progressive improvement in airflow, reduction in resistance, and prolongation of fluency." (Singer 1988, op cit.)

These observations were applied to the tracheo-oesophageal speech failure group and two surgical techniques are described to release the "spasm" in the P-E segment. Firstly, a pharyngo-oesophageal constrictor myotomy has been advocated, ensuring that the myotomy corresponds to the upper sphincter (P-E segment) which will vary between 4 and 6 cm. Singer suggests that the myotomy should encompass the entire region; cricopharyngeus and inferior pharyngeal constrictor muscles. A second, and more controversial approach, is to use pharyngeal plexus neurectomy. The nerve dissection occurs at the level of the superior laryngeal nerve and the greater cornu of the hyoid bone. The identity of the nerves is confirmed by electrostimulation, which produces a brisk fasciculation of the corresponding constrictor muscle. It is suggested that the identity of the pharyngeal plexus is easier to confirm at the time of initial laryngectomy, rather than in subjects who have already undergone laryngectomy and developed spasm as a late complication with failure to develop speech (Singer, 1988, personal communication). As the surgical technique evolves, it may therefore be that neurectomy will be applied at the time of laryngectomy and myotomy reserved for the secondary state in order to prepare the laryngectomised pharynx for speech restoration.

1.3.6.3.Pharyngeal Constrictor Tone.

Constrictor spasm is virtually unknown following primary puncture, irrespective of the method of pharyngeal reconstruction. It probably occurs as a result of neuromuscular incoordination following the surgical disruption in the normal function of the constrictors. Laryngectomy alters the insertions of both the middle and inferior constrictors; particularly the inferior. The inferior constrictor anatomically has two parts: the thyropharyngeus and the cricopharyngeus. The anatomy of cricopharyngeus is disputed; usually it is described as the part of the inferior constrictor that is inserted onto the cricoid cartilage and is characterised posteriorly by an absence of the posterior pharyngeal raphe. It is, however, directly continuous with the striated circular fibres of the upper oesophageal sphincter, with which it is physiologically integrated. Following laryngectomy with a conventional repair of the constrictors, the middle constrictors are sutured together anteriorly and superiorly to the supra-hyoid muscles. The thyropharyngeus muscles are similarly sutured anteriorly but it is unlikely that the cricoid fibres of the cricopharyngeus are repaired due to their shortness. The upper oesophageal sphincter itself is not normally surgically disturbed.

The physiology of the sphincters during phonation also changes. During laryngeal phonation the lower constrictors are closed. With oesophageal speech, the constrictors must initially relax to allow oesophageal air charging and then subsequently contract for phonation.

Similarly, the physiology of deglutination is also disturbed. The second phase of swallowing consists of a wave of peristalsis in the constrictors followed by a reflex relaxation of the upper oesophageal sphincter. The opening of the sphincter was initially thought to be under nervous control but is now thought to occur following mechanical dilation of the sphincter by the bolus of food delivered by the peristaltic wave. Following laryngectomy, the normal peristalsis of the pharynx is disturbed and the normal reflex relaxation of the upper oesophageal sphincter may be disturbed, resulting in dysphagia which can only be overcome by compensatory hypertrophy of the newly repaired constrictors. It is therefore possible that constrictor spasm may be the result of such compensatory hypertrophy. This concept is supported by the findings at secondary myotomy, where it is the surgical impression that the fibres of the upper oesophagus are markedly stenosed in those subjects who have severe hypertonicity or spasm diagnosed during videofluoroscopy (Cheesman, 1988, personal communication).

This is further discussed in section 6.4.

1.4 Questions to address in this project:

There is a need to examine the physiological characteristics of the reconstructed pharynx in laryngectomees to attempt to assess the potential for speech acquisition, whether oesophageal or T-E puncture speech. Questions to address in our series of subjects are:

A. Why some laryngectomees fail to develop oesophageal speech.

B. Are the reasons for oesophageal speech failure likely to militate against successful speech rehabilitation with a T-E puncture.

C. How to reliably assess a subject's potential for SSR and to classify them in such a way that will determine their subsequent clinical management.

D. Which subjects are suitable, physiologically for a SSR procedure, and which need further surgery, prior to SSR, to adequately prepare the pharynx for SSR.

HYPOTHESIS:

The hypothesis to be tested in this thesis is that there is a continuum of tonicity in the reconstructed pharynx of laryngectomised subjects and that this anatomical / physiological state determines the success or failure of speech rehabilitation.

CHAPTER 2: LITERATURE REVIEW

2.1 Assessment

2.1.1 Purposes of Assessment

> "The basic goal of the initial speech evaluation is to assess comprehensively the patient's assets and liabilities and to assemble data from multiple sources to establish a realistic plan of treatment."

With this statement, Martin (1979) has outlined the purpose of an initial assessment. Information may be gained from medical/surgical reports, the operative notes, radiotherapy reports, the general medical history and from direct questioning and physical examination of the patient.

First, the operative records will give details of what structures were removed and on what structures remain. Whether the patient underwent a simple laryngectomy or whether he underwent associated surgery - e.g.unilateral or bilateral neck dissection (ie.removal of the lymph chain on either side of the neck). If there was a radical neck dissection, was the spinal accessory nerve sacrificed or saved? If it was sacrificed, shoulder dysfunction will be marked and this can impede mobility in the neck/shoulder region. This, in turn, may result in restrictive movements of the articulators for air-trapping activities prior to speech production.

Some patients undergo additional surgery involving the tongue and various maxillofacial structures. If the laryngectomee additionally underwent a partial glossectomy, there will be a marked articulatory difficulty as well as a problem with using the tongue to inject air into the oesophagus for speech.

If the laryngectomee's cancer resulted in the resection of the pharynx, as well as the larynx, there may have been a laryngo-pharyngectomy operation performed with a replacement of the pharynx by a stomach (gastric) pull-up or myocutaneous skin flap replacement. This will considerably alter the nature of the segment used as a vibratory source for speech and may result in a breathier, more "gurgly" speech quality; ie. speech of a deeper pitch, as in extreme creak of a laryngeal speaker, due to the nature of the more flaccid tissue (stomach walls) which are available for vibration.

Alternatively, no speech at all is produced as there are no tissues in suitable apposition to act as a vibrator for speech production.

When considering the functioning of the four basic processes used for speech production - namely, respiration, phonation, resonance and articulation - one can view speech production as an integrated system. Laryngectomy results not only in removal of the voice-producing mechanism (the vocal cords), but also elimination of the continued use of the trachea and lungs to supply air for speech. This is fully described in Section 1.2.2. and will not be elaborated here.

In the physical system after laryngectomy, the air supply for speech now becomes the oesophagus, phonation is produced at the re-constructed pharyngo-oesophageal (P-E) segment and resonance is produced by the nasal, oral and re-constructed pharyngeal cavities. The articulators (lips, teeth, tongue and palate) should remain essentially the same unless sacrifice of the recurrent laryngeal nerve (CN XII) has become necessary, in which case poor lingual strength and range of movement would result.

A structured oral/physical examination can provide valuable data about the range and strength of articulators, both in isolation and in co-ordinated movements. This physical examination may be compared with the detailed operative notes and any variation in results accounted for.

2.1.2 Methods of Assessment

Martin (op cit) pointed out the usefulness of viewing each patients speech and development on two levels - "macroscopic" and "microscopic". He referred to "macroscopic" as covering overall ratings of global speech proficiency, whereas "microscopic" referred to focal, more discrete assessment of specific skills needed to achieve that global speech proficiency. This idea had earlier been used by Wepman et al (1953), who indicated the value of assessing speech at two levels - "voice skills" (microscopic levels) and "speech skills" (which corresponded broadly to Martins's macroscopic level). "Voice" referred to the oesophageal sound, produced by air passing the P-E segment and Wepman et al, suggested a hierarchy of development of this sound as it becomes more controlled and then can be used for speech; i.e. shaped into words, first monosyllabic and then polysyllabic, then used in phrases and sentences. (See Table 2).

LEVEL	OESOPHAGEAL SOUND PRODUCTION	SPEECH PROFICIENCY
7	None	No speech
6	Involuntary only	No speech
5	Voluntary part of the time	No speech
4	Voluntary most of the time	Vowel sound differentiated Monosyllabic speech
3	At will	Single word speech
2	At will with continuity	Word grouping
1	Automatic	Good oesophageal speech

Ratings: - High = Inferior speech
Low = Superior speech

Table 2. The Wepman Scale of Speech Development.

Periodic testing of voicing proficiency ("microscopic assessment" in Martin's terms) can generate data from which to plan future therapy. Regular re-assessment can be used to evaluate the developing voice and speech skills, to modify or change the emerging speech of the laryngectomee and to identify and prevent "bad habits" from developing. Martin (op cit) stated that it is easier to identify and change undesirable patterns early in therapy than to try and alter well established malpractice once it has become habitual.

The majority of studies in the literature on assessment concentrate on assessing the development of communication skills and, largely, this has meant oesophageal speech. There are various descriptions of the procedures of assessing oesophageal speech and/or artificial larynx speech. These various forms of assessment may be divided into:

1. Informal observation;
2. Periodic testing of certain voicing skills;
3. Development of rating scales;
4. Overall functional communication profiles;
5. Objective measures, using instrumentation.

Any, or all, of these assessments may be useful, but it is essential to try and provide some objective measures of improvement - i.e. measures which are unbiased, replicable and consistent across testers and on repeated testing of the same subject.
"Informal observation" provides a "macroscopic" view of effectiveness and acceptability in communication; attitude towards therapy; psychological state and physical condition. Darley, Aronson and Brown (1975) use this term when describing assessments of motor speech performance and indicate the problems inherent in using diverse adjectives in examination reports. Informal observation relies on the clinical impressions and experience of the therapist concerned and is thus biased by the therapist's own experience and standards. Speech evaluation is purely descriptive at this level.

The use of labels such as "good" or "average" for oesophageal speakers has only limited currency as, unless the assessors are from similar clinical backgrounds and experience, comparisons are difficult. An "average" speaker at one clinic may be "good" at another. Thus more objective measures are needed.

Berlin (1963a) outlined the underlying criteria for successful oesophageal speech acquisition. These included achieving good consistency (ability to sustain voice throughout an utterance); long duration (ability to achieve a continuous note in a sustained way); and short latency of oesophageal sound production. (Latency refers to the time taken between air being injected into the oesophagus by the articulators and the sound being returned for speech). These emerging skills Gardener (1971) suggested should be regularly tested to chart the laryngectomee's progress towards fluent, intelligible speech. This use of "periodic testing of certain skills" by Gardener (op cit) indicates a step further towards objectivity.

From his clinical experience, Gardener (op cit) produced "norms" for successful speech acquisition - e.g. laryngectomees who became good speakers made twenty successful attempts on consonant-vowel (CV) speech combinations after fourteen lesson days. Poor speakers took twenty four or more lesson days. Based on these established "norms", Gardener (op cit) developed an oesophageal speech assessment which relied on the periodic testing of certain skills to chart progress in the individual laryngectomee. However, these measures referred to the acquisition of underlying skills which were felt to be necessary for good oesophageal speech, but not to their performance or usage.

"Developmental rating scales" were devised to correct some of the aforementioned problems inherent in the use of descriptive labels, such as "poor" and "good" oesophageal speakers. Wepman et al (1953 op cit) indicate four reasons for the existence of developmental rating scales:

A. To provide a common frame of reference for discussing prognosis and progress.

B. To judge the stage of recovery at any one time.

C. To enable the patient to visualise progress and to foresee the forthcoming stages in the development of good speech.

D. To serve as a motivating device.

Wepman et al produced a seven point rating scale, reflecting clinical observations made over many years of studying the development of speech after laryngectomy (See Table 2). As stated earlier, this encompassed their view of the bi-modal aspect of rehabilitation: progression of oesophageal sound production and then the acquisition of speech proficiency. It was their strongly held premise that better results were obtained in speech acquisition if the laryngectomee was "held back" from attempting speech production until he had achieved good control over individual sound production. Achievement at level 5 was felt by Wepman to be important prognostically as (1) the musculature for sound production is evident and (2) the voluntary control of oesophageal sound is possible. He felt that progress beyond this level was largely a product of practice and motivation.

"Functional communication profiles" provide a "macroscopic" view of rehabilitation, but in a more detailed and structured way than by informal observation. Damste (1979) developed an extended score sheet for judging the quality of oesophageal speaking habits, aimed at advanced users of oesophageal speech who wanted to check whether there was any room for improvement. This list was designed by a joint committee of one laryngectomee, two logopedists and a phoniatrist and it included a social skills inventory to assess the amount of speech used and in what settings. The laryngectomee rates his own performance, with a clinician, on a 1-5 point scale, where 1 = very poor and 5 = excellent. Although valuable for the information gained as to the attitudes and concerns of the laryngectomee, such self-assessment scales are not objective, which limits their research value.

Logemann et al (1980), presented the Northwestern Otolaryngology Communication Profile for head and neck cancer patients, which is also a functional communication profile. It was designed to measure both the amount and the mode of speech production in a variety of situations. It is suggested that the profile be presented both pre- and post-operatively in order to quantify the success of the patient's rehabilitation. This is done in terms of how closely the patient approaches his pre-operative level of communication function and the rate of progress towards that gaol. Five situational categories are included:

Home/With friends/At work/Telephone/In other situations.

By questionnaire, the laryngectomee indicates the mode of communication he uses in each situation:

Esophageal speech/Pencil and paper/ Mouthing/pointing or gesturing.

The value of all these assessments depends on the laryngectomee developing oesophageal speech. Thus, if he does not produce speech, he will not move from level 7 or 6 of the Wepman rating scale. However, this will not give any indication as to why he has failed to proceed in the assumed natural progression towards oesophageal speech acquisition. Large proportions of laryngectomees do not acquire oesophageal speech. Reported percentages of "failed" speakers post-laryngectomy vary enormously, as stated in section 1.3.5., and thus methods of assessing both oesophageal speakers and those laryngectomees who are having difficulty in achieving speech need to be found.

2.2 Failure to develop speech post-laryngectomy

Increasing use of radiotherapy and initial conservation surgery (eg. partial laryngectomy) has reduced the number of "ideal" candidates for acquiring oesophageal speech. Today's laryngectomee is likely to have had an advanced neoplasm which was treated by many and intensive therapies. Such a medical history leads to functional difficulties. After intensive radiotherapy, muscle atrophy and fibrosis are well recognised sequelae, which can interfere with the requisite anatomical and physiological base for oesophageal speech development and clinical experience shows that such patients have more difficulties.

With advances in skin grafts and flap reconstructions, more radical surgery is feasible, but this often results in long delays in healing and post-operative complications such as fistulae or skin/tissue breakdown. Consequently, the proportion of laryngectomees failing to acquire oesophageal speech may be rising. Salmon (1983) stated that ...

> "clinical observations suggest that side effects of radiation therapy such as scar tissue, poor healing, throat pain, dryness of mucosa, and the formation of oedema have a detrimental effect on the acquisition of voice. Thus, in countries such as Great Britain, where radiotherapy is used curatively as a first stage for most patients with cancer of the

larynx, we should anticipate even lower percentage rates for those who achieve esophageal voice..."(p. 144)

Duguay (1979) neatly categorised the problems in failure to develop oesophageal speech as falling into four areas:

1. Anatomical/Physiological
2. Psychological/Social
3. Teaching/Learning
4. O.G.K. (Only God Knows)

These categories were based on his extensive clinical experience. Other than (4) above, these groupings are not mutually exclusive and any/or of the other problem areas may exist in the "failed" oesophageal speaker. However, by using objective measures with instrumentation, it may be possible to exclude anatomical/ physiological problems. If, for example, "good" oesophageal speakers have measurable, definable pressure and resistance characteristics when speaking using the reconstructed pharynx, then one may predict the physiological prerequisites necessary for successful speech to result.
Should the laryngectomee fall outside the range experienced in "good" oesophageal speakers, then one might predict that no, or poor, speech production will occur. Should the laryngectomee fall within the range for "good" speech to result, but is not progressing well, then one might perhaps exclude anatomical/ physiological reasons and look towards (2) or (3) above. Objective measures can perhaps provide the most reliable predictive information with this population as it is transferable across clinicians, unlike clinical experience.

In considering the other types of assessment; informal observation will provide the information that speech has not developed and the experienced clinician may have an impression as to why this has occurred. While it may be possible to investigate the veracity of this "hunch", in reality the resources may not be available in every clinical setting. Periodic testing of certain skills and/or developmental rating scales will simply provide, on repeated use, the information that the laryngectomee is not progressing along a predictable path in speech development - but no information as to the cause of that "failure", is available from these assessments. Functional communication profiles will give an overall impression of the laryngectomee's day-to-day communication, but this may not indicate improvement in speech production and would not provide any reasons for a failure in speech development. In a recently published study, examining forty-eight "failed" oesophageal speakers (ie. those scoring 5 or more on the Wepman scale), using videofluoroscopy, Cheesman et al (1986), found physical barriers to oesophageal speech development in all of the subjects studied. Early published studies of the reasons for failure to develop oesophageal speech often did not assess the subject anatomical/physical potential for speech acquisition, concentrating instead on psychological and social reasons by using attitude questionnaires and clinical judgements.

Although motivation is relevant to acquiring any skill, the underlying physical propensity for oesophageal speech acquisition has to be present for success to result.

With the requisite anatomy/physiology many laryngectomees learn to speak without any regular speech therapy. Priest (1984), himself a laryngectomee, reported his own experience of speech developing "naturally" postoperatively and cited similar experiences from other laryngectomees. It is of interest that in the study by Cheesman et al, although not reported, the twenty laryngectomees who were operated on primarily and simultaneously given a myotomy (a muscle-cutting procedure, as described in 1.3.6.) through the upper oesophageal sphincter developed oesophageal speech effortlessly on the tenth post-operative day with no speech therapy. This would seem to provide further evidence for the fact that good post-laryngectomy voice acquisition depends on the surgeon's ability to control the repair of the pharynx. (Cheesman 1987, personal communication). It seems, therefore, increasingly likely that a fundamental reason for failure to achieve oesophageal speech may be in the area of anatomical/ physiological malfunctioning, with psychological, social, teaching and learning reasons becoming secondary, and often masking the underlying anatomical problems.

Robe (1956), by questionnaire and using his clinical experience, examined the role of certain factors in oesophageal speech acquisition in thirty two laryngectomees: twenty nine male and three female. The factors he considered included:

1. The type of operation
2. The site of the pseudoglottis (P-E segment)
3. Co-ordination of speech with respiration
4. Amount and type of speech training

He concluded that ..

> "the number of speech lessons did not influence speech fluency" (p. 184) and.."the type of surgery did not, in itself, determine voice results" (p. 186).

However, in this study, the laryngectomees were not examined to ascertain their physiological potential for speech production and therefore it is difficult to exclude anatomical problems as reasons for failure.

Gilchrist (1973) questioned fifty laryngectomees from the Royal National Throat, Nose and Ear Hospital in London, and then divided them into two groups:

A. Adequate - good speakers

B. Inadequate - poor speakers

He compared the two groups and his clinical impression was that the main reason for the poorer speakers was a lack of motivation. However, again this study did not use any objective measures of speech proficiency and did not examine any anatomical/physiological problems which may have been present.

Damste (1956) in a single case study of a pharyngeal speaker, outlined the prerequisites for success in learning oesophageal speech as being: a strong desire for speech; sufficient intelligence; physical strength and "not too much anatomical damage" (p. 998). These conclusions furthermore were supported by his clinical impressions from examining questionnaires returned by laryngectomees attending his clinic.

Shames et al (1963) in their ad hoc follow-up study found the following factors to have a significant correlation with speech proficiency: education; age; an intact cricopharyngeus muscle; the length of time before achieving voice; the amount of speech training (c.f. Robe above) and personality.

In reviewing the published studies on failure to achieve oesophageal speech, it is difficult to compare results. Most studies have relied on the author's clinical impressions of reasons for failure to achieve oesophageal speech, based on variously-designed attitude questionnaires, and often the authors have not clearly defined what is meant by a "failed" speaker in the study described.

2.3 Objective Assessments

Taking objective measures using instrumentation is a way of focusing, at a microscopic level, on particular attributes of speech and whether the physiological prerequisites for acquiring oesophageal speech are available to a particular laryngectomee.

A survey of papers from the literature provides valuable information on the parameters to study that may be relevant to oesophageal speech production. Previous studies have examined the following aspects of oesophageal speech production: frequency; intensity; duration; pressure; airflow; maximum phonation time; and EMG studies.
Three main categories of papers can be distinguished: studies of normal (laryngeal) speech; studies of oesophageal speech; and studies after surgical rehabilitation of speech.
In addition to objective acoustic and aerodynamic measures, a way of assessing the anatomy and physiology of the reconstructed pharynx is desirable. Cinefluoroscopy and/or videofluoroscopy is invaluable for viewing, under X-

ray, the size, shape and co-ordination of the pharynx and pharyngo-oesophageal (P-E segment) dynamically when the laryngectomee attempts to swallow or speak. This technique involves the laryngectomee swallowing a solution of Barium Sulphate (Baritop) to coat the oral, pharyngeal and oesophageal tract and a radiologist filming the areas concerned at rest and during swallowing and speech. "Spot" or static films are taken (radiographs), together with the videotaped sample, to give combined information on the anatomy and the inter-relationship of structures during speech.

While videofluoroscopic films are being taken, it is possible to take simultaneous acoustic and aerodynamic measures of oesophageal speech production.

2.3.1 Cinefluoroscopic Studies of Alaryngeal Speech.

The early, and extensive, work of Diedrich and Youngstrom (1966) led the field in this area of study for many years. They describe their study of twenty seven oesophageal speakers who underwent cinefluoroscopy (video was not available at that time) and lateral spot films while undergoing various speech manoeuvres and having oral air pressure studies taken. Speech was recorded on magnetic tape, during a connected reading passage. Inhalation and exhalation oral breath pressure was measured during five trials with an oral manometer. Six spot films were taken for each subject at rest and during phonation of the vowels /i/; /a/; and /u/.
The aim of the study was to investigate :

> "the morphology and dynamic movement patterns of the lips, tongue, palate, pharynx and esophagus during post-laryngectomy speech." (p.58).

The areas they examined included: area and width of the hypopharynx; cervical level of the P-E segment; morphology (i.e. shape) of the P-E segment; length of the P-E segment; movement of the neoglottis; width of the oesophagus; palatal function; air intake; graphic analysis of articulators; speed of air intake; pre-phonation period; chronological age and months post-operative; surgical data; audiological evaluation; arthritis; and oral manometric measures.

From this extensive study, they concluded:

> "Esophageal speech skill following laryngectomy was not related to morphological function of the reconstructed pharynx and pharyngo-esophageal junction. These results would support those who believe that psychological factors are probably more important to the development of good speech than the anatomical factors." (p.61).

Unfortunately, although describing the P-E segment, no attempt was made in this study to assess its physiology during phonation or to assess the opening pressure of the P-E segment which was needed to achieve phonation in the subjects studied. It is not therefore the case that this study examined the "morphological function" of the reconstructed pharynx and P-E segment; rather it described the morphology (i.e. form) of the structures during manoeuvres, but no attempt was made to assess their potential for function. Additionally, all the subjects examined had acquired œsophageal speech and, without examining subjects who had failed or who had difficulty in speech acquisition, it is difficult to conclude that the morphology plays little or no part in success or failure to develop speech.

This work has influenced many clinicians and is often quoted in defense of the premise that failure to develop oesophageal speech after laryngectomy is more likely to be related to psychological problems than to anatomical/physical ones, but this conclusion from the study appears to be based on a misleading premise.

2.3.2 Acoustic Properties of Oesophageal Speech

Various studies objectively measuring acoustic properties of oesophageal speech - frequency, intensity and timing - have been published.

2.3.2.1 Frequency (Pitch)

Curry and Snidecor (1961) indicate the difference between frequency and pitch in the following way:

> "the frequency of an auditory stimulus can be measured in the complete absence and independently of any listener.
>
> The term "frequency" should be considered as a distinctly physical attribute of the auditory stimulus. The unit is cycles per second (or Hertz) "Hz". Pitch is an auditory experience identified with the listener. Pitch should be thought of as the listener's personal reaction to the pure frequency stimulus; the unit of pitch is the "mel"" (p. 415).

Thus, pitch is the psycho-perceptual correlate of frequency. Frequency is measured; this, in turn, is perceived as pitch e.g. one hears middle "C" and the octave below as perceived pitch, but these are represented physically as 256Hz and 128Hz.

Since laryngectomy always results in the sacrifice of tissue essential to normal speech production, clinicians and investigators have assumed that the principal factors affected by laryngectomy are those related to the voicing source. Hence, there have been a number of studies to specify the fundamental frequency (Fo) and phonation time characteristics of oesophageal speech. The fundamental frequency characteristics are well understood and among these studies the three of Snidecor and Curry (1959, 1960) and Curry and Snidecor (1961 op cit) are primary references for this topic.

Although the fundamental frequency characteristics of oesophageal speech are well recorded, the nature of the anatomical and physiological mechanisms underlying the regulation of the fundamental frequency and vocal intensity are not fully understood. Damste (1958, op cit) believed that the larynx of a normal speaker and the P-E segment of the laryngectomee operated in essentially the same way, ie. by the Bernoulli effect. He stated..

> "the normal vocal cords lie against one another along a rather great depth during phonation in the chest register, as it is clearly shown by frontal tomograms of the larynx. The subglottic pressure forces them apart. As soon as the air-bell has passed, they close again, beginning at the bottom. This closure is not only effected by the elasticity of the cords but in particular by the Bernoulli effect, the decreased pressure on the medial side of the cords when a rapid current of air passes through. Thus, the air current is , so to say, cut into slices by the glottis, the air penetrating through the glottis in separate puffs of air. Fundamentally, this is not different for the pseudoglottis. Only the distance the air-bells have to cover through the glottis is not a few millimetres, but a centimetre or more" (p. 43).

Determination of the frequency of sound in oesophageal speakers is made more difficult by the fact that only rarely are the sound signals purely periodic. More commonly, the vibrations of the P-E segment are aperiodic, so that it is difficult to speak of a tone. The occurrence of aperiodicity was attributed to three causes by Damste (op cit). The first reason was due to variations in sub-neoglottic pressure. The volume of air in the oesophagus is only small (80ccs); any fold in the mucous membrane below the level of the P-E segment may easily influence the supply of air. Secondly, the length and elasticity of the P-E segment are not so constant and adjustable as in the normal glottis. The third, and most important, cause is the accumulation of mucous above the mouth of the oesophagus, a handicap for most laryngectomees. Thus the air is forced in a highly irregular way through the P-E segment through a varying layer of secretions. It is therefore noise that is produced, rather than a tone.

There is some confusion in the literature as to whether oesophageal speech can truly be said to have a fundamental frequency, as it involves aperiodic sound. Often, what is being recorded is a regular note at the P-E segment (vibrator) with aperiodic overtones. Additionally, the results of analysis using "good" oesophageal speakers may be very different from analyzing those who have difficulty in producing sound. This point is discussed in the paper by Curry and Snidecor (op cit).

Most current data using sound spectography, suggest that the average Fo of a male oesophageal is 65Hz, or half that of the normal male laryngeal speaker. The female range is 80-87Hz. Spectrographic studies (Van den Berg and Moolenaar-Bijl, 1959) have shown mean fundamental frequency ranges of 50-100Hz. Damste (op cit) studied twenty five laryngectomees and tape recorded them during production of monosyllabic words; sustaining vowel sounds; producing the highest and the lowest achievable pitch and in spontaneous speech.

In sonographic analysis of these recordings, they had the same range with a median of 69Hz; Kytta (1964) found a mean of 50.40Hz in eighteen subjects studied. On the other hand, Perry and Tikofsky (1965) using narrow band spectography in conjunction with oscillographic wave tracings, found a mean Fo for vowel sounds in oesophageal speakers to be 29-37hz, with 56% of the subjects studied being below 30Hz.
Damste (op cit) indicated that...

> "the low fundamental tone of the oesophageal voice, together with the fact that it is rich in overtones, is the reason that it is very suitable for an accurate production of vowels" (p.44).

Shipp (1967), Hoops and Noll (1969) and Weinberg and Bennett (1972) provide data that verify these early observations. Weinberg and Bennett (op cit) revealed that the average Fo characteristics of oesophageal speech may vary as a function of speaker sex and may highlight the extensive variation in average Fo exhibited by oesophageal speakers. They indicate the need to control for acoustic differences between male and female oesophageal speakers. Many of the earlier studies (Damste 1958, Shipp 1967) did not state the sex of the subjects used in their studies.

Isshiki's findings (1964) for normal laryngeal speakers are relevant to the consideration of development of a higher pitch in oesophageal speakers. To obtain increased intensity at a higher pitch, the flow rate through the larynx is raised. Since oesophageal speakers have only one-third the flow of laryngeal speakers, it is not surprising perhaps that they cannot obtain good intensity levels at higher pitch.

The work of Van den Berg et al (op cit) discusses certain aspects of increasing the frequency variability of oesophageal speech.

> "Pitch and intensity are correlated in oesophageal speech. A low pitch is produced with a low intensity of the voice; a high pitch with high intensity. This increase of the pitch at increasing intensity, ie. at increasing the flow rate through the P-E segment, is caused by the Bernoulli effect of air which escapes through the narrow opening. The walls are sucked toward each other by the negative pressure in the P-E segment and this effect increases the overall stiffness of the muscular structure, which then vibrates at a higher frequency as its mass remains the same." (p. 304)

Curry and Snidecor (1961) compared the "usual pitch" for each of Damste's (op cit) twenty subjects with the median frequency of each of their six superior oesophageal speakers.

The results of the Curry and Snidecor studies - that the average Fo for male oesophageal speakers is 65 Hz and the magnitude of the Fo variability is four semitones - indicates that the amount of Fo variation produced by oesophageal speakers is equal to, and often greater than, that produced by laryngeal speakers. Damste's 20 speakers, considered as a group, had a median frequency of 67.5Hz compared with 63.3Hz for Curry and Snidecor's six superior speakers. The individual Damste subjects showed a much wider range of frequency production than did the Snidecor group; the highest "usual frequency" in any of Damste speakers was 185Hz; this value compared with 80.0Hz for the highest median frequency of any one of the Curry and

Snidecor subjects. One possible explanation for this result is the sex of the subjects. Unfortunately, in his study, Damste did not specify the speaker sex. All of the Curry and Snidecor subjects were male.

Snidecor and Curry (1959) selected six superior oesophageal speakers by common consent of eight skilled and experienced judges who were asked to rate the effectiveness of ten audiotaped samples of oesophageal speakers. The judges rated aspects of pitch, intelligibility, rate, loudness, quality, and articulation. From this, six "superior" oesophageal speakers were selected. All were male, of at least four years post-laryngectomy. The subjects were recorded on a dual channel magnetic tape recorder. On one channel was recorded an intensity modulated 1000 cycle tone as an air change signal. The second channel recorded speech - viz. "The Rainbow Passage" (Fairbanks) which was recorded " as if speaking to an audience of 25 people."

Pitch curves were plotted from the frequency measurements and measures of pitch thus computed, using a phono-graphic technique (this involved making temporal measurements from a photographic tape record with a simple analogue computer and checking these results against a one-tenth of a second time line recorded on photographic tape), originated by Metfessel (1927). The central tendency and variability for frequencies of the superior oesophageal speakers were contrasted with a set of similar measures in a group of superior laryngeal speakers i.e. "normals" as provided by Pronovast (1942).

The expected lower fundamental frequency level was recorded in the oesophageal speakers; a difference of almost exactly one octave was noted (63.27 Hz for oesophageal speakers compared with 132.1 Hz for normals, corresponding to the musical notes C2-C3). Both the mean and median frequencies used by oesophageal speakers were one octave lower than normals.

However, somewhat surprisingly, it emerged that when the pitch variability was compared, 13.21 tonal range was recorded as a mean total frequency range for the superior oesophageal speakers compared with 10.05 tonal range for normal laryngeal speakers. Thus, the amount of fundamental frequency variation present in connected discourse produced by oesophageal speakers is equal to, and often greater than, that produced by laryngeal speakers.

One possible explanation for this surprising result may lie in the fact that, as Damste (op cit) observed, the measured frequencies and perceived pitch at lower frequencies are not readily detected. Thus one should re-evaluate the premise that oesophageal speakers have a restricted pitch range; it may be that the frequency range is wide, but at too low a pitch to be detected. The distribution of frequency usage in the Curry and Snidecor (op cit) study showed greater variability for oesophageal speakers than for laryngeal speakers; standard deviation values were 7.0 and 4.3 tones per second respectively.

Damste (op cit) noted that ...

> "the fundamental frequency of oesophageal speech is difficult to determine. Frequency is low and the sound is very complex; in other words, the fundamental tone is accompanied by a large number of relatively strong overtones. The ear is not so sensitive in the frequency range of the fundamental, much more sensitive in the range of the partials, and the latter contain a large part of the total sound energy" (p.42).

It is doubtful if the average fundamental frequency level is raised much, if at all, above 66-70Hz (Snidecor and Curry, op cit). Stevens and Volkman (1940), presented experimental evidence to clearly indicate that in the frequency range below 100Hz the perceptual or pitch aspects of a stimulus are greatly reduced when compared with the frequencies above 100Hz.

As described earlier, Snidecor and Curry (op cit) in their study showed a mean total frequency range for oesophageal speakers to be 13.21 tones, compared with 10.5 tones for adult and normal speakers. Despite this apparently favourable comparison between the frequency ranges of the two subject groups, the oesophageal speakers were judged to have a "restricted pitch range" when the tape recordings were evaluated. The frequency measure indicated a considerably greater movement than was apparent in pitch to the listener. This significant lack of agreement between the measured (frequency) and the perceived (pitch) aspects of low frequency speaking performances can be readily understood from the results of Stevens and Volkmans experiment.

Due to the particular frequency-pitch relationship in the frequency range of oesophageal speech, extensions of the pitch range for the oesophageal speaker may not be as productive perceptually as would be the case at the higher frequencies which are characteristic of the normal speaking range.

2.3.2.2 Intensity (Loudness)

Normal laryngeal loudness is between 65-79dB at a comfortable level with a range of 0-95dB (Davis, 1979). Oesophageal speech, by contrast, has an average range of 40-50 dB (Edels, 1984) so it is not surprising that many laryngectomees complain of not being heard against background noise. Objective measures can be made using a sound level meter or by using a VU meter on a tape recorder. Alternatively, sound pressure level (SPL) may be recorded, as in the study of Hoops and Noll (1969). They studied twenty-two male laryngectomees who had no foreign accent or dialect variation from general American, who used oesophageal speech as their primary mode of communication and who were capable of reading the "Rainbow Passage" (Fairbanks). Each subject was recorded on colour motion film and rated by thirty experienced judges (speech therapy students) and thirty naive listeners (who had never heard an oesophageal speaker before). Each judge was asked to rate each speaker on a seven point rating scale on the basis of general communication effectiveness. The relationship between listener judgements and acoustic attributes of frequency, intensity and rate was investigated.
The vocal sound pressure level of the subjects was investigated using a Bruel and Kjaer power level recorder and a statistical distribution analyzer. The speech samples were supplied to the power level recorder, which had been

calibrated with a 80 dB test tone, by the tape recorder. The power level recorder activated the statistical distribution analyzer which samples per position on the tape recorder at a sampling rate of 10 per second, thus producing a frequency of occurrence distribution of sound pressure level for each subject's reading of the "Rainbow Passage." From this frequency distribution, the following values were obtained for each subject: (i) Mean sound pressure level; (ii) Standard deviation sound pressure level. The mean for the twenty-two subjects ranged from 57.02 dB to 67.57 dB. The average mean pressure level was 62.40 dB. This did not correlate with oesophageal speech proficiency.

Snidecor and Isshiki (1965), in examining six male oesophageal speakers, showed that a superior male oesophageal speaker was capable of intensity levels of 86 dB with a range of 20 dB for both upward and downward intensity sweeps. They noted that pitch and intensity vary..

> "almost entirely by the airflow with secondary participation through the action of the pseudoglottis" (P-E segment).

Later they state:

> "since there is much less control of the pseudoglottis in esophageal speakers than for the glottis in normal subjects, it is assumed that the mean flow rate is of greater importance in controlling the intensity than in normal speakers." (p.215).

Isshiki's (1964) findings regarding airflow rates, pitch and intensity in laryngeal speakers may be related to this statement.

He stated that at a very low pitch, the intensity of the normal voice is raised by increasing the glottal resistance, rather than by increasing the flow rate; furthermore, with low resistance of the glottis, increase in flow rate would cause the glottis to remain open and the voice would become weak. As the laryngectomee cannot achieve an increase in P-E segment resistance, he has to provide a compression in an alternative way. Clinical experience has shown that digital compression of the neck at the level of the P-E segment will increase vocal intensity (Logemann 1983). Zinner and Fleshler (1972) found in measuring pressure changes at three sites in the oesophagus during intensity tasks, that medial pressure in the oesophagus was significantly raised when the patients attempted production of vowel sounds at greater intensity (see later comments). These studies would seem to indicate the relevance of including, if possible, aerodynamic as well as acoustic measurements when undertaking objective measurements of oesophageal speech production.

2.3.2.3 Duration (Time)

In oesophageal speakers, duration has two components: the time needed to charge the oesophagus with air and the length of sound produced on that air change by the controlled release of air through the vibrating P-E segment. At either end, the oesophagus is protected by a sphincter. Superiorly the P-E segment separates the hypopharynx from the oesophagus. Inferiorly, the cardiac sphincter separates the oesophagus from the stomach. The P-E segment is anything from 0.8 to 1.2 inches long, while the cardiac sphincter is 2 ins long (Edels, Y., op cit). The oesophagus lies behind the trachea and the lungs within the thoracic cavity, where it is subject to alterations in pressure during pulmonary respiration. Unlike the trachea, which retains a permanent patency in both normal and oesophageal speakers, the oesophagus is collapsed at rest. It normally opens to allow the passage of food by a peristaltic wave. The oesophagus at rest has a negative pressure registering between -4.0 and -7.0mmHg below atmospheric pressure. This may increase to -10 to -20mmHg, during pulmonary inspiration. Whereas, prior to surgery, the oesophagus played no part in voice, for the laryngectomee it must now act as a new air reservoir; expulsion of its air acting as an initiator for vibration at the P-E segment. Due to the limited air reservoir available to be used as an initiator for phonation, the laryngectomee must re-inflate his phonatory system more frequently than the normal (laryngeal) speaker. The speed and ease with which he can accomplish this feat is reflected in the smoothness and continuity of his connected speech (Berlin 1963a). Short latency and good duration were two skills which appeared to Berlin to have.. " a face validity relationship to adequate oesophageal speech" (p.42).

In his study, six judges were trained to time laryngectomees (using a stopwatch), from the time the laryngectomee signalled he was inflating his oesophagus until the time the examiner perceived vocalisation, which was defined as a sound of 0.4sec or longer. 38 male subjects were assessed and divided into good (N = 28) and poor (N = 10) speakers on the basis of recordings of a standard interview. Three experienced speech therapists and two otolaryngologists, at separate sittings, were asked to rate the oesophageal speakers on a 5 point scale where 1 = excellent and 5 = very poor oesophageal speech. Subjects whose mean ratings were 2.8 or better were put into the "good" speaker group. Subjects whose mean ratings were 4 or worse were classified as "poor" speakers. Four ratings fell initially between 3.0 and 3.9 and, as they would not be fitted into either group definitively, these subjects were dropped. Another patient developed an oesophageal stenosis so was also dropped, leaving a total of thirty three subjects. "Good" oesophageal speakers were able to develop 0.2 sec latency (ie. time needed to inflate the oesophagus). Good speakers showed almost no range or variability (r= 0.2-0.6 sec; s.d.= 0.03 sec). By contrast, "poor" oesophageal speakers had a mean latency as high as 1.3 sec, a range of 0.2 - 2.0 sec and a standard deviation of 0.88 sec.

In sustaining duration of the vowel /a/, the range of duration in good speakers was between 2.2 sec and 3.6 sec; a longer duration did not appear to be necessary for good speech. Once a laryngectomee surpassed 1.8 sec, the examiners felt that this indicated an encouraging prognosis for ultimate development of good oesophageal speech (Berlin, op cit).

Adequate duration of sound is an important skill to develop as an oesophageal speaker cannot articulate many words on a short, uncontrolled burst of sound. Christensen and Weinberg (1976) examined vowel duration characteristics of ten oesophageal and nine normal speakers. High quality tape recordings were made of both oesophageal and normal speakers uttering vowels spoken within various consonant environments. Stimulus materials were 32 symmetric CVC syllables eg./pip/. Eight consonants (/p/; /t/; /k/; /b/; /d/; /g/; /s/; /z/) were combined with four vowels (/i/; /e/; /a/; /u/). The stimuli were recorded within the sentence frame "________is a word".

Overall vowel durations of oesophageal speakers were consistently longer than normals, indicating that oesophageal speakers do not compensate for their striking diminution in air supply for speech by decreasing vowel duration.

Zinner and Fleshler (op cit) looked at pressure changes at three sites in the oesophagus during length (duration) and loudness tasks. They assessed eleven male laryngectomees, used videofluoroscopy to confirm the catheter site and used pressure transducers to measure at one-third, two-thirds and at the base of the oesophagus. Their subjects phonated on a vowel sound /a/ for a mean of 1.45 sec (range 0.44 - 3.43 sec). There was no significant difference at the levels of the oesophageal pressure recorded during varying lengths of phonation; the changes in pressure only occurred during the loudness tasks, when the medial pressure was significantly raised compared to the proximal and distant levels of recording.

2.3.3 Aerodynamic Studies of Oesophageal Speech

These studies may be divided into different sections, depending on which objective measurements are to be studied.

2.3.3.1 Pressure

In reviewing the literature concerning pressure measurements during oesophageal speech production, it is difficult to compare results due to methodological difference in the studies reported. Additionally, three different recording measures are used - mmHg or cm H_2O or KPa. (1 mmHg = 13.6 cm H_2O. 1KPa = 9.8cm H_2O). Thus, all reported results have to be converted to a common scale for comparison to be possible. (See Table 3).

INVESTIGATOR	PRESSURE RECORDED mmHg	CIRCUMSTANCES
Damste (1958)	3.68 - 73.50	Manometry in the pharynx & oesophagus during speech manoeuvres
Vrticka (1963)	40.00 - 80.00	Trans-nasal insufflation Speech tasks not stated
Seeman (1967)	10.00 - 50.00	Trans-nasal insufflation
Zinner (1972)	11.80 - 68.70 (mean = 40.20)	Duration tasks
	20.30 - 83.01 (mean = 52.20)	Intensity tasks
Winans (1974)	13.00 - 39.00	Intraluminal manometry at rest

Table 3. Results of recordings of pressure measurements.

In reported studies, experimental conditions have varied considerably, both in the terms of the type of measurements taken and in the tasks / manoeuvres undertaken.

Some experimenters have measured resting cricopharyngeal pressure (Winans et al, 1974), in the belief that it was this muscle which predominated as the vibrator for speech post-laryngectomy. Others (Seeman 1967), have measured air pressure during trans-nasal insufflation testing. This involves passing a small bore (14FG) rubber catheter (tube) via the nose and oropharynx into the upper oesophagus. Air is passed into the oesophagus via the catheter and the subject opens his mouth for "passive" phonation - i.e. he is instructed not to try for sound, but simply not to resist the air returning from his oesophagus and vibrating for speech at the P-E segment. A side arm normally records the pressure during various speech manoeuvres.

Some experiments have used pressure measures independently of speech (e.g. Damste, 1958), while others (Zinner et al, op cit) included simultaneous acoustic measures of length of phonation and loudness (intensity) or attempted to study pressure measures and correlate these with speech fluency (Bozymski, 1972).

Different experimental conditions and measurement techniques may account for the divergent observations among investigators. In the published literature, the following pressure studies are relevant to this experiment:

Thirty eight laryngectomees had the tension of the upper oesophageal sphincter determined by Seeman (op cit) by using a trans-nasal insufflation test. Repeated cinefluoroscopy was used to examine "the gradual functional adaption of the oral parts of the gastro-intestinal tract". The tension values of the upper oesophagus did not change. Fifty three laryngectomees (75.71%) had good voice; eight (11.43%) attained basic communication; in four (5.7%) cases, the speech was insufficient for practical communication and in five subjects (7.14%) the speech was not evaluated due to their lack of attendance. Of the thirty eight subjects in whom oesophageal pressure measures were taken, twenty four cases (63.16%) had low readings (40mmHg); nine cases (23.68%) had increased (40-80mmHg) readings and in five cases (13.16%) the recordings were high (89mmHg). He noted that..

> " in patients with a high tension of the upper oesophageal sphincter, the rehabilitation was difficult in most cases " (p.3)

Damste (1958 op cit) measured pressure simultaneously in the oesophagus and pharynx and the sound produced.
He used two trans-nasal catheters, one in the oesophagus and one in the pharynx, connected to a manometer. In using such a catheter, there is marked interference in the voicing system, although Damste stated that.."the catheters were never a serious impediment to speech"(p.15). In twenty four subjects, there was a pre-phonatory phase in which the pharyngeal pressure went to 30-40cm H_2O.

This, Damste stated was ..

> "five times as high as the pressure during the hold in the articulation of plosives and twice as high as the highest pressure found in swallowing."(p.22).

An average oesophageal air pressure of 30cm H_2O was found in good speakers during phonation. However, in one subject a recording of as low a pressure as 5cm H_2O was made, and some were as high as 100cm H_2O. Unfortunately, the sex of the subjects is not stated, neither are the "good" oesophageal speakers defined, except to state that this..

> "was based on the subjective impression of the speech therapist and the author." (p.12).

Zinner and Fleshler (op cit) measured intraoesophageal pressures during phonation in eleven laryngectomees, whose ages ranged from 49-71 years. This study examined pressure changes at three sites in the oesophagus during length and loudness tasks. Fluoroscopy was used to confirm the catheter site and three simultaneous pressure sites were recorded at three levels: one-third, two-thirds and at the base of the oesophagus.

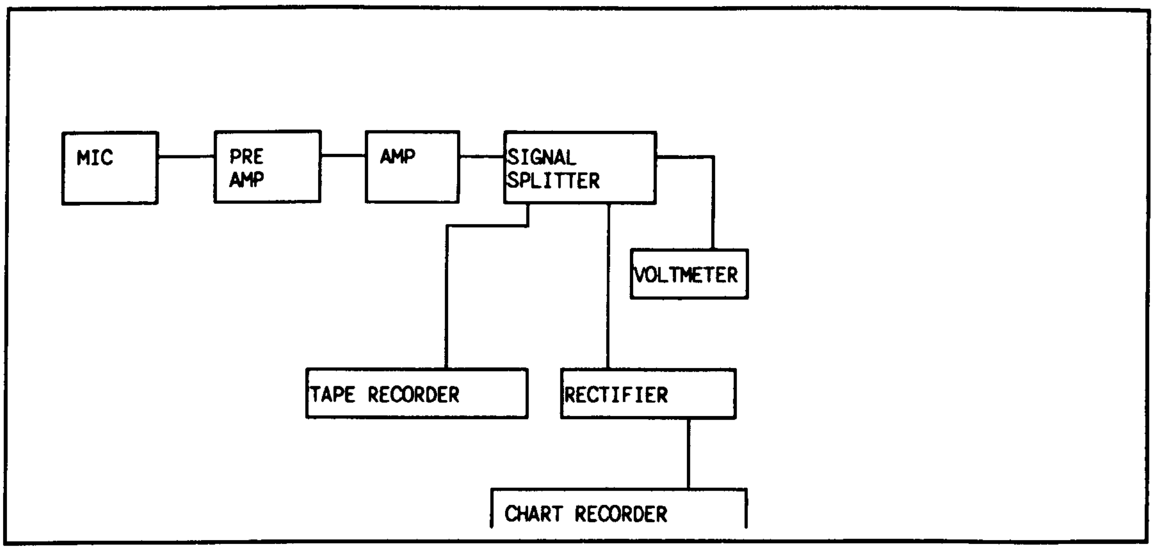


Figure 7. Equipment design for Zinner & Flesher Pressure Measurements.

Subjects were asked to phonate for five times as long as they could and eight times as loudly as possible, using a voltmeter for feedback. This examination provided information on the magnitude of pressure changes, the variation in different subjects, the levels in the oesophagus and the consistency of control in speech. Individual variation in recordings was considerable, but no significant difference was recorded at the different levels of oesophageal pressure on different trials during the length of phonation tasks. However, during loudness tasks, the medial pressure recorded was significantly higher than proximal or distal pressures during maximum intensity tasks. This would seem to indicate that loudness is achieved by increase in muscular tension/pressure at the P-E segment. (See earlier comment on Isshiki's study).

Due to the tremendous variability of pressure recordings in her Ph.D study, Salmon (1965) questioned whether good and poor oesophageal speakers could be predicted on the basis of intra-oesophageal measures. However, she did not use acoustic measures, which may have yielded significant information. Her subjects were six adult males. Three were classified as "good" speakers and three "poor" on a sustained phonation task as had earlier been described by Berlin (1963a, 1965). Pressure sensors recorded variations above, below and within the P-E segment with cinefluoroscopy used to confirm the sites of recordings.

Vrticka and Svoboda (1963) examined 70 laryngectomees and classified their speech proficiency on the Robe scale (see 2.1.2). They recorded the oesophageal pressure in 38 of those subjects and found that 24 subjects had low (ie <40 mmHg) pressure measurements and had good voice; 9 subjects had increased pressure measures (40-60 mmHg) and basic voice production, whereas 5 subjects had high (>80 mmHg) recordings and poor speech production. Unfortunately, 27 subjects could not be recorded.

More recently, Baugh et al (1987) used air insufflation testing to obtain intra-oesophageal pressure measurements from twenty one consecutive laryngectomised subjects who were candidates for surgical speech rehabilitation. The authors associated pre-operative pressure measurements with post-operative results and found that all subjects demonstrated at least limited sound production with insufflation at 1 and 3 litres per minute flow rate from a compressed air source. Subjects who demonstrated intra-oesophageal pressures of < 20 mmHg developed fluent tracheo-oesophageal (T-E) speech. Subjects with intermediate pressure recordings (21-39 mmHg) developed non-fluent T-E speech which did not improve with speech therapy. In contrast, subjects with intra-oesophageal pressures > 40 mmHg did not develop speech and were non- speakers. Intra-oesophageal pressure measurements were significantly different between T-E speech fluency groups (p <0.01) and were associated with a specific speech fluency group. The researchers in this paper did not use videofluoroscopy as they felt that...

> " the subjective nature of the assessment, the significant radiation exposure, and the unproven efficacy of videofluoroscopy to predict the tonicity of the pharyngeal constrictors during oesophageal insufflation preclude its use as a clinical method for predicting T-E speech fluency at this time "

The results of the pressure studies pertinent to this thesis are summarised in Table 3.

2.3.3.2 Volume of Air Use

Snidecor and Isshiki (1965) assessed six male oesophageal speakers and measured volume and airflow using pulmonary and acoustic measures. Three manoeuvres were undertaken: speech, swallowing and breathing. During this time, simultaneous recording of the rate and volume of oronasal airflow, respiratory movements and voice signal was taken. Four recordings were made on each laryngectomee during the following tasks: reading the Rainbow passage (Fairbanks, 1940); sustaining phonation on the vowel /a/; counting as far as possible on one air charge; articulating CVC syllables; swallowing and breathing. Analysis of the data revealed the following:

1. Total air volume used for the 51 word passage ranged from 372 - 1115 cc in contrast to a laryngeal speaker who used 3020 cc. The poorest speaker used the least air. Air volume was not related to the type of speaker in continuous speech, rather the airflow rates on respiration were important. These rates were determined by the volume of air trapped, the amount of pressure in the oesophagus and the resistance supplied by the P-E segment. Volume per syllable ranged from 5cc to 16cc; figures substantially higher

than those which had been earlier reported. The laryngeal speaker used 43cc per syllable.

2. Ratio of phonated time to total time ranged from 0.38 to 0.57 which is below mean figures of 0.6 to 0.75 reported for normal (laryngeal) speakers but which, nevertheless, indicates a good level of efficiency.

3. Flow rates for sustained /a/ in superior oesophageal speakers were recorded at 20 - 75cc/sec for soft and medium voice and 85 - 100cc/sec for loud voice with a mean flow rate of 72cc/sec for "easy" phonation. These values are much greater than for most oesophageal speakers. The flow rate in laryngeal speakers during a comfortable sustained /a/ was 70-180 cc/sec.

Snidecor and Isshiki conclude from this that ..

> "good effective airflow with a high rate of flow is to be encouraged. The rate of flow for insufflation must be rapid and under complete voluntary control. The rates of outward flow may approximate as high as one third that of the normal speaker." p.242).

For various reasons, the results from the studies here reported are difficult to compare; the experimental conditions vary considerably, and the subjects are not always the same sex; indeed, in some studies, the sex of the subjects is not stated (e.g. Damste, 1956). In many studies, the researchers have not defined their criteria in determining "good" from "poor" oesophageal speakers.

2.4. E.M.G. Studies

Shipp (1980) studied eighteen male laryngectomees and recorded EMG activity from the inferior constrictor and cricopharyngeus muscles simultaneously during recordings of voicing on a single vowel sound using oesophageal speech.

During the inflation of the oesophagus, prior to phonation, there were bursts of muscle activity recorded from both muscle groups but poorer speakers seemed to show less control of differential muscle contraction than adequate talkers. No typical or model muscle patterns were found for subjects during the phonatory portion of the voicing task. Consistency of pattern within each subject was extremely high during a given procedure and on repeated procedures. Shipp concluded that ...

> "each laryngectomised talker adopts a phonatory method that is unique to him and consistent with his pre - operative anatomy and physiology."

The pattern variations obtained across speakers during phonation tasks were too numerous to classify.

Despite this somewhat surprising conclusion, it is evident that there is a pattern of pharyngo - oesophageal manoeuvres which need to be present before successful speech can result. In our recent studies (Cheesman et al, 1986; Perry and Edels, 1985; Perry et al, 1987) we have indicated that, by using videofluoroscopy during three different tasks - swallowing, attempted phonation, and air insufflation testing - it is possible to classify all oesophageal speakers thus far examined on a continuum of tonicity in the reconstructed pharyngo - oesophagus and that this has implications for their success or failure in the acquisition of oesophageal speech as well as their success or failure for T-E speech restoration.

What remained to be tested was whether these observed qualitative differences could be measured quantitatively. In order to do this, a combination of acoustic and aerodynamic measures needed to be taken during speech manoeuvres.

CHAPTER 3: METHODOLOGY

3.1 Introduction

In 1983, following the work of Blom and Singer in Indianapolis and their attempts at surgical reconstruction of voice after laryngectomy, it was decided to begin a similar programme of surgical voice restoration at Charing Cross Hospital, London.

There were many subjects available; mainly laryngectomees at the hospital who had failed to achieve satisfactory oesophageal voice and did not wish to use the artificial larynges (electronic communication aids, see 1.2.3.) which were available as alternative modes of communication. These seemed to provide a source of potential candidates for surgical speech restoration and, as the procedure was reversible, this was considered to be a feasible group of subjects to whom we could offer surgical voice restoration (Blom - Singer procedure).

If one compares laryngeal, oesophageal and tracheo-oesophageal speech (involving a surgically created fistula between the trachea and oesophagus), it is evident from Table 1 that there is a need for a competent P-E segment to vibrate and produce sound for shaping into speech. Therefore, it was judged necessary to devise a method of assessing the potential of the P-E segment to vibrate and produce sound prior to suggesting to the subject that a surgical procedure be undertaken.

The anatomy and function of the pharynx and upper oesophagus after laryngectomy can be assessed by using a combination of fluoroscopy video recording and static (spot) films.

Fluoroscopy allows the movement of the P-E segment to be viewed while observing the subject undertaking various speech manoeuvres. This can be stored on video tape as a permanent record of movement. The spot films were chosen to give a more anatomically detailed but static image of the P-E segment. The advantages of this combined system are that the dosage of X-rays is low, that the information is easily stored and reviewed and that the level of precision/detail is high. It was judged to be the system that would provide clearest and most appropriate information with which to test this investigator's hypothesis.

The examinations at Charing Cross Hospital were undertaken by a team consisting of a radiologist, a speech therapist and an otolaryngologist, all of whom observed both the fluoroscopy screen and the subject during the examination. The oral tract, pharynx, P-E segment and oesophagus were filmed in the lateral position during the acts of:-

swallowing
attempted phonation
attempted phonation during air insufflation.

All the procedures were recorded on videotape and at least two spot films (radiographs) were taken during each of these activities. In this way, one could later review the dynamics of swallowing (using the videotape) and the anatomical inter-relationship of the laryngectomee's oral, pharyngeal and oesophageal structures using the spot films, checking that the structures were in similar inter-relationship on both spot films during all three manoeuvres.

The radiological appearances and the quality of speech were discussed and noted in the course of the examination. The video recordings and static films were stored for later review. Individuals who were judged to be poor oesophageal speakers were classified in a way that is described (see 3.4.).

It is often difficult to obtain satisfactory lateral radiographs of the pharynx and upper oesophagus with standard fluoroscopy equipment, especially after laryngectomy as the lower part of the neck (where the tracheo-stoma is situated) is thinner and more radiolucent than the face and the upper part of the neck (Zaino, 1987). However, a few simple modifications were made to standard fluoroscopic equipment which made it possible to obtain diagnostic radiographs of the P-E segment in virtually all patients. These are described in section 3.2.

Baugh et al (1987) commented that..

> "The subjective nature of the (videofluoroscopic) assessment, the significant radiation exposure and the unproven efficacy of videofluoroscopy to predict the tonicity of the pharyngeal constrictors during oesophageal insufflation preclude it's use as a clinical method for predicting tracheo-oesophageal speech fluency at this time" (see 2.3.2.9).

The radiation dose during the whole examination was estimated at 5-10 cGy which, although being high for a diagnostic procedure, was much less than the therapeutic dose of 5,000 - 6,000 cGy already received by most British laryngectomees.

This is often not the case with American laryngectomees as, in America, radiotherapy is used far less pre-operatively as a curative procedure. Consequently, the "significant radiation exposure" to which Baugh et al referred is not such a major consideration in our subjects. The radiation dose was limited to the neck and thoracic inlet, so the bone marrow dose was low.

3.2 Equipment

A standard fluoroscopy unit (Siemens Sirescop 3) was used, with the table in the vertical position and the foot removed. Most subjects were examined whilst standing as their shoulders tended to drop slightly in this position, which was an advantage in viewing the soft tissues of the pharynx. Those unable to stand for the duration of the examination were seated.

The distance between the table and the image intensifier measured 45 cm, which allowed most subjects to be examined in the true lateral plane. Those with broad shoulders were placed with their shoulders slightly oblique, but the P-E segment and neck were always screened in the lateral position. The level of the tracheostoma was marked with a radio-opaque marker; initially this was a metal dessert spoon, taped to the para-stomal skin, but after experimentation, a radio-opaque letter /s/ was used which was able to mark the stoma position relative to the P-E segment.

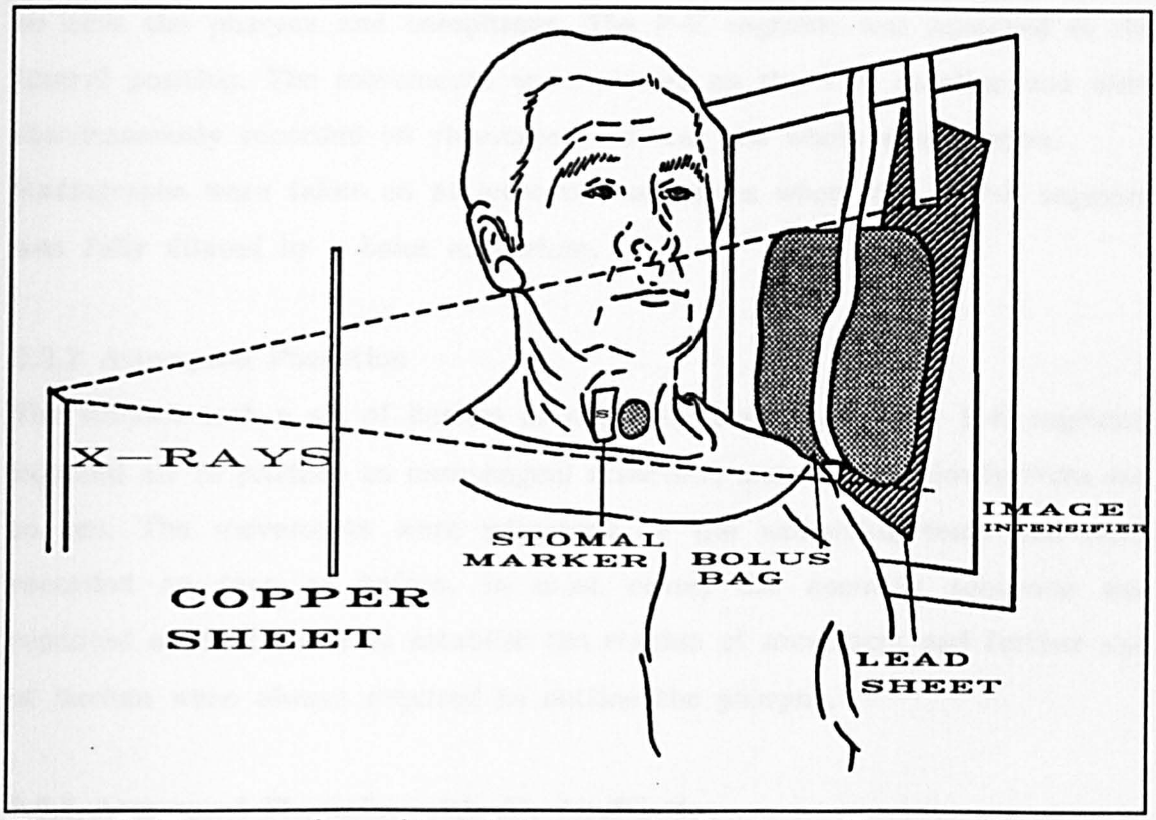


Figure 8. Diagram of a subject being examined using a Temporarily Modified Fluoroscopy Unit.

Three additions were made to the fluoroscopy unit before the examination. A copper sheet (0.50mm thick), measuring 30 x 30 cm., was attached to the table top to filter the X-ray beam; the field size reaching the image intensifier was reduced with a lead rubber sheet (0.5mm lead equivalent) which prevented flooding of the image intensifier with the non-attenuated part of the beam passing in front of the patient's neck; and a bolus bag containing polyvinyl acetate beads and powdered cellulose, was placed over the lower part of the neck to increase its radio-opacity.

These temporary modifications, shown in Fig.8, improved the fluoroscopy images and spot films and made it possible to obtain good diagnostic images of the P-E segment in all cases.

3.3 Examination

3.3.1 Swallowing

The subject took several swallows of barium sulphate suspension from a cup to coat the pharynx and oesophagus. The P-E segment was screened in the lateral position. The movements were viewed on the T.V. monitor and were simultaneously recorded on videotape - as was the whole examination. Radiographs were taken on at least two occasions when the P-E segment was fully dilated by a bolus of barium.

3.3.2 Attempted Phonation

The subject took a sip of barium to coat the mucosa of the P-E segment, injected air to produce an oesophageal reservoir, and counted slowly from one to ten. The movements were witnessed by the examining team and were recorded on tape as before. In most cases, the counting sequence was repeated once or twice to establish the rhythm of movement and further sips of barium were always required to outline the pharynx.

3.3.3 Attempted Phonation with Air Insufflation

A flexible tube (catheter) was passed through the subject's nose, via the pharynx, through the P-E segment (see Fig 9).

Figure 9. Air Insufflation Test to Show Position of Catheter.

Although some examiners (Blom, 1985; Baugh, 1987) have stressed the importance of the catheter tip being at a level of 27cm measured from the nares, as these investigators have not always used air insufflation under fluoroscopic conditions, it is more important to stress the tip should be under a vibrating P-E segment. This, in practice, varied from 20 cm to 27 cm from the nares in this series of subjects and, if present, was clearly visible using videofluoroscopy.

Originally, a radio-opaque guide wire was needed to assess the base of the catheter and its relationship to the P-E segment. Eventually, a specially modified 14 FG bronchographic rubber catheter was used and this could be readily seen on fluoroscopy as it had a radio-opaque tip and two side holes close to this. The tip was positioned below the level of the narrowest part of the P-E segment and the examiner insufflated the catheter to fill the oesophagus (Fig.9). The P-E segment was coated with a sip of barium and the subject counted from one to ten as before, but this time did not need to inject air him/herself.
The procedure was usually repeated with the catheter tip at different levels and the point of maximum vocalisation was noted. Radiographs were taken, usually two at each position of the catheter, to obtain a more detailed picture of the catheter tip, tracheostoma and P-E segment.

3.4 Classification

The classification system which was devised at Charing Cross Hospital for identifying failed oesophageal speakers was evolved from an initial study of fifty post - laryngectomy subjects; eight "good" and forty-two "poor" oesophageal speakers (Cheesman et al, op cit). The good speakers volunteered for the examination and their co-operation was essential as there was no adequate description of the radiological appearance of the P-E segment after laryngectomy, especially during air insufflation testing, in subjects with "good" oesophageal speech. All good speakers were at Level 1 scale of the Wepman scale of speech proficiency (Wepman et al, op cit. See 2.3.)

Poor oesophageal speakers (Wepman level 4 or more) were classified according to the radiological appearance of the P-E segment after laryngectomy during swallowing, attempted phonation and air insufflation. They were divided into four categories, namely:-

Hypotonic
Hypertonic
Spastic (spasm)
Stricture

on the basis of the criteria discussed below. The radiological appearances of good and poor oesophageal speakers are summarised in Table 4.

ACTIVITY	Barium Swallow	Phonation	Air Insufflation Test	Result
HYPOTONIC	Pharynx to oesophagus dilates. Fast motility.	No P-E segment visible. Weak voice.	Whispery, no phonation	SLIGHT POSITIVE
GOOD OESOPHAGEAL SPEAKER	Pharynx to oesophagus dilates. Fast motility.	One P-E segment visible. Good oesophageal air reservoir. Good voice.	Good voice - like own.	POSITIVE
HYPERTONIC	Slight narrowing in tract visible but dilates.	One tight or two P-E segments visible. Poor voice.	Better voice on gentle air blowing.	POSITIVE
SPASM	Slight narrowing in tract visible but dilates.	Very tight P-E segment. Oesophageal air reservoir good. Poor/no voice.	No voice achieved. Very tight P-E segments. Explosive release.	NEGATIVE
STRICTURE	Hold up to barium. Little/no dilation. Reports dysphagia. Slow motility.	Little/no air penetrates oesophagus.	No voice achieved. Oesophagus inflates but no P-E segment opening.	NEGATIVE

Table 4. Radiological Classification of Good and Poor Speakers.

3.4.1 Good Oesophageal Speakers

The characteristic radiological feature of this group was the presence of a narrowed or vibrating segment during phonation. The segment was up to 3 cms in length and the anterior and posterior walls were separated by a few millimetres during phonation. The P-E segment dilated fully during swallowing. There were sometimes slight irregularities of the anterior and posterior walls, but a striking radiological feature was often a "pseudo-epiglottis" which arose from the anterior wall of the pharynx at the base of the tongue. This feature was common to both good and poor oesophageal speakers and seemed to be of no functional significance (fig. 10). It was observed in approximately 85% of all subjects.

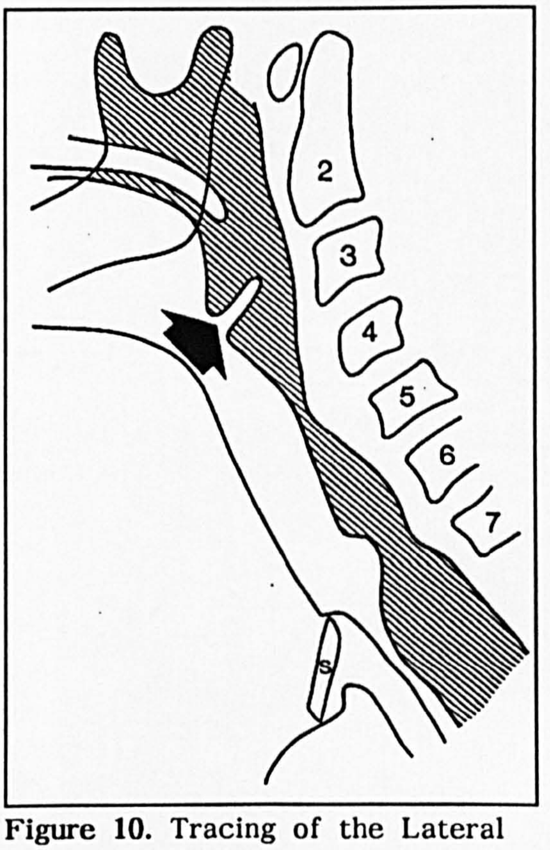


Figure 10. Tracing of the Lateral Radiograph of a good oesophageal speaker during phonation. The cervical vertebrae have been numbered 2-7 and the pseudo-epiglottis is arrowed. The stoma is marked with a radio-opaque marker /s/.

The P-E segment dilated widely on swallowing and, although there could be irregularities in the anterior wall (including a pseudo-epiglottis) there was no hold-up to the passage of barium which passed rapidly into the upper oesophagus (Fig. 11).

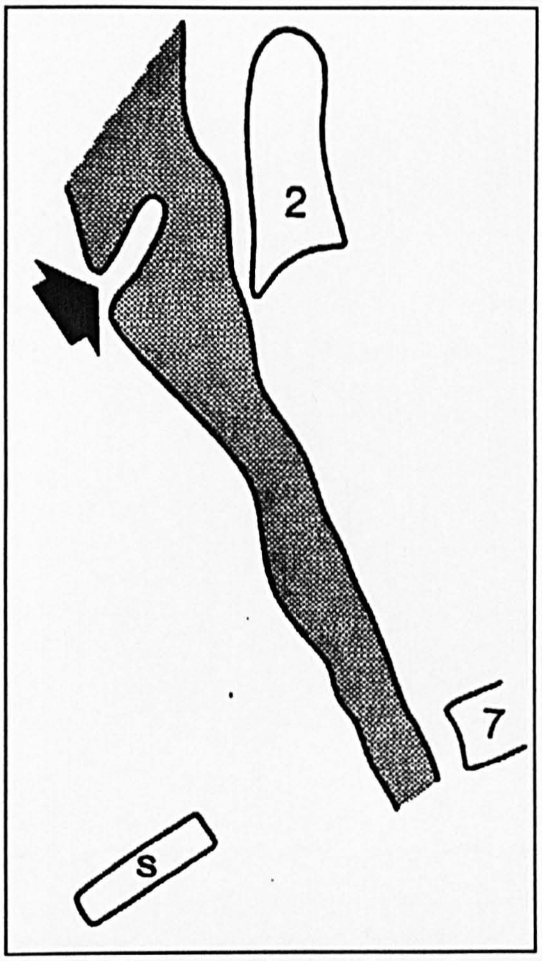


Figure 11. A good oesophageal speaker during barium swallow showing a pseudo-epiglottis (arrowed) & a widely dilated P-E segment.

3.4.2 Hypotonic

The characteristic feature of this group was that the anterior and posterior walls of the P-E segment remained widely separated during attempted phonation and air insufflation (fig 12).

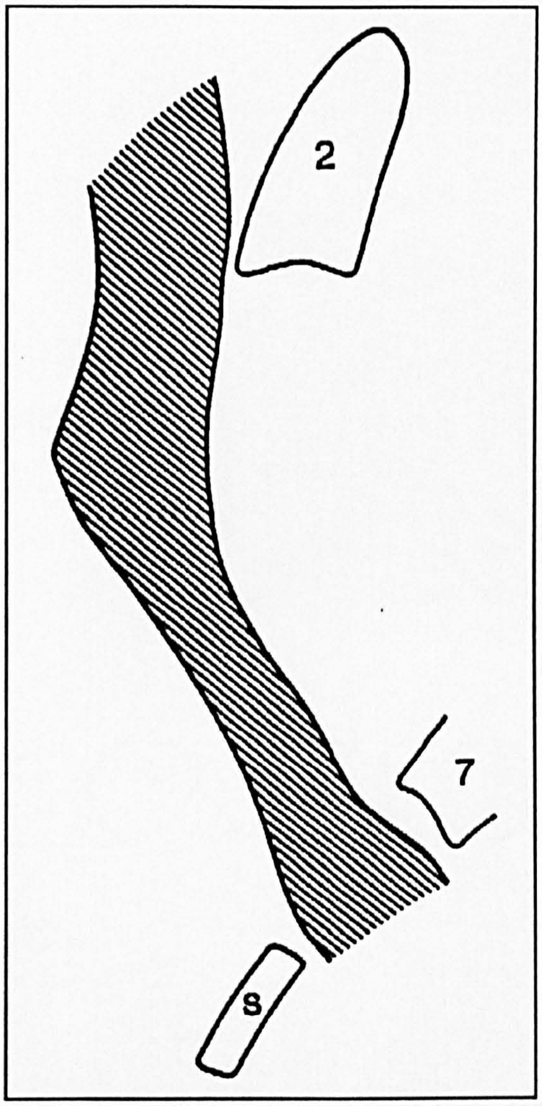


Figure 12. Attempted phonation in a subject classified as hypotonic. The P-E segment is widely dilated with no vibrating segment and no voice production.

There was no evidence of a narrowed or vibrating segment unless pressure was applied to the anterior part of the neck, usually with the subject's finger (i.e. digital pressure), when the anterior wall of the P-E segment was displaced posteriorly so that it contacted the posterior wall. This produced a vibrating segment and voicing occurred (Fig 13)

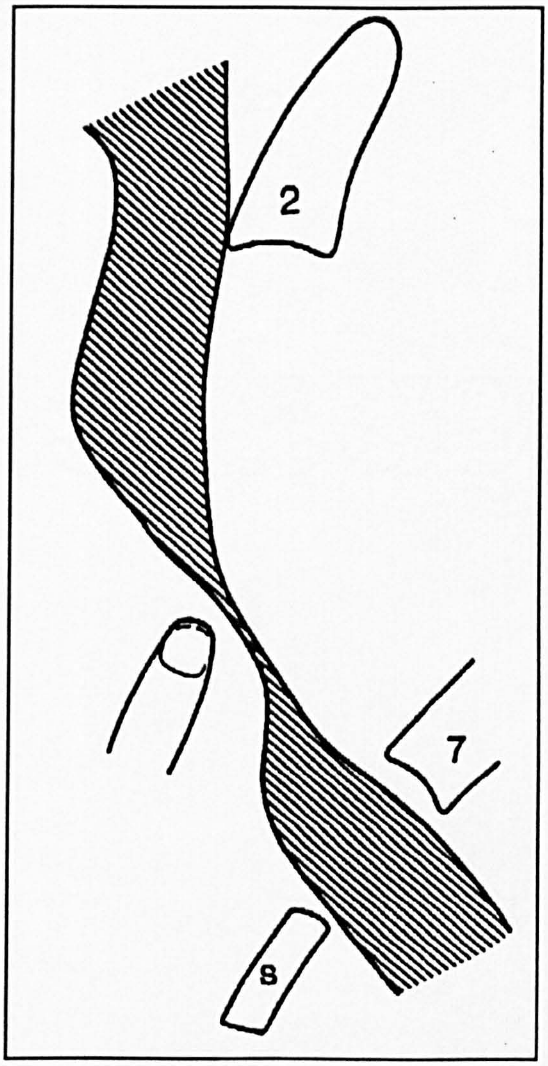


Figure 13. Digital pressure applied during attempted phonation in the same subject as shown in fig 12. A vibrating segment is thus visible & voicing occurs.

3.4.3 Hypertonic

The characteristic radiological feature of this group was the presence of one or two severely narrowed segments during attempted phonation which persisted during air insufflation. If two segments were present, one was usually narrower than the other and seemed to be functionally more significant (Figs. 14, 15)

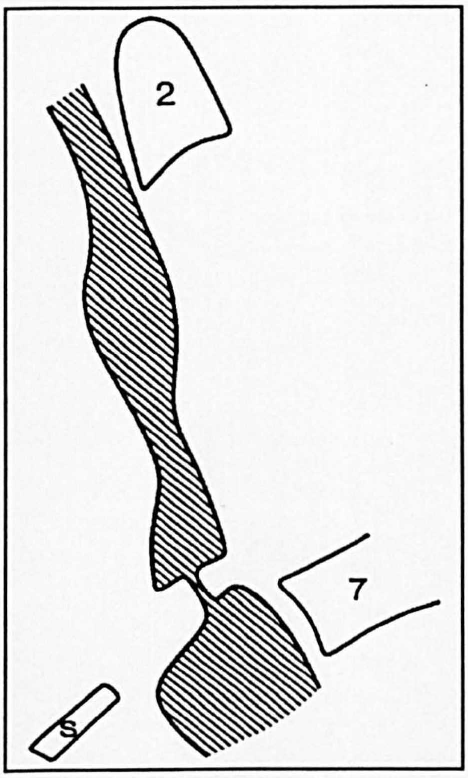


Figure 14. Attempted phonation in a subject classified as hypertonic showing a severely narrowed segment. The voice was poor.

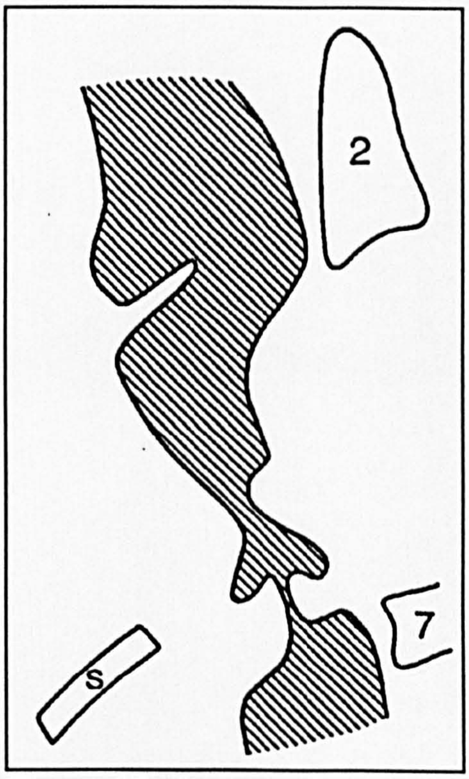


Figure 15. Attempted phonation in a hypertonic subject with two narrowed segments and a poor voice.

Gentle air insufflation through the catheter (by introducing the air more slowly) produced slight dilation of the narrowed segment(s) and resulted in phonation if the tip of the catheter was just below the narrowed segment. Vigorous air insufflation produced spasm and no voice.

The outline of the P-E segment during swallowing was variable, but there was never any hold-up to the passage of barium. The anterior wall may be irregular and a pseudo-epiglottis was often present. The posterior wall may demonstrate a filling defect due to pharyngo-oesophageal spasm on some swallows (Fig 16). This means that barium is seen to lodge in the pharynx despite repeated swallows. The diagnosis of muscle spasm should not be made unless the P-E segment can be seen to dilate fully on at least one occasion during swallowing. It can be difficult to differentiate between persistent muscle spasm and stricture, but if the subject can be persuaded to take multiple swallows of barium, the narrowed segment will invariably dilate at least once and this can be recorded on videotape. It is then possible to investigate this phenomenon.

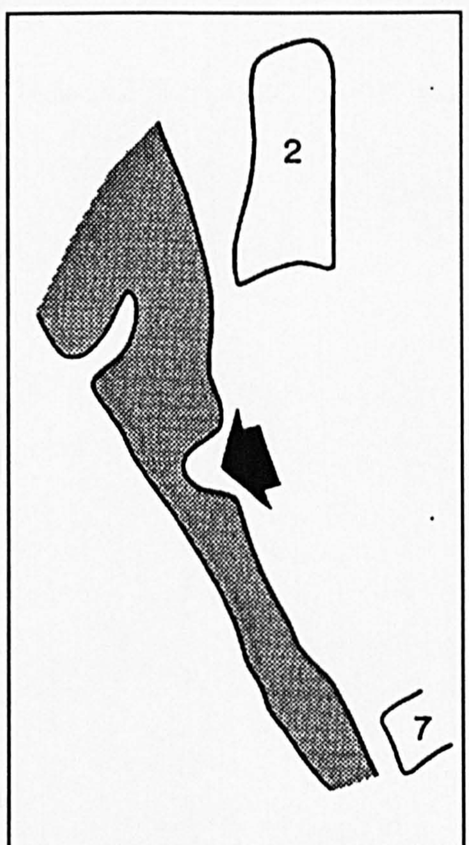


Figure 16. Barium swallow of a hypertonic subject showing a small posterior filling defect (arrowed) due to P-E Spasm. A pseudo-epiglottis is present.

3.4.4 Spastic (Spasm)

The features of the hypertonic group were all present but in a more extreme form and it was sometimes difficult to distinguish between hypertonicity and spasm in an individual subject, as the more severely affected subjects in the hypertonic group merged with the more mildly-affected subjects in the spasm group.

The main feature which distinguished subjects classified as being in the spasm group was that the narrowed segment(s) which were visible during attempted phonation (Fig 17) become spastic during air insufflation, occluding the P-E segment and producing marked oesophageal and gastric distention (fig 18). The air was released in sudden explosive bursts and there was no useable voice, despite the air being introduced very slowly through the catheter.

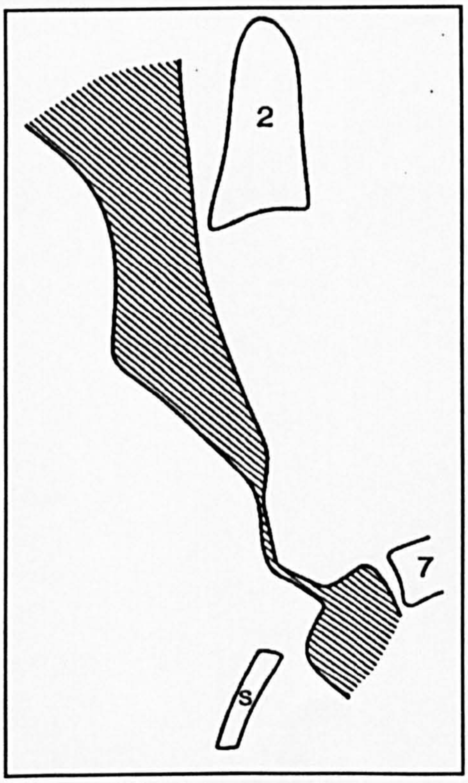


Figure 17. Attempted phonation in a patient with spasm showing a long spastic segment. Voice was not produced.

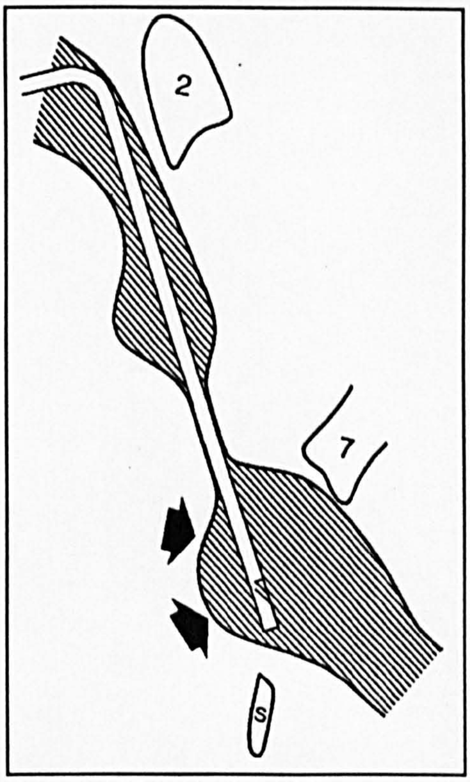


Figure 18. Air insufflation in subject shown in Fig 17. Occlusion of the P-E segment can be seen with marked dilation of the upper oesophagus (arrowed).

The dilated upper oesophagus could be clearly seen on fluoroscopy and the oesophageal contractions and explosive releases of air could be seen and heard by the examining team. Spot films taken during air insufflation showed marked oesophageal dilation and the narrowed segment(s) were narrower and longer than in the hypertonic group (Fig 16).

3.4.5 Stricture

The distinguishing feature of this group showed during the Barium swallow where there was constant narrowing of a short segment with a hold-up to the passage of Barium. The narrowed segment had the same outline on swallowing, phonation and air insufflation and never dilated. There was no vibrating segment and no voice. A stricture is caused by scar tissue formation, and is usually the effect of extensive radiotherapy and surgery resulting in formation of fibrous tissue (Fig 19).

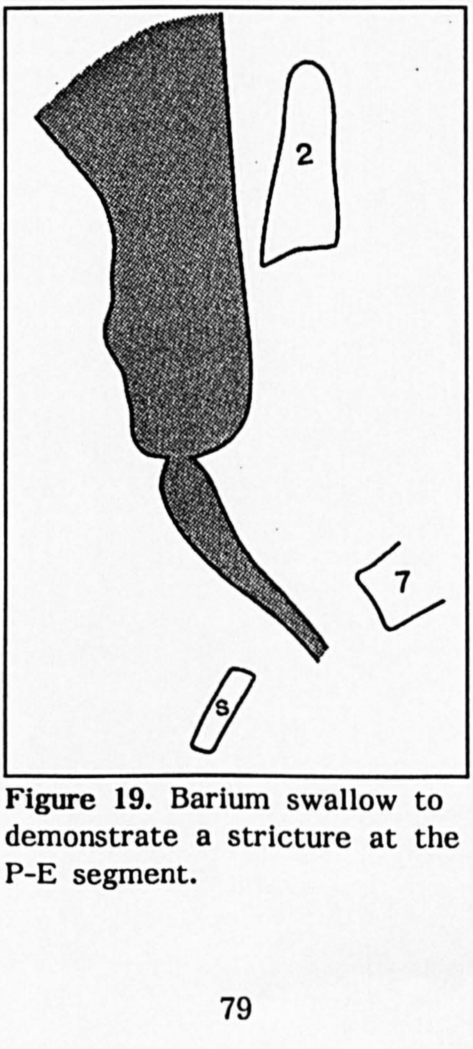


Figure 19. Barium swallow to demonstrate a stricture at the P-E segment.

Subjects in this group all suffered from dysphagia, which was worse for solids than liquids.

The advantage of this classification is that it suggests anatomical/ physiological causes of failure in poor oesophageal speakers. This classification, albeit subjective, has the merit of attempting to refine the interpretation of air insufflation testing from an "all or nothing" phenomenon into one which takes account of the multifaceted nature of the responses.
It was recognised however, that there was a need to quantify the reactions described above and thus objective measurements were developed as part of this study.

3.5 Objective Measurements

A survey of the literature (see section 2.0) provided valuable information on the parameters to study which might be relevant to successful oesophageal speech production. Previous studies examined the following aspects of oesophageal speech production: frequency, intensity and duration of speech produced (i.e. acoustic measures); pressure and flow generated while producing that speech (i.e. aerodynamic measures); and EMG studies of muscle groups involved in oesophageal speech production. In most published studies of alaryngeal speech only one or two of the variables listed were studied and these have been discussed in the literature review previously (2.0). A number of studies referred to the inter-relationship between acoustic variables and aerodynamic changes and it was thus expected that a study including both acoustic and aerodynamic parameters, simultaneously measured, could be expected to provide critically important data for better understanding of oesophageal speech production.

3.6 Pilot Study

3.6.1 Subjects

Six subjects classified as "hypotonic" on our system and six subjects classified as in "spasm" were fully investigated in the pilot study and the results analysed as a way of validating the two extremes of the hypothesised spectrum. These subjects were compared with one good oesophageal speaker and all thirteen were investigated at the monthly videofluoroscopy clinic at Charing Cross Hospital, in the diagnostic X-ray unit. All subjects in the study group had been referred as possible surgical speech restoration candidates. The specific procedures were explained to the subjects and an informed consent obtained. Prior to investigation, each subject completed an intake data sheet (Appendix 1) which provided background case history details as well as medical/ surgical information.

3.6.2 Videofluoroscopy

Videofluoroscopy was used to assess the size, shape and co-ordination of the pharynx, P-E segment and oesophagus dynamically as well as to ascertain the correct placement of the radio-opaque catheter prior to air insufflation testing as previously described. It was then possible to take simultaneous acoustic and aerodynamic measures during oesophageal speech production.

3.6.3 Airflow (In)

The first variable to be controlled in this pilot study was the airflow for the insufflation test. This was performed using a compressed air source at flow rates of 0.5 and 5.0 litres per minute as measured by a Thorpe flow meter at the head of the cylinder. Plastic connecting tubing (14 FG in diameter) via a Y-connector was used to connect the compressed air source to the modified 14FG bronchographic rubber catheter for insufflation via the S.E. laboratory 4.86 differential pressure transducer.

3.6.4 Pressure

The pressure transducer was calibrated using a manometer for the ranges 80 - 200 mmHg and the calibration signal stored on channel 1 of the 4 channel recorder. The pressure below the P-E segment (i.e. intra-oesophageal pressure) and acoustic measures were continuously recorded at the time of opening of the P-E segment (i.e. at initial onset to voicing) and during sustained phonation of the vowel /a/ . This was recorded as ranges and stored on the 4 - channel recorder. To obtain pressure measures during connected speech, subjects were asked to count from one to ten and again simultaneous pressure and acoustic recordings were taken.
The Gould 8188-440-09 4 channel thermal (hot wire) chart recorder was used for hard copy of the pressure signal. The speed of paper used for recording from the chart level recorder was 10cm/sec. The pressure was recorded with the S.E. transducer and the signal then processed using a Gould pressure amplifier. The signal strength could be observed on an oscilloscope and permanently stored on channel 1 of the Racal 4 channel tape recorder using frequency modulation (FM) instrumentation tape.

3.6.5 Acoustic Measures

The acoustic recordings were made using a Bruel and Kjaer microphone at a distance of 10 cm from the subject's mouth at a signal gain from 3 - 4. The subject was instructed to keep his/her head still while speaking and the microphone/mouth distance was re-measured at the end of the recording. These recordings presented particular problems as the frequency of oesophageal speech recorded was in the 80 - 130 Hz waveband. Thus it was necessary to remove all background low-frequency noise from the recordings, otherwise it was difficult to detect signal from noise when analysing the tape.

The recording was stored on channel 3 of the Racal 4 - channel recorder for later analysis. From the signal, it was anticipated that frequency, intensity and duration of phonation of the sustained vowel sound could be measured. To provide a reference level for later analysis, an 80 dB pure tone was recorded on the audiotape, 4v pk - pk, at 80 Hz.

3.6.6 Airflow (Out)

Finally, airflow out from the mouth during the production of the vowel sound /a/ was recorded on a separate channel of the 4 - channel recorder and printed out on the chart paper. In order to do this, the subject used a mouth piece for recording and a nose clip to avoid nasal emission of air. The mouth piece was connected by plastic tubing to a flow meter head and thence to the chart recorder. A separate connection took the flow signal from the flow meter to the 4 - channel recorder for permanent storage of the signal.

3.7 Investigative Procedure

The sequence of events was as follows:-

1. The barium swallow was recorded on videotape as previously described.
2. Attempted phonation by the subject her/himself - using oesophageal speech - was recorded on videotape and spot films taken.
3. The modified bronchographic catheter was placed in situ via the nose and oropharynx into the oesophagus just under the P-E segment. The siting was checked and recorded using videofluoroscopy prior to air insufflation testing.
4. The catheter was connected to the recording apparatus and the subject was instructed to open his/her mouth "as if you are going to say /a/", but not to try for sound, simply to let the air come back and not to resist it while phonating.

Flow rates from 0.5.to 5.0 ltr/min were used to insufflate the oesophagus from the compressed air source.

If no sound was heard at 0.5 ltr/min, the flow was increased to 1.0 ltr/min and checked on the fluoroscopy screen to ascertain that the air was indeed inflating the oesophageal walls and that the catheter tip, for example, had not become blocked with Barium.

5. The microphone was positioned in its stand, 10cm from the subject's mouth. Pressure and acoustic signals were recorded during three attempts at phonating on the vowel /a/ during insufflation of the oesophagus.
6. The subject was then instructed to try and mime counting to ten when the compressed air was felt at the back of his/her throat. Repeat recordings of pressure and sound were taken, again during the three attempts at counting. During both these sets of recordings, all background low frequency noise was removed.
7. The subject was instructed to try and produce the /a/ sound again and it was explained that, this time, the flow of air from his/her mouth was to be recorded, so a mouthpiece needed to be worn to record this and a nose clip, to avoid any escape of air via the nose. Pressure of air was again recorded at this time to check with the earlier recordings for accuracy.

Again 0.5 - 5.0 ltr/min airflow was put into the oesophagus via the catheter. The flow and pressure were recorded on the chart level recorder, and the signals stored on the 4 - channel recorder as previously described. With this apparatus, it was possible to record pressure of air needed for onset and duration of sound to be produced; the frequency, intensity and duration of sound produced and the flow of air needed to sustain the phonation. In recording these measures, simultaneously as far as possible, it was hoped to ascertain whether there were quantitative differences among the groups previously described qualitatively.

3.7.1 Parameters to be Measured

During the investigative procedure 3.7. above, the following aerodynamic measures were taken for each subject:-

1. The opening and sustaining pressure of the P-E segment during vocalisation.
2. The airflow required for speech.

The following additional acoustic measures were calculated:

3. The fundamental frequency (of sustained vowel sound) and ranges of frequencies used (in counting tasks).
4. The length of phonation (of sustained vowel sound). As stated in 3.7., a known flow of between 0.5 and 5.0. litres per minute was passed whilst the aerodynamic and acoustic measures were taken.

These results are shown in Tables 5 and 6.

OPENING PRESSURE OF THE P-E SEGMENT RECORDED AT ONSET TO PHONATION

GROUP 1 - HYPOTONIC	GROUP 2 - SPASM
Mean (mmHG)	Mean (mmHg)
S1 = 30.30	S7 = 93.00
S2 = 12.30	S8 = 78.30
S4 = 29.30	S9 = 56.00
S5 = 28.30	S10 = 86.00
S6 = 30.30	S11 = 91.70
Range 12.30 - 32.60 mmHg	Range 56.00 - 98.00 mmHg
Mean = 21.70 mmHg	Mean = 83.30 mmHg

NB - This compares with the Tonic speaker from whom a recording of 35.00 mmHg was obtained.

T - TEST RESULTS

Variable		Group 1	Group 2
Mean		27.133	83.833
S.D.		7.406	15.209
T. statistic	=	8.210	
Degrees of freedom	=	10	
Significance:		0.0005 ($p < 0.0005$)	

Table 5. Results of Pilot Study: Pressure.

SUSTAINED PHONATION ON VOWEL /a/

GROUP 1 - HYPOTONIC	GROUP 2 - SPASM
Mean (sec)	Mean (sec)
S1 = 5.90	S7 = 1.40
S2 = 4.90	S8 = 0.65
S4 = 15.80	S9 = 2.80
S5 = 9.20	S10 = 0.80
S6 = 4.80	S11 = 3.20
Range 4.80 - 15.80 sec	Range 0.25 - 3.20 sec
Mean = 21.70 mmHg	Mean = 83.30 mmHg

NB - This compares with the Tonic speaker who sustained /a/ for 6.6 seconds.

T - TEST RESULTS

Variable		Group 1	Group 2
Mean		8.783	1.517
S.D.		4.465	1.214
T. statistic	=	3.847	
Degrees of freedom	=	10	
Significance:		0.003 ($p < 0.0005$)	

Table 6. Results of Pilot Study: Duration of Phonation.

3.8 Outcome of Pilot Study & Revised Experimental Procedures.

Following a pilot study from October 1986 to February 1987 (i.e. five data collection clinics), the recording apparatus was refined and modified and the experimental procedure was revised to make best use of available resources.

Additionally, data analysis proved to be more problematic than had been anticipated.

Methodologically, problems arose in the following areas:

3.8.1 Acoustic Recordings.

Background noise in the X-ray room was a major problem during the acoustic recordings as the interference "hum" from the air conditioning unit had frequency components (50 - 100Hz) close to the fundamental frequency of oesophageal speech (approx 80 Hz). Therefore it was difficult to ascertain acoustic signal from background noise and the deployment of narrow band filtering would have been of limited use. The air conditioning noise had to be removed by ensuring that each recording was taken with the air conditioning and radiographic machinery turned off.
Acoustic recordings would have been more reliable if it had been possible to use a mask microphone to filter out background noise, but this was not available. In fact, once the obvious background noise had been removed, the high quality microphone ensured that good acoustic recordings were obtained and stored on channel 3 of the 4 channel recorder.

It was originally anticipated, following the success of recordings in the pilot study, that all the 68 subjects who were finally recruited for the full study would have recordings taken when phonating on the vowel /a/ during air insufflation testing. However, in reality, in the final study there were problems in that three subjects in the "spasm" group could not achieve any sound at all, despite recordings of intra - oesophageal pressure of 60 mmHg and more and the subjects were complaining of gastric distention and feared vomiting so that the insufflation had to be abandoned. A further one subject in this group could only phonate once in three attempts. Hence, in Table 7 (Appendix III), there are missing recordings for the four subjects in Group 4 (spasm), but a total of 193 sound samples could be obtained.

3.8.2 Pressure Recordings.

The slow response time of the pressure transducer / airline system caused a time delay between the acoustic and pressure recordings, making it difficult to ascertain the exact pressure at which vocalisation commenced. After discussion with Weinberg (personal communication, Groningen,1987), this problem was overcome by reducing the air input flow to a constant 0.5 ltr/min., thus reducing the rate of pressure build up (dp / dt). The pressure was recorded on channel 1 of the 4 channel magnetic tape recorder and on the paper output from the chart recorder.

Another potential problem with the apparatus was to ensure that air in the tube to the patient did not compress so that a sharp change in input would have been slowed down (i.e. a poor high frequency response). This was a further reason for using videofluoroscopy to assess the air/line response.

3.8.3 Flow recordings.

A pneumotachograph, attached to a mouthpiece, monitored the air volume / flow rate leaving the mouth during speech for each subject recorded. Additionally, a nose clip prevented nasal emission of airflow. This presented a problem, as most subjects intensely disliked using the mouthpiece, which consisted of a large rubber flange fitting around the gums and was uncomfortable to wear. The flow was recorded and stored on channel 4 of the 4 channel tape recorder. In examining the results from the subjects in the pilot study, the flow rates were not significantly different. It was difficult for the subjects to retain a suitable intra - oral seal for the recordings to be accurate and they found it difficult to sustain a vowel sound whilst retaining the mouthpiece in position.

As it was difficult to overcome these technical problems, and the initial results did not reveal significant differences across the thirteen subjects examined, it was decided to abandon this aspect of the recordings after the pilot study was complete. Had a mouth mask been available with a flow meter embedded, then flow would have been a useful parameter to include.

The revised apparatus design is shown in Fig.20.

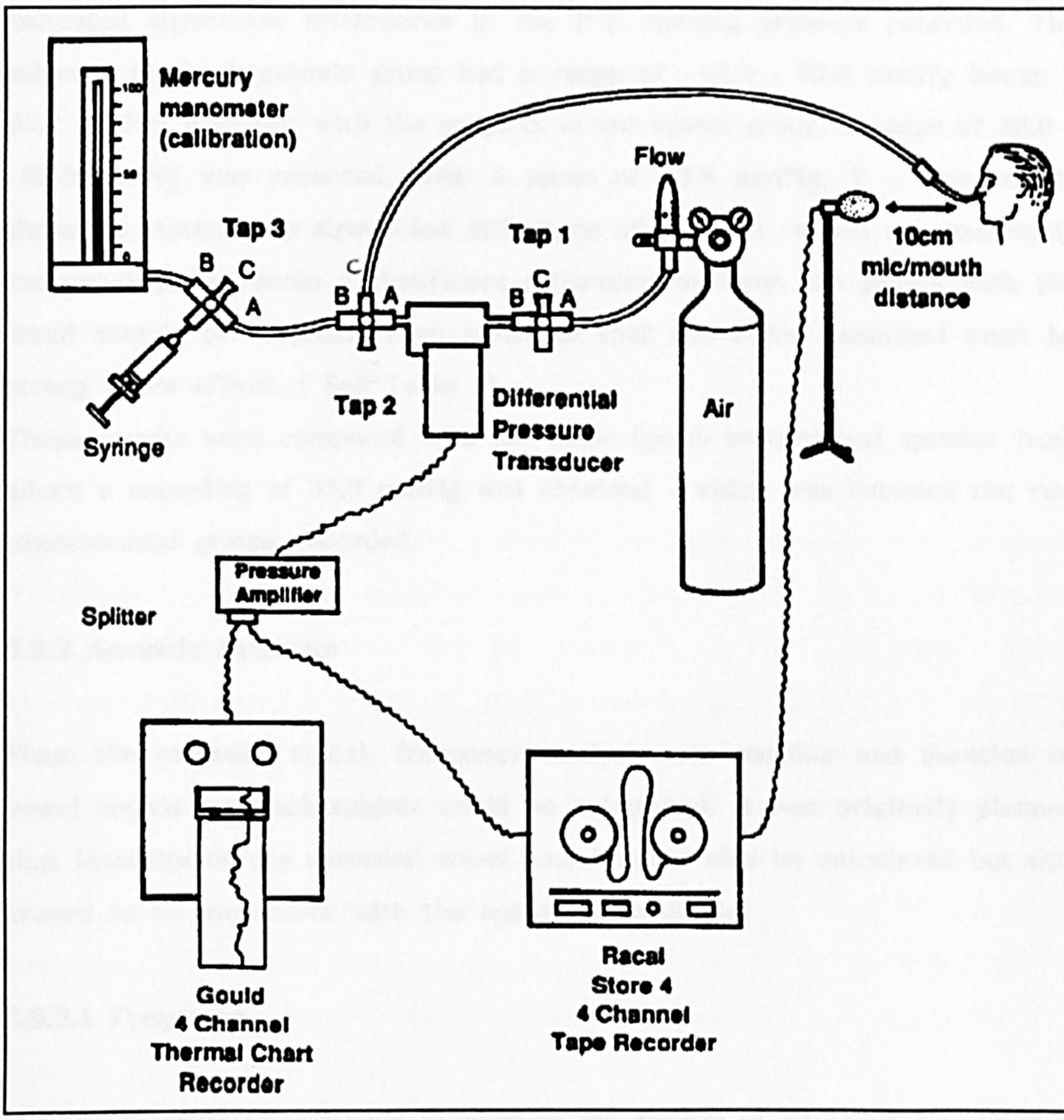


Figure 20. Apparatus for measurement of pressure & for acoustic recordings.

3.9 Methodology for Processing Data & Recording Results

The results of the pilot study revealed a number of problems in data processing.

3.9.1 Pressure

This was recorded for each subject as described in 3.6.3. In analysing six subjects classified (from videofluoroscopy) as being in the spasm group and six subjects in the hypotonic group, the results from the chart recordings indicated significant differences in the P-E opening pressure recorded. The subjects in the hypotonic group had a range of 12.3 - 32.6 mmHg (mean = 21.7 mmHg) whereas, with the subjects in the spasm group, a range of 56.0 - 89.0 mmHg was recorded, with a mean of 83.8 mmHg. T - test results showed a statistically significant difference of $p < 0.01$ which was extremely encouraging. To obtain a significant difference between the groups with the small sample of subjects taken indicates that the event examined must be strong in its effect. (See Table 5)

These results were compared with the tonic (good) oesophageal speaker from whom a recording of 35.0 mmHg was obtained - which was between the two experimental groups recorded.

3.9.2 Acoustic Measures

From the recorded signal, frequency analysis was possible and duration of vowel sounds for each subject could be calculated. It was originally planned that intensity of the recorded vowel sound would also be calculated but this proved to be impossible with the apparatus available.

3.9.2.1 Frequency

It proved extremely difficult to analyse the frequency of sound produced as the noise emitted by subjects is complex, aperiodic sound, with no true "tone." This problem was noted by Damste in 1958 in his analysis of oesophageal speech (see 2.3.1.1). The accumulation of mucous above the P-E segment means that air is forced in a highly irregular way from the oesophagus, via the P-E segment, through varying layers of secretions. Thus, noise is produced, rather than a clear tone. This observation is even more pronounced when considering the sound produced from air insufflation as the catheter in situ passes through the vibrating (P-E) segment so one might expect that this further changes the vibratory characteristics of the segment. However, air insufflation remains the only non - operative way of assessing the P-E segment characteristics.

An informal discussion with experienced speech therapists supported the clinical observation that the two groups could be distinguished on the basis of pitch differences. Without exception, the recordings of hypotonic subjects were consistently rated as being lower in pitch than those from subjects in the spasm group. No attempt was made to conduct an experimental analysis of these observations. Instead, machinery was investigated as a means of objectively quantifying these perceived pitch differences in terms of frequency analysis.

Initial analysis using Visispeech I proved unreliable as there was no clear fundamental visible; the frequency was not measurable at low frequencies (approx. 80 Hz) with the software available in this programme. Using spectrographic analysis proved to be more promising. Figs.21 - 24 show clear pattern differences between the subjects in the two groups analysed. Although low frequencies were present in both groups, higher frequencies (sampling at 80 - 8000 Hz) were more evident in the spasm group. Additionally, there were formants present in the good oesophageal speaker that were in the same frequency band widths expected for laryngeal phonation of the vowel /a/ as given by Baken, R.J.(1987). These formants could also be seen in some of the subjects in the spasm group.
This was perhaps surprising, given the fact that the vocal tract above the level of the glottis has changed characteristics following laryngectomy due to the surgery performed. One might expect the pharynx to have changed its formant characteristics due to the changed anatomy after surgery (Cheesman, 1986).

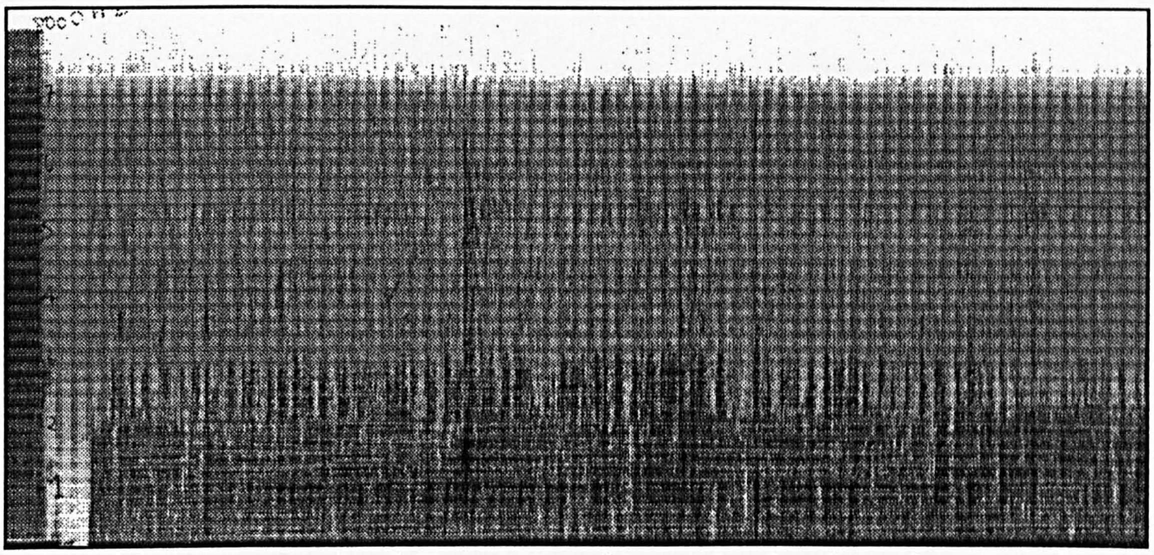


Figure 21. Spectrograph of Hypotonic subject, showing long duration of sound & lack of clear formants.

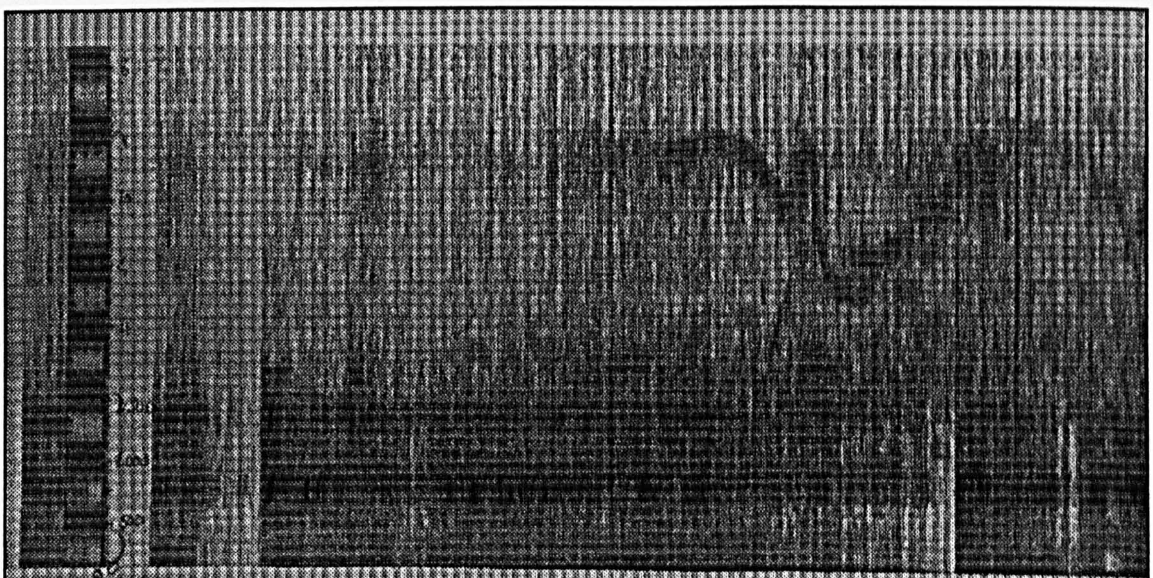
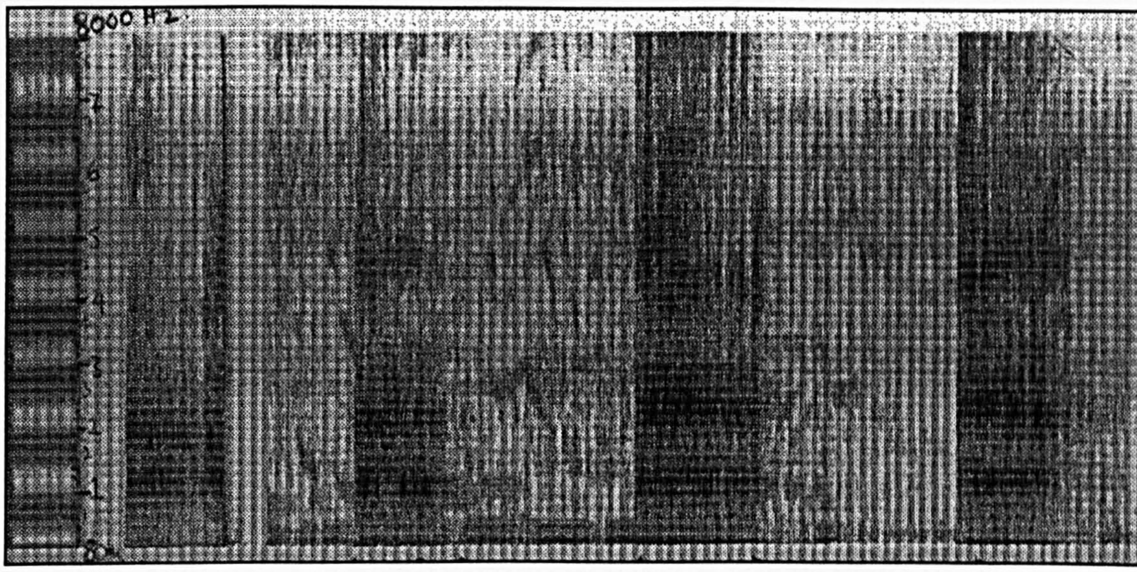
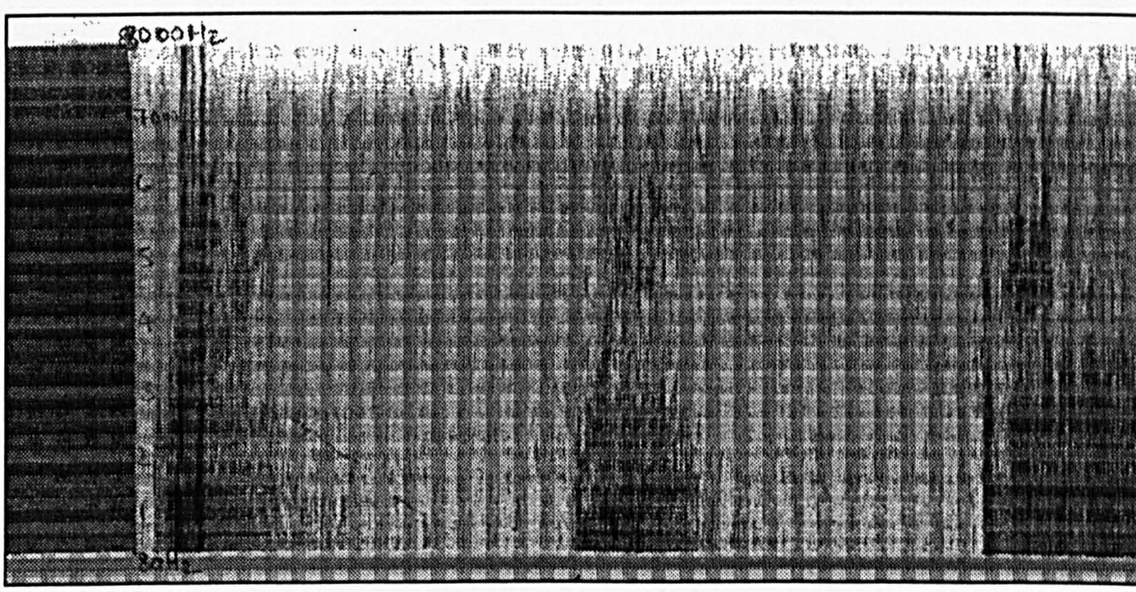


Figure 22. Spectrograph of Hypotonic subject showing the long duration of sound & lack of clear formants.



-- a-- -- a -- -- a -- -- a --

Figure 23. Spectrograph of Spasm subject. The formants can be clearly seen. The length of sound is shorter than Figs 21/22.



-- a -- -- a -- -- a --

Figure 24. Spectrograph of a Spasm subject. The formants can be clearly seen. The length of sound is shorter than Figs 21/22.

There was no clear fundamental frequency evident in any subject examined; this was probably due to the nature of the vibratory (P-E) segment which does not produce periodic sound. Rather, it was aperiodic sound that was produced with overlying sound from accumulated saliva etc. in the pharynx. The bursts of sound produced by the subjects in the spasm group were clearly delineated from the constant, low frequency sound produced by the subjects in the hypotonic group. (Figs. 21, 22 c.f.23, 24).

These findings indicated that further, more detailed, analysis of the sound signals produced, using a computer- based system, might produce valuable data about the differences in sound quality generated by the subjects in the two examined groups.

3.9.2.2 Fourier Analysis

Due to the complexity of the sounds generated, it was decided to use Fourier Analysis to determine the components of the wave forms. This is a dissection technique whereby a complex waveform is split into its component simple harmonic vibrations. Any steady state waveform can be generated by the addition of a series of sine and cosine waves of different frequencies and magnitude. The converse is also true, that any complex waveform can be analysed in terms of its frequency, amplitude and phase components. In practice, the time axis (time domain) is divided into equal points and the amplitude at each point is determined; number pairs are thus obtained i.e. analogue to digital conversion. The waveform is now in digital form and can be operated on by a computer algorithm called the Fast Fourier Transform (FFT) (Cooley and Turkey, 1965). This works on the principle that when a signal is multiplied by a sine or cosine wave, a cross product is obtained which has a finite value if the signal contains components at the frequency of the sine or cosine wave by which it has been multiplied. If it does not contain such components, the cross product is zero. The signal is sequentially operated on in this way, starting with the fundamental frequency and working through the harmonics, the values of each cross product being stored. This enables the coefficients of the Fourier series to be determined and, from these, the amplitude and phase of the harmonics. Fourier analysis does not add any new information, but converts that present in the time domain to the frequency domain.

Baken, R.J. (1987) described the principles of the Fast Fourier Transform (Cooley and Turkey, op cit) which makes it possible to implement a purely mathematical analysis even on relatively small computer systems. It is a three dimensional plot of the change in spectral content of the sound, over a period in time. In other words, the frequencies that are present in a sound and how their levels change with time.

After examining various possible software systems, it was decided to use the LYNEX system for frequency analysis. This is a 16 bit stereo sampling instrument, and studio post - production unit with a built - in 8 track digital mixer.

Although not designed for this project, it was felt that the software could potentially be used for the analysis of the sound signals recorded from these subjects as it could sample and store sounds and offered FFT analysis as part of the software programme. For a full explanation of the system see Appendix 2.

In its original form the sampling for the FFT was 0 - 25 KHz. However, it became obvious in sampling the recordings from the thirteen subjects in the pilot study, that the FFT plot was skewed towards the lower frequencies, 0 - 3 KHz with very little/no sound being visible in the 3 - 25 KHz frequencies. Consequently, the manufacturers of the Lynex agreed to re-write the software, producing a 0-6 KHz scale to utilise the whole screen. i.e. the formants in the 0- 6 KHz frequency band could be sampled with more accuracy, ignoring the frequencies above this level. Once this was complete, the results of frequency analysis were more encouraging.

3.9.2.3 Duration.

The Lynex programme previously described was used to accurately analyse the duration of sound samples of the vowel /a/ obtained from each subject. See Fig 25 for the apparatus used.

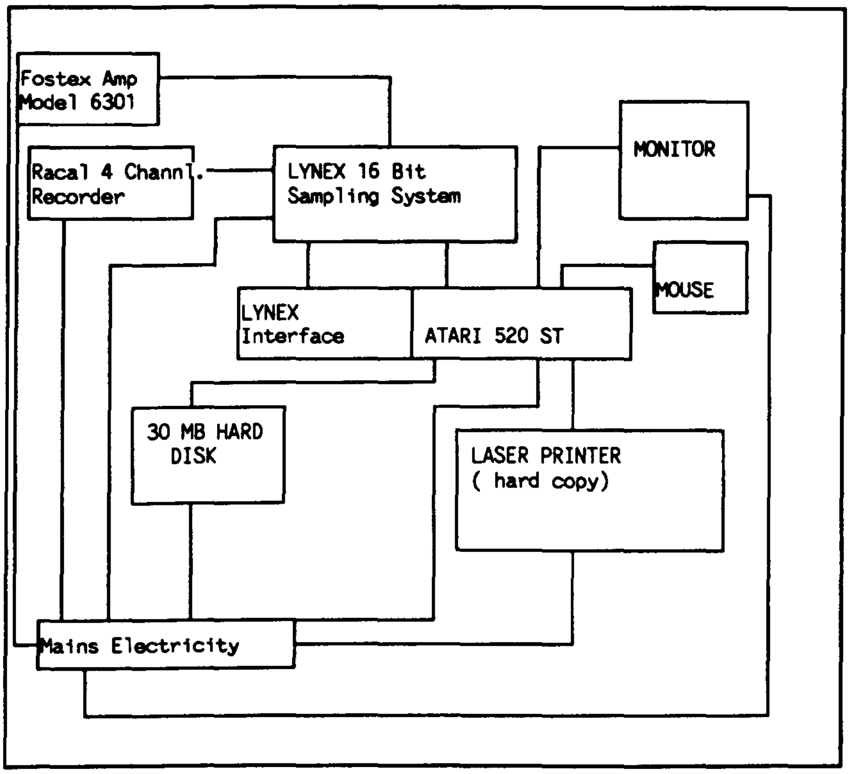


Figure 25. Apparatus for analysis of stored acoustic signal.

By setting the sample frequency to 12KHz, a maximum of 43.2 seconds of sound could be viewed on the screen. In using M1 and M2 cursors to mark the beginning and the end of the sound occurring on the scope trace, the length of phonation could be accurately measured. This could be confirmed by replaying the sample through the amplifier to ensure that the viewed oscilloscope trace is, indeed, a vowel sound produced rather than, for example, a non - speech sound e.g. burp. For each subject, three samples of /a/ were analysed in this way and documented.

This system was particularly useful in analysing samples from subjects in the spasm group as the duration of sound was often less than 0.5 sec and thus impossible to measure by auditory means alone.

Each sound sample of /a/ was stored on floppy disks, which meant six /a/ samples per disk i.e. three samples of two subjects. The disks were catalogued "Spasm 01; 01A; 01B" etc and thus could be recalled easily (Fig 26) from the memory. A 30 MB hard disk was used to back up the memory although this, in fact, became full during the main project before the 193 samples could be stored. Thus, 193 samples were stored on the floppy disks, but only 183 samples on the hard disk. This illustrates one of the minor problems with Lynex -it used a large quantity of memory to undertake this project where the sounds sampled had to be stored.

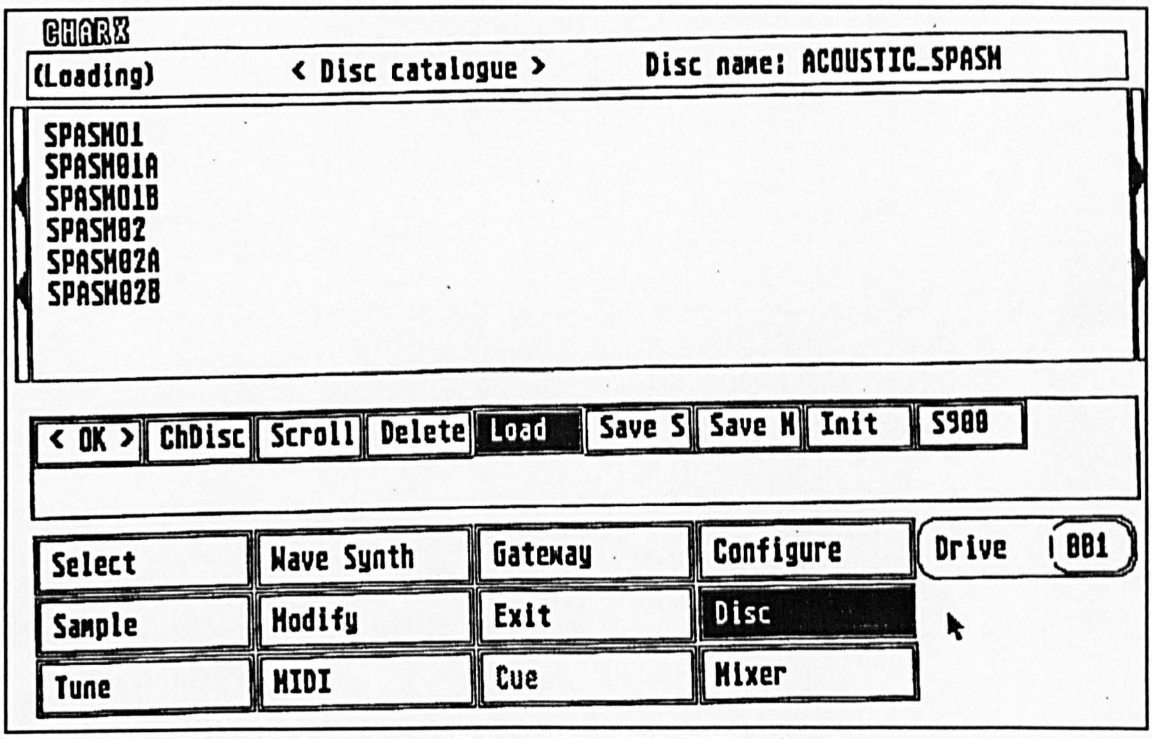


Figure 26. Screen illustrating Spasm 01, 01A,01B selection.

There were statistically significant differences in the two groups in the pilot study when comparing duration of phonation on a vowel sound. The hypotonic subjects were able to sustain phonation for a mean of 8.78 sec with a range of 4.8 - 15.8 sec, whereas the spasm group had a mean length of phonation of 1.5 sec with a range of 0.25 sec - 3.2 sec. A student's T - test was used to test for significance. This produced a statistically significant difference across the groups with a t of 0.003 ($p < 0.01$). These readings were compared with the tonic (good) speaker who could sustain /a/ for 6.5 sec. - which was higher than the spasm group but within the range of that produced by the hypotonic subjects (Table 6).

3.10 Statistical Analysis of Results

After discussion with the statistician at Charing Cross Hospital (Ken Macrae, personal communication, 1989), it was decided to use the Mackintosh computer programme STATWORKS to store and statistically analyse the results of this project.

The programme, which offers all possible statistical formulae, proved to be easy to use for analysing the results from the pilot study and was used for data storage and statistical analysis in the full study.

CHAPTER 4 : RESULTS.

4.1 Introduction.

68 subjects in total were examined, 13 female and 55 male. The age range was 40 to 85 years. In this results section, we will consider, for all 68 subjects examined, the following: the total pressure and duration recorded in all subjects - as shown in Table 7 (Appendix III) as Pressures 1 and Duration 1 respectively. The mean recordings for each subject are shown as Mean Pr 1 and Mean Dur 1 in the same Table. Additionally, the characteristic patterns of the frequency measures will be described.

4.1.1 Hypotonic.

15 subjects in the hypotonic group, Group 1, were recorded; 4 female and 11 male. The age range was 53-80 years, with a mean age of 66.4 yrs. The opening pressures are shown as Pressures 1 Group 1 and the duration of phonation as Duration 1 Group 1 in Table 7 (Appendix III).
The hypotonic subjects were then asked to phonate using digital pressure (see 3.6.1) to increase the tone in the P-E segment and re-recorded. These results are shown as Pressures 2 and Duration 2, Group 1 in Table 7 (Appendix III).

Some subjects were unable to produce sound during these manoeuvres. Subject 11 appeared to have no tone in the P-E segment. The sound was inaudible, although the segment was seen to open during air insufflation when viewed fluoroscopically. Pressure measurements were recorded although the sound was inaudible, being whispery in quality and thus difficult to ascertain from background stoma noise on replaying the recording. Once digital pressure was applied, the sound was audible and pressure and duration could then be measured (Pressure 2 and Duration 2 in Table 7, Appendix III).

During addition of digital pressure, some subjects over-occluded and could then not produce sound, so acoustic recordings could only be taken from 7 of the 15 subjects (Pressure 2, Duration 2, Group 1 in Table 7, Appendix III).

4.1.2 Tonic.

In the Tonic group (Group 2), all 20 subjects were recorded. There were 2 female and 18 male subjects; age range was 40-82 yrs with a mean age of 61.35 yrs. Opening pressures of the P-E segment and duration of phonation during /a/ can be seen as Pressure 1, Duration 1, Group 2 in Table 7, Appendix III).

4.1.3 Hypertonic.

In the Hypertonic group (Group 3) all 13 subjects were recorded; 3 female and 10 male. The age range was 48-85 yrs with a mean of 63.65 yrs. Recordings obtained from subjects in this group are seen as Pressure 1, Duration 1, Group 3 in Table 7 (Appendix III).

4.1.4 Spasm.

In the spasm group, (Group 4) 20 subjects were recorded; 4 female and 16 male. The age range was 48-85 yrs. with a mean age of 63.65 yrs. Four of the subjects could not phonate on /a/ during air insufflation as they complained of gastric distention in three cases (Subjects 65, 66, 67,) and this could be viewed on the fluoroscopy screen. In fact, subject 66 did produce some sound when attempting /a/ but it was of extremely short duration (0.74 sec). Subject 68 could also not phonate during air insufflation. He complained of impending syncopy so the recording had to be abandoned. For all these subjects, the fluoroscopy showed spasm at the P-E segment and the pressure recordings are the maximum of those recorded, although little/no phonation occurred.

All subjects in this group were offered a surgical myotomy to release the P-E segment (see section 1.3.6.1) for improved vocalisation. 4 subjects underwent this procedure and they were re-recorded after surgery. These results are shown as Pressure 2, Duration 2, Group 4 in Table 7 (Appendix III).

4.2 Pressure Recordings : Description of Chart Recordings.

4.2.1 Hypotonic. Group 1.

The characteristic form of the chart recording for hypotonic subjects is shown as in fig 27. for subject No. 12.

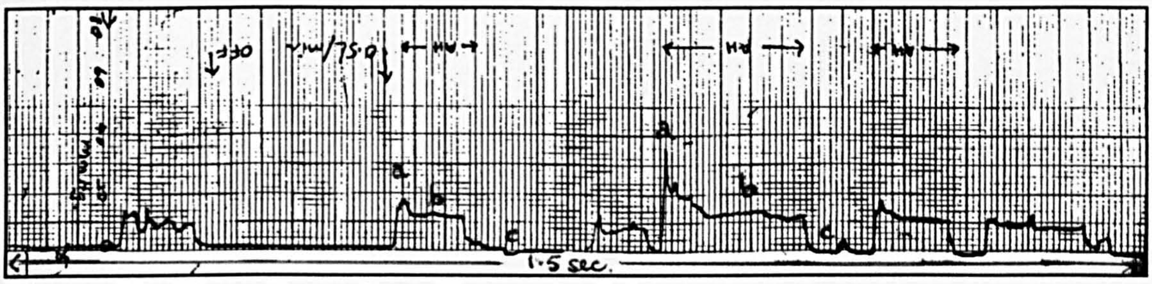


Figure 27. Chart Recording: Hypotonic Subject.

There is a gradual build up of pressure, ranging between 12.0-29.33 mmHg with a mean opening pressure of 19.27 mmHg and s.d. of 4.92 (Table 8).

Variable: Mean PR 1		Observations: 15	
Minimum:	12.00	Maximum:	29.33
Range :	17.33	Median:	18.33
Mean:	19.27	Standard Error:	1.27
Variance:			24.24
Standard Deviation:			4.92
Coefficient of Variation:			25.55
Skewness:	0.31	Kurtosis:	-0.88

Table 8. Statistics: Opening pressure recordings in Hypotonic Subjects.

The pressure peaks as the P-E segment opens (a, in fig.27) when sound is heard. The pressure then drops slightly (b = sustaining pressure) and is maintained throughout the phonation of /a/. Once phonation ceases, the pressure returns to normal (zero. See c in fig. 27).

When digital pressure is added, the opening pressure of the P-E segment significantly increases (fig.27 c.f. fig. 28) and the peaks of pressure shown on the chart recordings are shown to be higher (a).

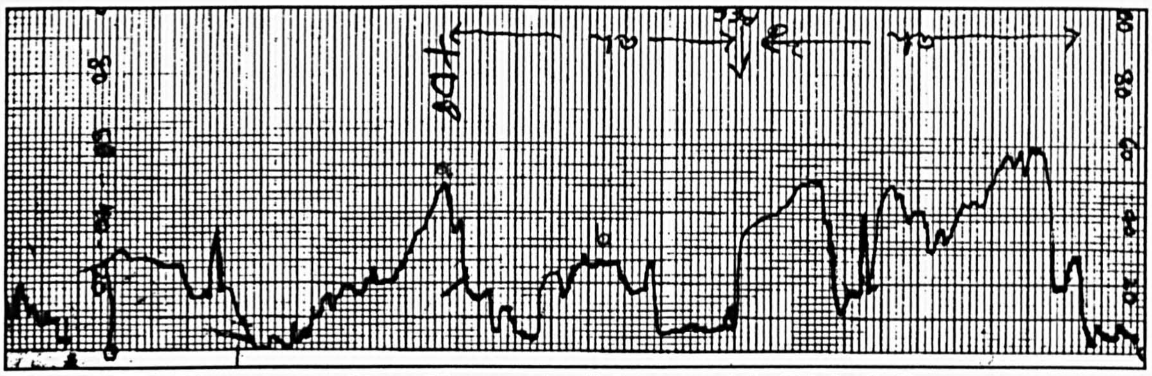


Figure 28. Chart recording: Hypotonic Subject with Digital Pressure.

The sustaining pressure during phonation is often similar in measurement, although more irregular in shape (b). The opening pressure with the addition of digital pressure rises, ranging from 22.67-64.0 mmHg with a mean of 40.51 mmHg and s.d. of 12.56 (Table 8a).

Variable: Mean PR 2		Observations: 11	
Minimum:	22.67	Maximum:	64.00
Range :	41.33	Median:	35.33
Mean:	40.51	Standard Error:	3.79
Variance:		157.66	
Standard Deviation:		12.56	
Coefficient of Variation:		30.99	
Skewness:	0.37	Kurtosis:	-1.20

Table 8A. Statistics: Pressure Recordings in Hypotonic Subjects with Digital Pressure.

4.2.2 Tonic. Group 2.

The tonic subjects (fig 29. showing subject No. 22) have a more irregular pressure pattern than the hypotonic group (1).

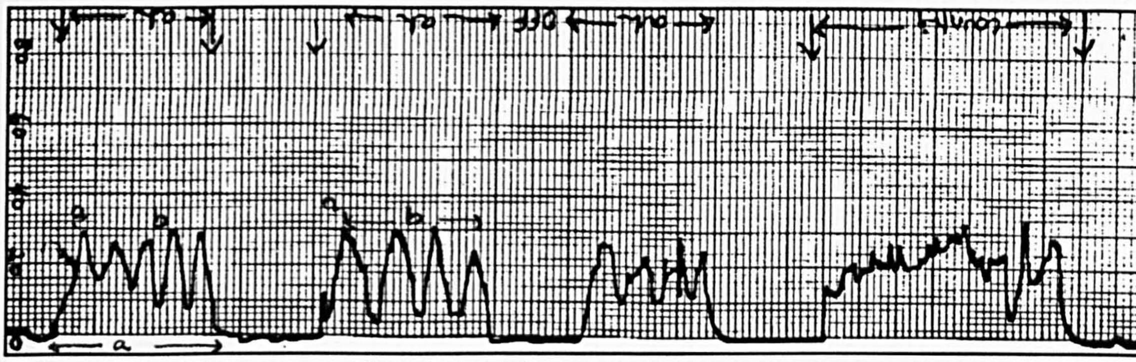


Figure 29. Chart recording: Tonic Subject.

There is a more gradual build up of pressure than in Group 1, (see (a) in fig 29) but the sustaining pressure is less constant,(b) although it is present throughout the sound. The chart recording is more staccato as the sound is sustained. The range of opening pressure recorded was 15.0-49.0 mmHg with a mean of 27.25 mmHg. s.d.= 8.05. (Table 9).

Variable: Mean PR 1		Observations: 20	
Minimum:	15.33	Maximum:	49.00
Range :	33.67	Median:	26.67
Mean:	27.25	Standard Error:	1.80
Variance:		64.78	
Standard Deviation:		8.05	
Coefficient of Variation:		29.54	
Skewness:	0.89	Kurtosis:	0.63

Table 9. Statistics: Tonic Subject.

4.2.3 Hypertonic. Group 3.

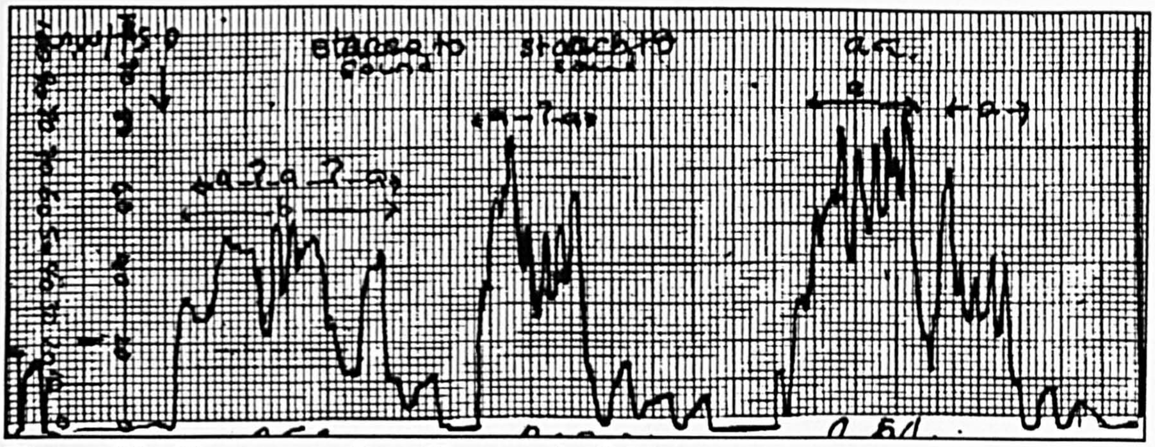


Figure 30. Chart recording: Hypertonic Subject.

The hypertonic subjects (e.g. subject No. 40 in fig. 30) have a higher opening pressure (a) recorded, ranging from 38.0-54.67 mmHg, with a mean of 45.56 mmHg and s.d. = 6.17. The pattern during sustained phonation (b) is very irregular with a fluctuating pressure (fig 30) See Table 10.

Variable: Mean PR 1		Observations: 13	
Minimum:	38.00	Maximum:	54.67
Range :	16.67	Median:	43.33
Mean:	45.56	Standard Error:	1.71
Variance:		38.07	
Standard Deviation:		6.17	
Coefficient of Variation:		13.54	
Skewness:	0.10	Kurtosis:	- 1.81

Table 10. Statistics: Hypertonic Subjects.

4.2.4 Spasm. Group 4.

Figure 31. Chart Recording: Spasm Subject.

The spasm subjects have short, sharp peaks of opening pressure (a) as a short, staccato /a/ sound is heard (fig. 31 shows subject No. 59). Although the pressure does not return to zero after phonation, as there is some air left in the oesophagus, it is only once the pressure sharply builds up and the P-E segment is released that sound is heard. The pressure recorded ranges from 39.33-99.33 mmHg with a mean of 78 mmHg. s.d.= 18.0 (Table 11).

Variable: Mean PR 1		Observations: 20	
Minimum:	39.33	Maximum:	99.33
Range :	60.00	Median:	83.00
Mean:	78.00	Standard Error:	4.03
Variance:		324.01	
Standard Deviation:		18.00	
Coefficient of Variation:		23.08	
Skewness: - 0.81		Kurtosis: -1.41	

Table 11. Statistics: Pressure recordings in Spasm Subjects.

After myotomy, the same subject exhibits different chart recordings (Fig 31 c.f. Fig 32); similar in appearance to the hypotonic subjects, but higher in recorded opening pressure.

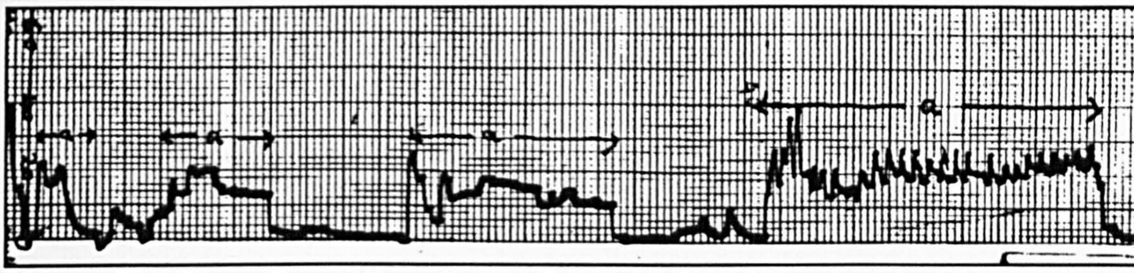


Figure 32. Chart Recording: Spasm Subject after Myotomy.

There is a sharp opening pressure recorded, followed by a lower sustaining pressure during phonation of /a/. Pressure recorded ranged from 28.0-37.3 mmHg with a mean of 32 mmHg recorded. s.d.= 3.95. See Table 11a.

Variable: Mean PR 2	Observations: 4
Minimum: 28.00	Maximum: 37.33
Range : 9.33	Median: 31.33
Mean: 32.00	Standard Error: 1.96
Variance:	15.39
Standard Deviation:	3.92
Coefficient of Variation:	12.26
Skewness: 0.35	Kurtosis: -1.88

Table 11a. Statistics: Spasm Subjects, after myotomy.

4.2.5 Statistical Analysis of Pressure Recordings.

A one way analysis of Variance between the 4 groups shows a statistically significant difference in the pressures recorded. See Table 12.

Source	Sum of Squares	Deg. of Freedom	Mean of Squares	F-Ratio	Prob>F
Between Groups	34102.40	3	11367.47	96.66	0.000
Error	7174.04	61	117.61		
Total	41276.45	64			

Table 12. ANOVA. Pressure Measurements Between Groups.

($F(3) = 96.66$; $p < 0.001$. One tailed). Thus, the observed pressure differences between the groups is highly significant.

In Group 1 (Hypotonic), a paired t-test shows a highly statistically significant difference in the pressures recorded with and without digital pressure. See Table 13.

Group 1.

Independent Samples...

Variable:	Mean Pr. 1	Mean Pr 2
Mean:	19.60	38.17
Std Deviation:	5.11	10.38
Observations:	10	10

t-statistic:	-5.08	Hypothesis:
Degrees of Freedom:	18	Ho: u1 = u2
Significance:	0.0005	Ha: u1 = u2

Table 13. Pressure Measurements with/without Digital Pressure in Hypotonic Subjects.

There is a statistically significant difference between the two groups ($t(18) = 5.08$; $p < 0.005$).

In Group 4 (Spasm) there is a highly statistically significant difference in the pressures recorded before and after myotomy, despite the small number (4) of subjects who underwent this procedure. Using a paired t-test, results show a significant difference between the two groups. ($t(6) = 10.30$; $p < 0.005$). See Table 14.

Group 4.

Independent Samples...

Variable:	Mean Pr. 1	Mean Pr 2
Mean:	88.17	32.00
Std Deviation:	10.17	10.38
Observations:	4	4

t-statistic:	10.30	Hypothesis:
Degrees of Freedom:	6	Ho: u1 = u2
Significance:	0.0005	Ha: u1 = u2

Table 14. Pressure Measurements Before/After Myotomy In Spasm Subjects.

There was no correlation between the age of the subjects and the pressure recorded. A Spearman rank correlation gave a negative value of $r = -0.03$. Thus, it is unlikely that as the subjects became older, they tend towards hypotonicity due to loss of muscle tone with age (Table 15).

Fig 33 shows graphically that the pressure measurements recorded from the four groups (hypo - tonic - hyper - spasm) are along a continuum. It can be further seen that the spasm group recordings are more variable than any of those from the other three groups.

Variable:	Age
Variable:	Pressures 1
Paired Observations:	68
Test:	Spearman Correlation
Statistic:	-0.03
Significance:	0.831

Table 15. Correlation between Age/Pressure Recordings.

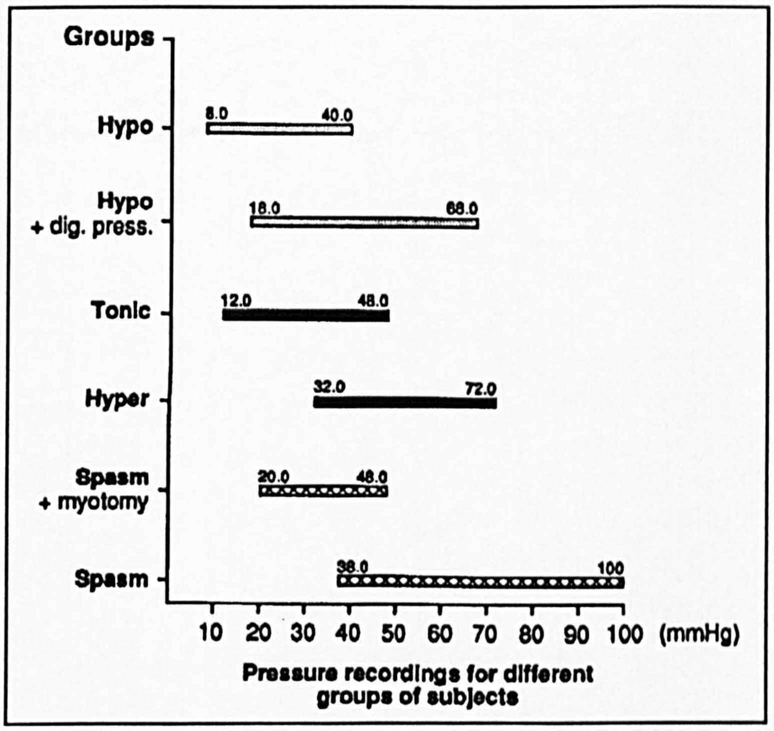


Figure 33. Chart of Pressure measurements in all Groups.

4.3 Acoustic Recordings: Duration of Phonation.

The duration of the vowel sound /a/ which was recorded varied considerably between groups. See fig. 34.

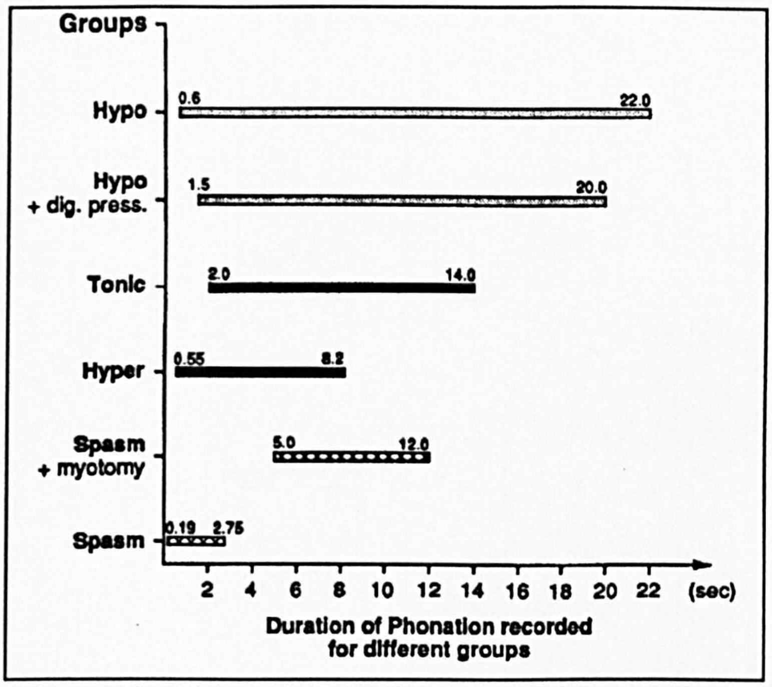


Figure 34. Chart of Duration Measurements in all Groups.

4.3.1 Hypotonic.

The duration of recorded vowel sound was 0.6-21.0 sec with a mean of 8.73 sec; s.d. = 5.64. See Table 16.

Variable: Mean DUR 1		Observations: 45	
Minimum:	0.60	Maximum:	21.00
Range :	20.40	Median:	6.73
Mean:	8.37	Standard Error:	1.46
Variance:		31.80	
Standard Deviation:		5.64	
Coefficient of Variation:		67.35	
Skewness: 0.81		Kurtosis: -0.33	

Table 16. Duration of Phonation: Hypotonic Subjects.

When digital pressure was added, the duration of sound changed from 1.9-14.0 sec with a mean of 6.23 sec; s.d. = 4.23. See Table 16a.

Variable: Mean DUR 2		Observations: 7	
Minimum:	1.93	Maximum:	14.00
Range :	12.07	Median:	5.33
Mean:	6.23	Standard Error:	1.60
Variance:		17.91	
Standard Deviation:		4.23	
Coefficient of Variation:		67.96	
Skewness: 0.60		Kurtosis: -1.06	

Table 16a. Duration of Phonation: Hypotonic Subjects with Digital pressure.

4.3.2.**Tonic.**

The duration of sound recorded ranged from 2.03-13.33 sec with a mean of 6.95 sec; s.d.= 3.06. See Table 17.

Variable: Mean DUR 1		Observations: 20	
Minimum:	2.03	Maximum:	13.33
Range :	11.30	Median:	7.50
Mean:	6.95	Standard Error:	0.62
Variance:		7.60	
Standard Deviation:		2.76	
Coefficient of Variation:		39.66	
Skewness: 0.03		Kurtosis: -0.38	

Table 17. Duration of Phonation Tonic Subjects.

4.3.3.**Hypertonic.**

The duration of sound recorded ranged from 1.03-7.73 sec with a mean of

2.68 sec; s.d. = 1.85. See Table 18. This was much shorter than either hypotonic or tonic subjects.

Variable: Mean DUR 1		Observations: 13	
Minimum:	1.03	Maximum:	7.73
Range :	6.70	Median:	1.76
Mean:	2.68	Standard Error:	0.51
Variance:		3.43	
Standard Deviation:		1.85	
Coefficient of Variation:		69.18	
Skewness: 1.45		Kurtosis: 1.43	

Table 18. Duration of Phonation: Hypertonic Subjects.

4.3.4. **Spasm.**

The duration of recorded sound was extremely short; 0.33-1.87 sec with a mean of 0.85 sec; s.d.= 0.41 (Table 19).

Variable: Mean DUR 1		Observations: 13	
Minimum:	0.33	Maximum:	1.87
Range :	1.54	Median:	0.72
Mean:	0.84	Standard Error:	0.10
Variance:		0.17	
Standard Deviation:		0.41	
Coefficient of Variation:		48.58	
Skewness: 1.19		Kurtosis: 0.58	

Table 19. Duration of Phonation: Spasm Subjects.

After myotomy, the duration of sound recorded in the same subject (No. 59) was longer; 7.0-8.93 sec with a mean of 8.06 sec; s.d.= 0.81. See Table 19a. This corresponds closely to the hypotonic / tonic group boundary.

Variable: Mean DUR 2		Observations: 4	
Minimum:	7.00	Maximum:	8.93
Range :	1.93	Median:	8.15
Mean:	8.06	Standard Error:	0.40
Variance:		0.65	
Standard Deviation:		0.81	
Coefficient of Variation:		10.04	
Skewness: -0.24		Kurtosis: -1.93	

Table 19a. Duration of Phonation: Spasm Subjects after Myotomy.

The results of pressure and duration are shown in Table 7 (Appendix III) for all groups.

Fig 33 shows graphically all groups studied and it can be seen from the duration measurements that again these overlap along a continuum.

4.3.5 Statistical Analysis of Duration of Phonation.

A one way analysis of variance between the four groups shows a statistically significant difference in the duration of sound recorded ($F(3) = 19.57$; $p < 0.001$. One tailed). See Table 20.

Source	Sum of Squares	Deg. of Freedom	Mean Squares	F-Ratio	Prob>F
Between Groups	611.87	3	203.96	19.57	0.000
Error	635.89	61	10.42		
Total	1247.77	64			

Table 20. ANOVA. Duration of Phonation Across Groups.

There is a negative correlation between pressure and duration recorded. Using Spearman's rank correlation, a negative correlation of r= - 0.67 was obtained (Table 21).

Variable:	Mean Pressure 1
Variable:	Mean Duration 1
Paired Observations:	65
Test:	Spearman Correlation
Statistic:	-0.67
Significance:	0.000

Table 21. Correlation between Pressure and Duration.

As the opening pressure of the P-E segment rises, the duration of sound produced becomes shorter. This can be seen graphically in Fig 35.

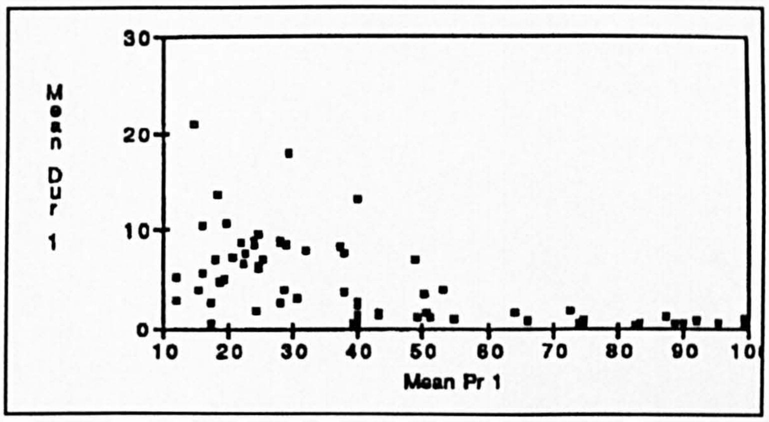


Figure 35. Graph of Correlation between Pressure & Duration recorded.

Group 1 subjects show no significant difference in length of phonation achieved with and without digital pressure, using a paired t-test. ($t(12) = 1.20$; $p > 0.001$). See Table 22.

Group 1.

Independent Samples...

Variable:	Mean Dur. 1	Mean Dur 2
Mean:	10.19	6.32
Std Deviation:	7.64	4.23
Observations:	7	7
t-statistic:	1.20	Hypothesis:
Degrees of Freedom:	12	Ho: u1 = u2
Significance:	0.253	Ha: u1 = u2

Table 22. Paired t-test Results in Group 1 Subjects with/without Digital Pressure.

This may be because the subjects, in using digital pressure, were somewhat erratic in their use of finger pressure and/or the number of subjects involved was small for a statistical significance to be investigated.

Group 4 subjects exhibit a statistically significant difference in length of phonation before and after myotomy, despite the small number (4) of subjects who underwent this procedure. Using a paired t-test, $t(6) = -17.72$. $p < 0.001$. Following myotomy, the acoustic and pressure measurements mirror closely that of the tonic group. See Table 23.

Group 4.

Independent Samples...

Variable:	Mean Dur. 1	Mean Dur 2
Mean:	0.75	8.06
Std Deviation:	0.16	0.81
Observations:	4	4
t-statistic:	-17.72	Hypothesis:
Degrees of Freedom:	6	Ho: u1 = u2
Significance:	0.000	Ha: u1 = u2

Table 23. Paired t-test Results in Subjects Before/After Myotomy.

4.4. Acoustic Recordings: Frequency Characteristics.

This was recorded as described in the Methodology section, and the recorded signal analysed using the Lynex software programme.

4.4.1.**Hypotonic.**

A very small amplitude signal is obtained from all the hypotonic subjects (See fig. 36 as a typical example from subject No. 6) i.e. the signal:noise ratio is low.

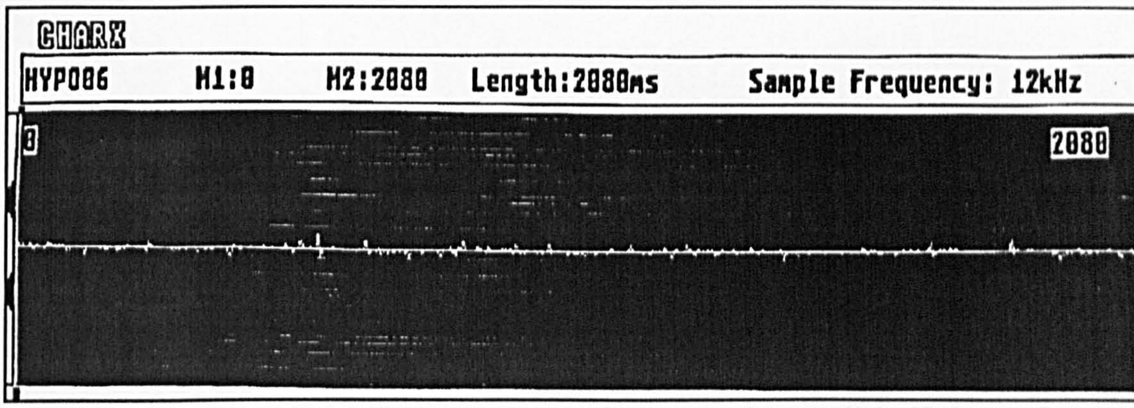


Figure 36. Hypotonic Subject.

This is audible as a quiet /a/ sound during air insufflation. If the sample signal is magnified, using the "zoom" facility of the Lynex, it can be seen that the signal trace is continuous throughout the sample (fig 37).

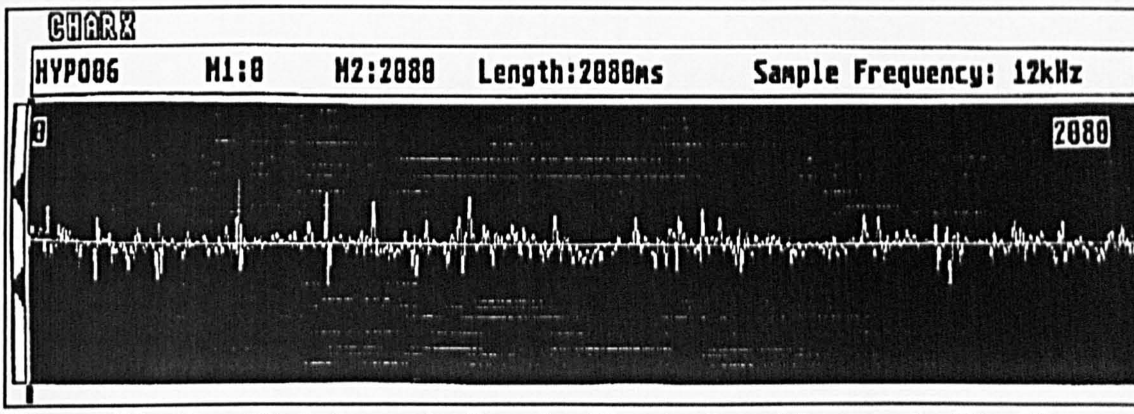


Figure 37. Magnified Acoustic Sample: Hypotonic subject as in Fig 36.

If a small sample of the sound (100ms) is expanded (fig 38), pulsile busts of sound are visible.

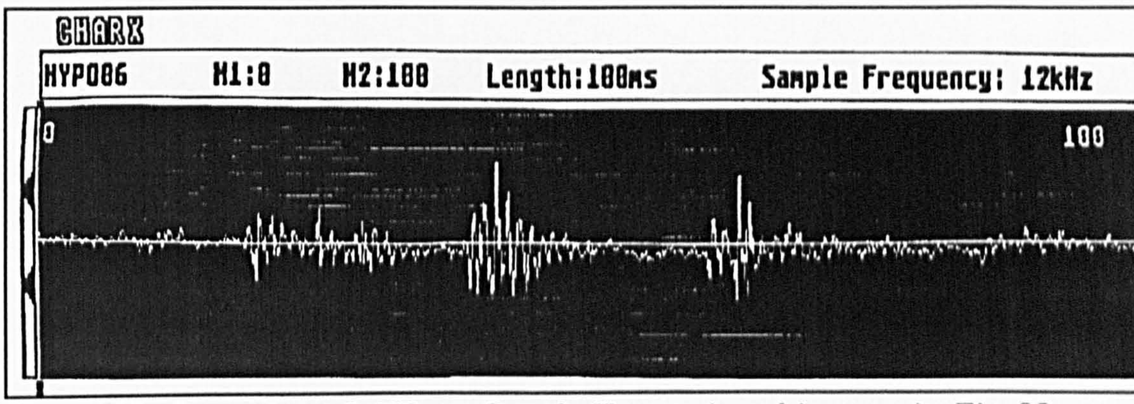


Figure 38. Expansion of 100ms. Sound: Hypotonic subject as in Fig 36.

If the waveform is further magnified (by expanding a 10ms sample, fig 39),

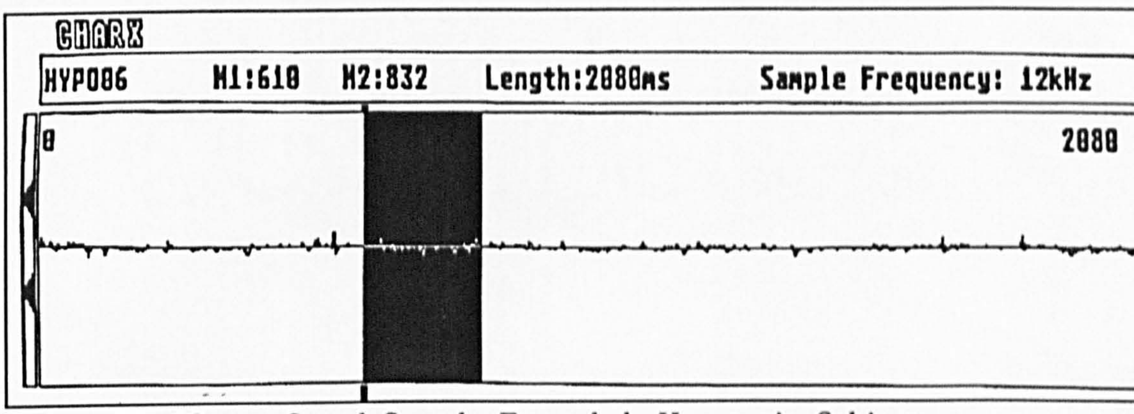


Figure 39. 10 ms. Sound Sample Expanded: Hypotonic Subject.

the trace is shown to be aperiodic (fig 40).

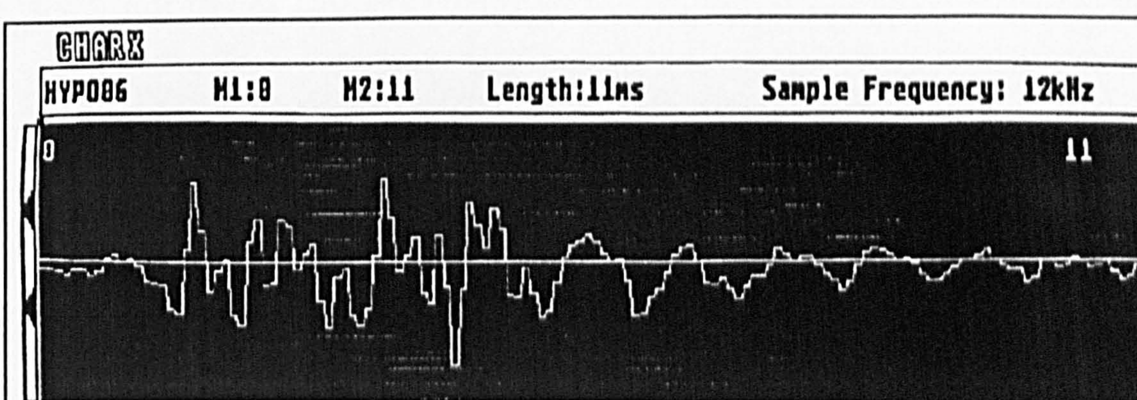


Figure 40. Hypotonic Subject showing Aperiodicity in recorded signal.

A FFT of 100ms of the signal demonstrates very weak formants by comparison with stronger background noise (fig 41).

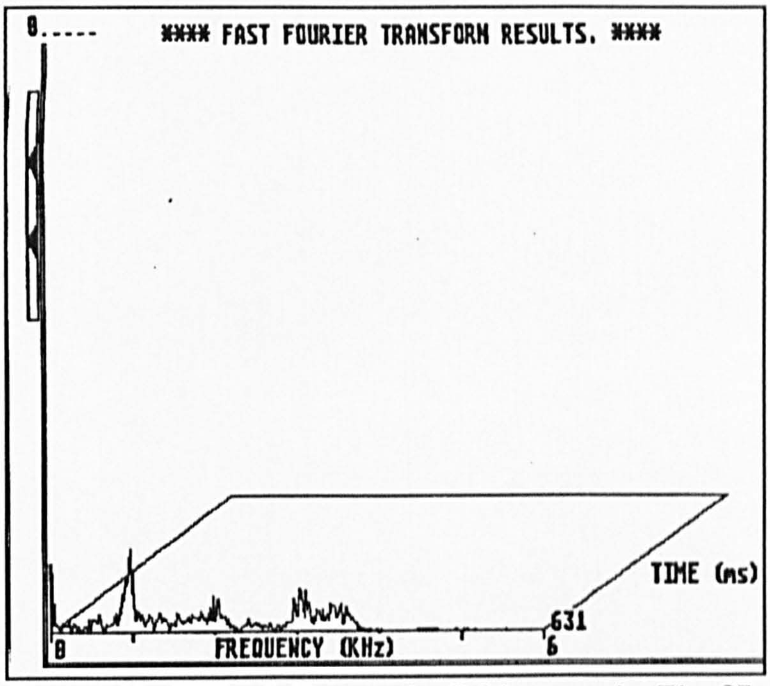


Figure 41. FFT of Hypotonic Subject as in Fig 37.

In two subjects from this group, there were no formants visible at all until digital pressure was applied.

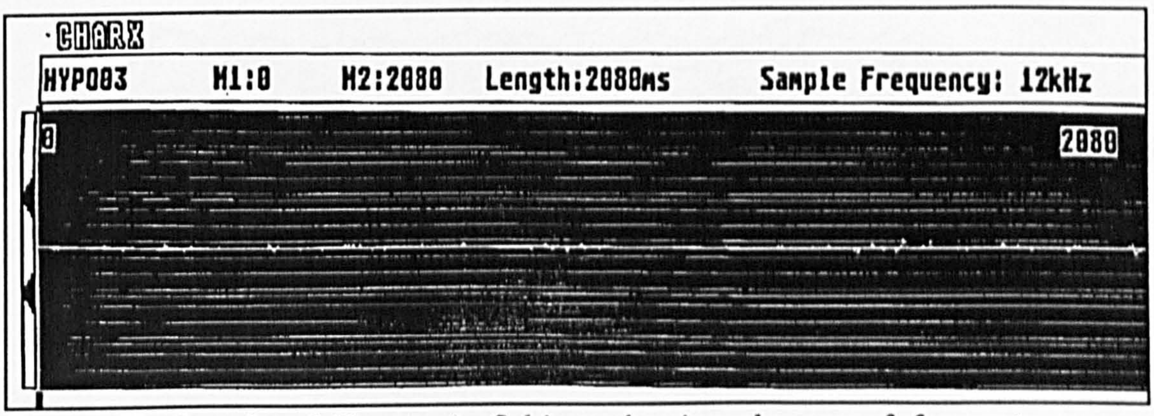


Figure 42. Sample of Hypotonic Subject showing absence of formants.

The above samples show the signal and FFT without digital pressure from subject No. 3; figs. 42i and 43i shows the signal and FFT in the same subject after digital pressure has been applied.

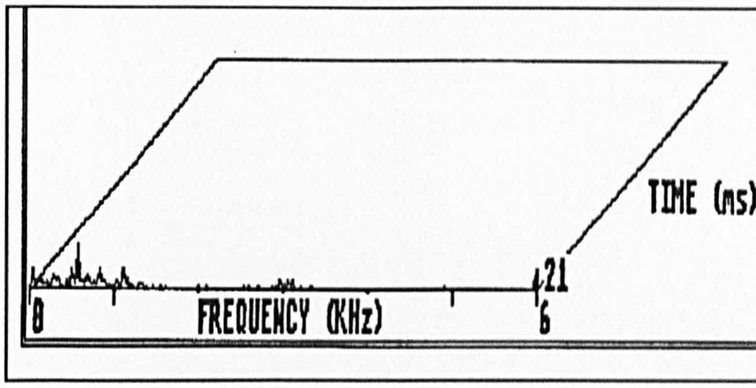


Figure 43. FFT of Hypotonic Subject with no digital pressure.

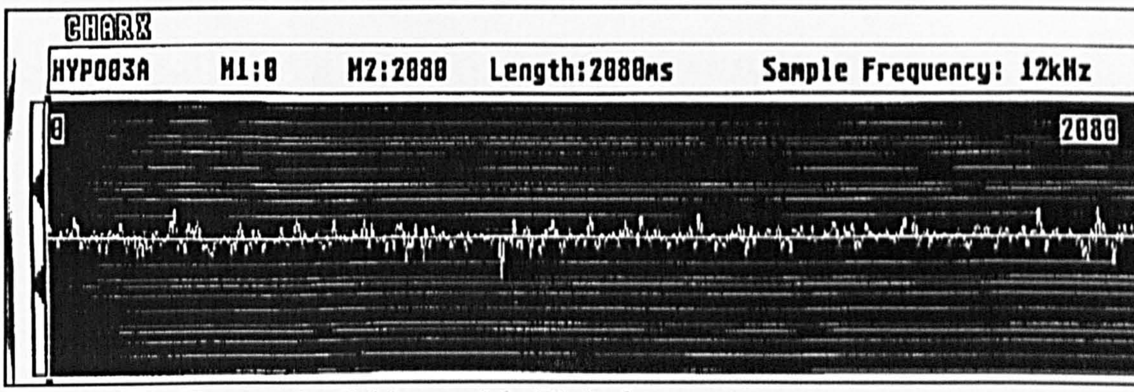


Figure 42i. Hypotonic Subject with Digital pressure.

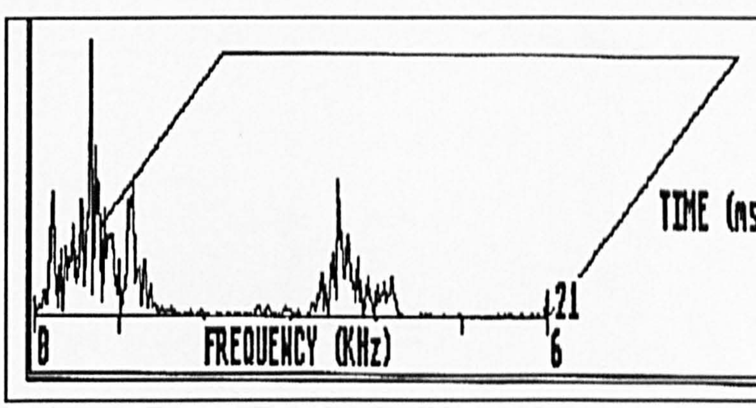


Figure 43i. FFT of Hypotonic Subject with Digital Pressure.

4.4.2 Tonic (good) speakers.

Continuous, pulsile waveforms which are non-sinusoidal are visible on all samples recorded for this group of subjects (e.g. figs 44, 45, 46 from subjects Nos. 31, 32 & 30).

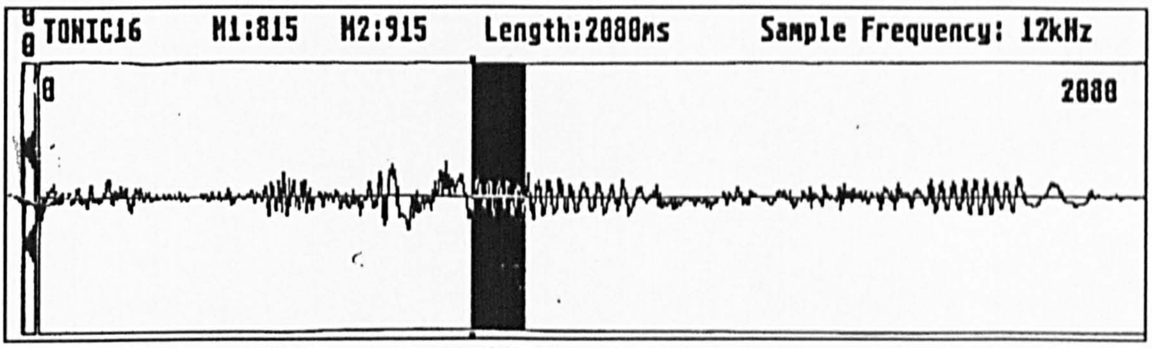


Figure 44. Acoustic Recording of Tonic Subject.

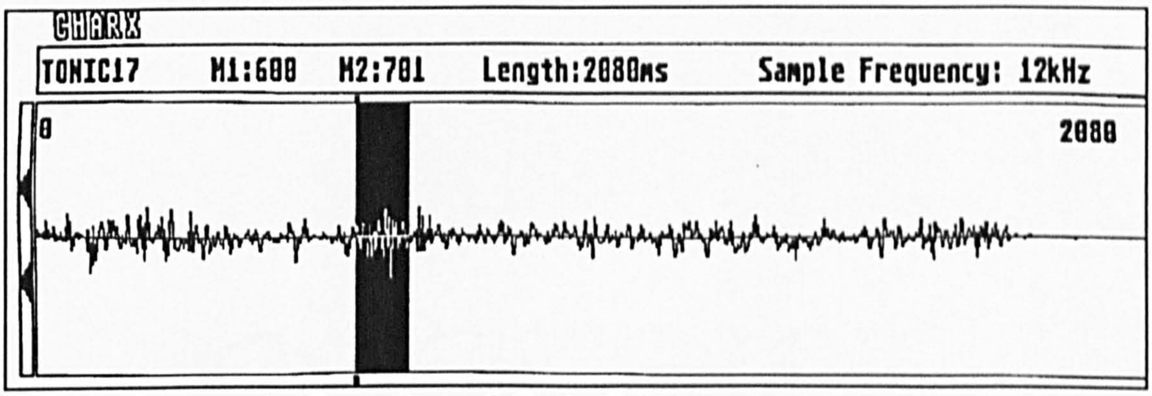


Figure 45. Acoustic Recording of Tonic Subject.

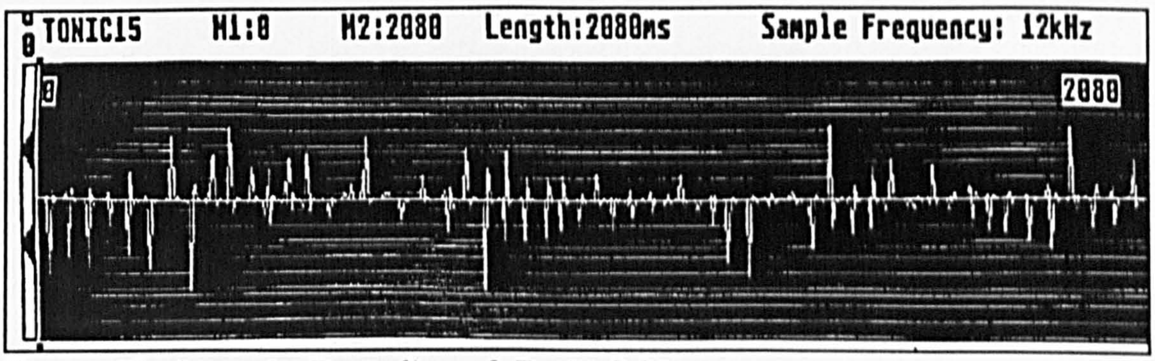


Figure 46. Acoustic Recording of Tonic Subject.

Figs 44, 45 and 46 show typical waveforms from Tonic subjects. It can be seen from Figs 44ii, 45ii and 46ii (p124) that there are considerable proportions of periodicity in each of these samples when short (10ms) sections of the sounds are expanded and magnified.

There was evidence of some regularity in the frequency of the pulses in the waveform.

If a small sample (100-140ms) of sound is magnified, the sound can be seen to be almost periodic in format (figs. 44i, 45i, 46i.).

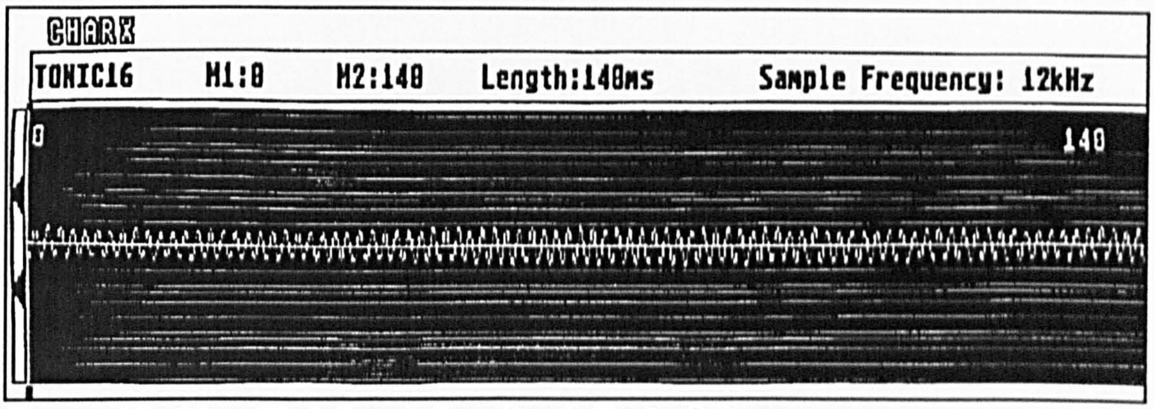


Figure 44i. Magnified Sound Sample of Tonic Subject.

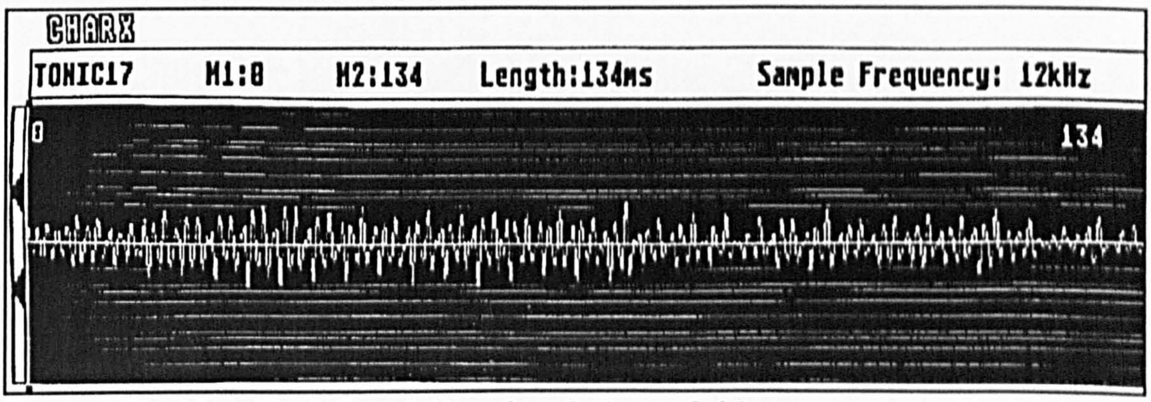


Figure 45i. Magnified Sound Sample of Tonic Subject.

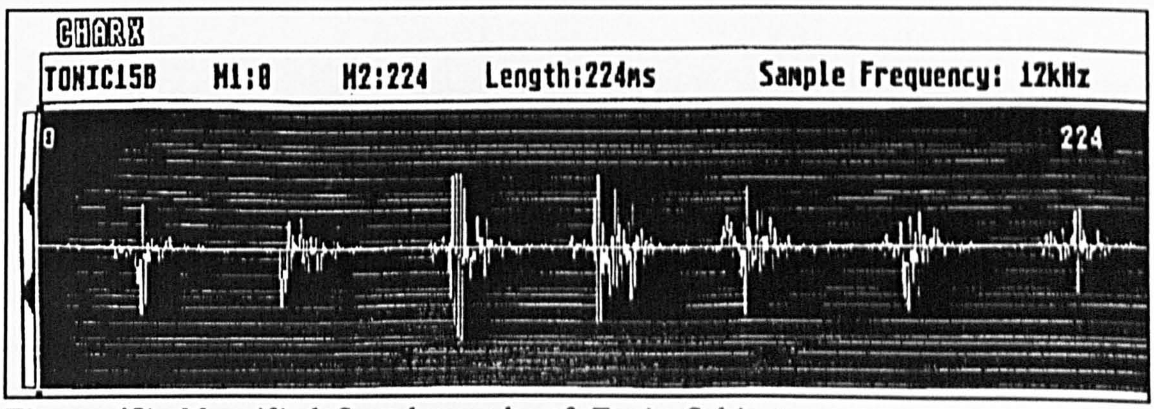


Figure 46i. Magnified Sound sample of Tonic Subject.

If the waveform is closely examined, by expanding 10ms of sound, the signal can be seen to be regular with evidence of periodicity (figs. 44ii, 45ii, 46ii).

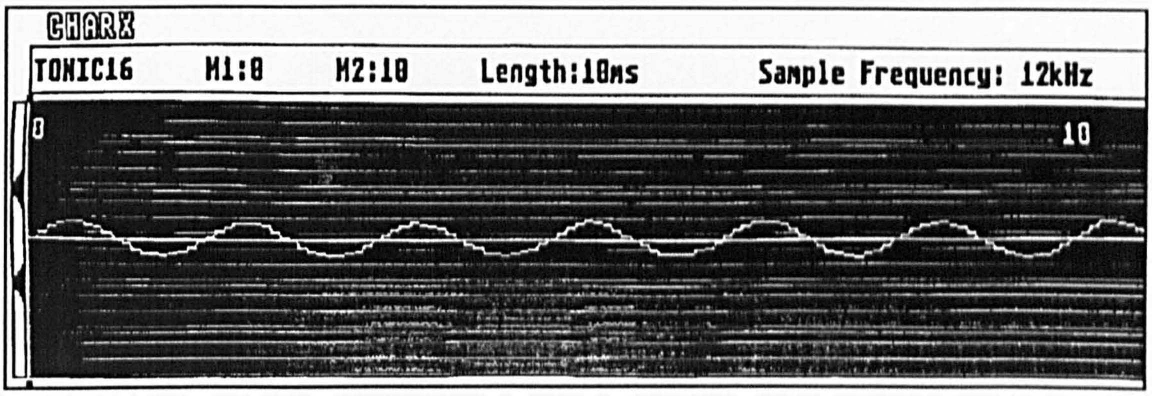


Figure 44ii. Further Expansion of Sound Showing Periodicity.

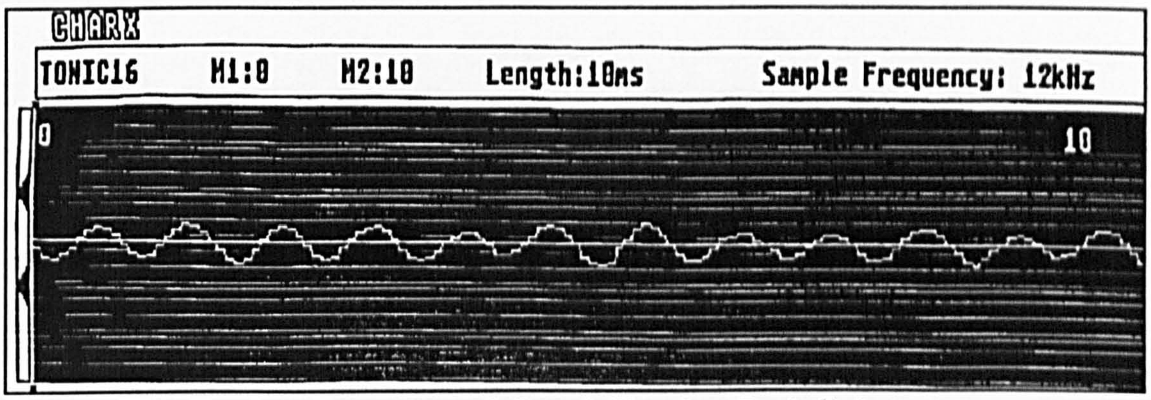


Figure 45ii. Further Expansion of Sound Showing Periodicity.

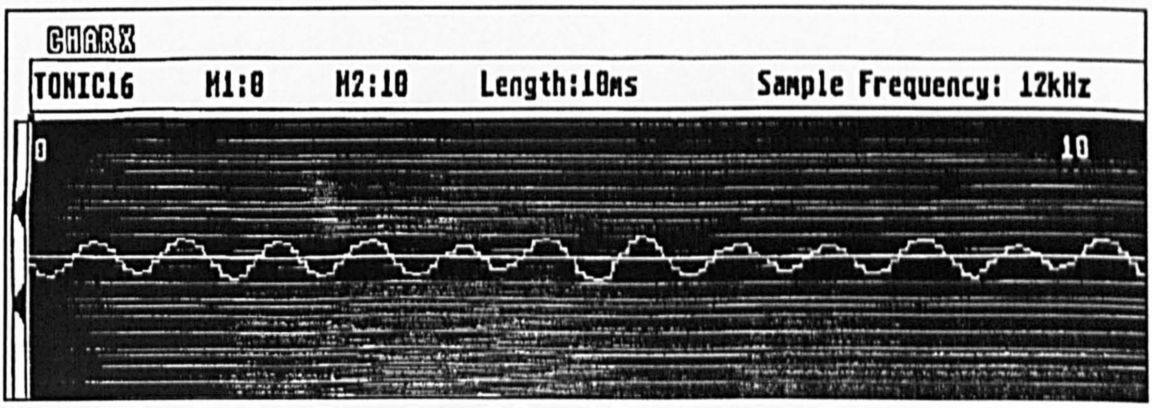


Figure 46ii. Further Expansion of Sound Showing periodicity.

The frequency recorded for subject 31 (male) ranges from 625 pulses per second (pps) to 1125 pps. Subject 32 (male) can be seen to exhibit sound of approximately 600pps and subject 30 (female) produces frequency at a level of 1000pps. There is evidence of regularity in the signal, although not continuously throughout the whole sample, for all tonic subjects. FFT for all subjects in the tonic group shows dominant formants with respect to background noise.

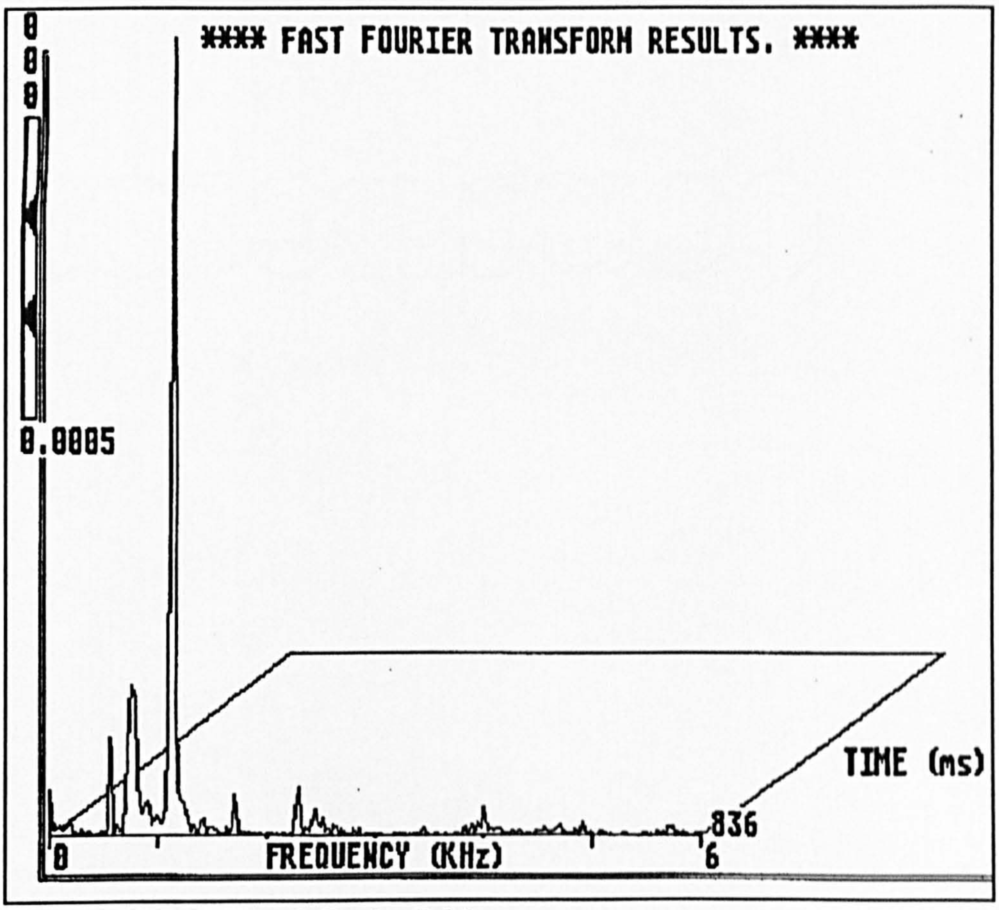


Figure 44iii. FFT of Tonic Subject.

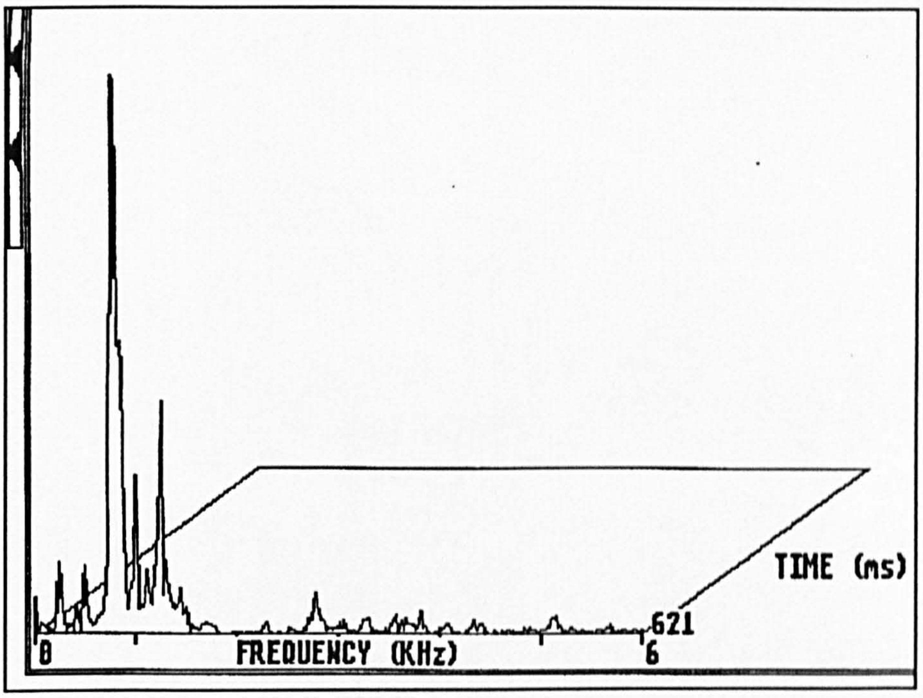


Figure 45iii. FFT of Tonic Subject.

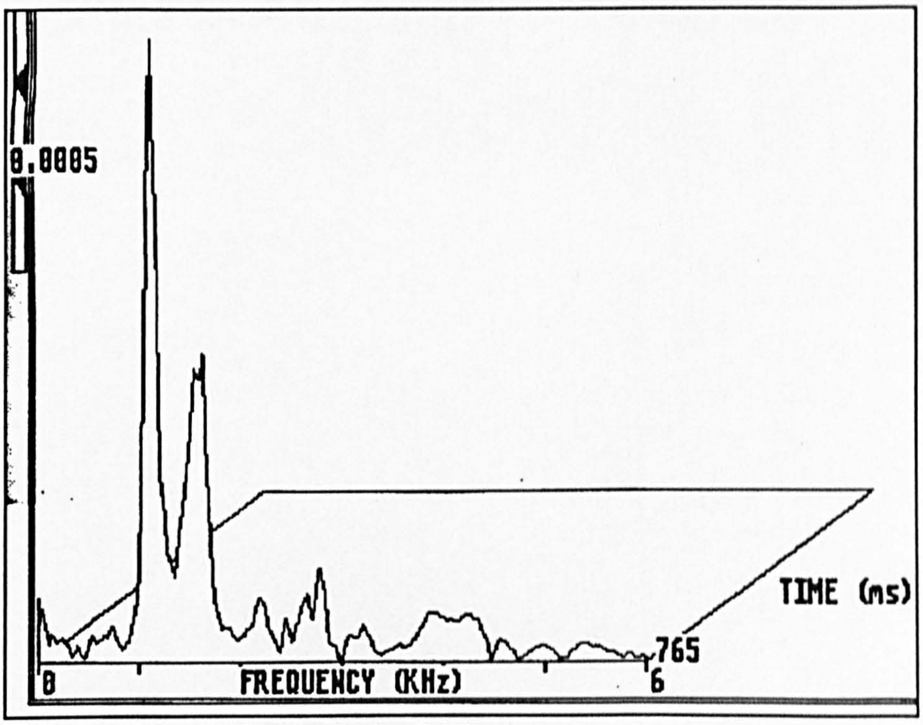


Figure 46iii. FFT of Tonic Subject.

The FFTs for the subjects illustrated in figs 44iii, 45iii, 46iii, show formant behaviour as illustrated in Table 24.

Subj No.	30	30a	30b	31	31a	31b	32	32a	32b
F1	1100	1100	1100	1150	800	750	850	850	850
F2	1600	1450	1450	750	1300	1450	1350	1250	1350
F3	2800	2800	2650	550	--	--	2850	2700	2850

Table 24. Formants Present in Tonic Subjects 30, 31 & 32.

For both subjects 30 and 32 there was remarkable consistency in the presence of F1 on each of the three attempts at /a/. In subject 31 there was a formant at 750 Hz on the first recording but this was not the dominant formant - that was at 1150 Hz. These correspond to formants at 124, 730, 1090 and 2440 Hz for F0, F1, F2, F3 for normal male laryngeal speakers; 212, 850, 1220, 2810 Hz for normal female speakers. Thus there seems to be little evidence that formants in laryngectomised speakers correspond to those found in laryngeal speakers.

4.4.3 Hypertonic

By comparison with the hypotonic and tonic subjects, the waveforms of all subjects in this group are irregular in shape. The amplitude of the signal is large (fig 47 showing subject No. 37) and shows considerable fluctuation.

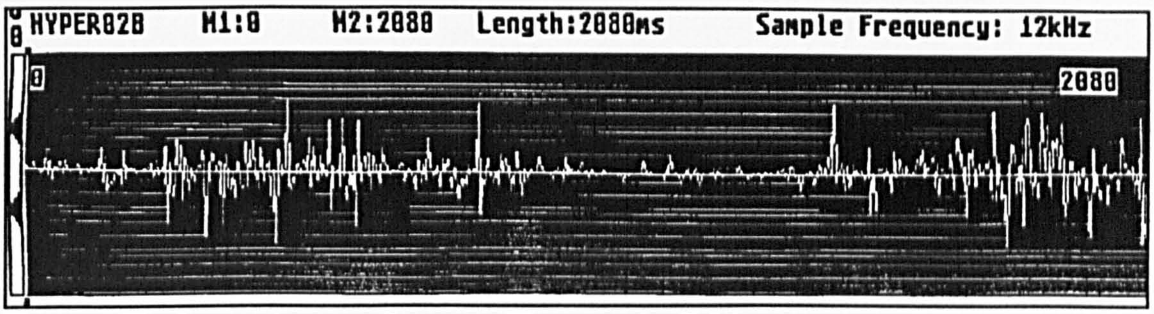


Figure 47. Acoustic waveform in Hypertonic Subject.

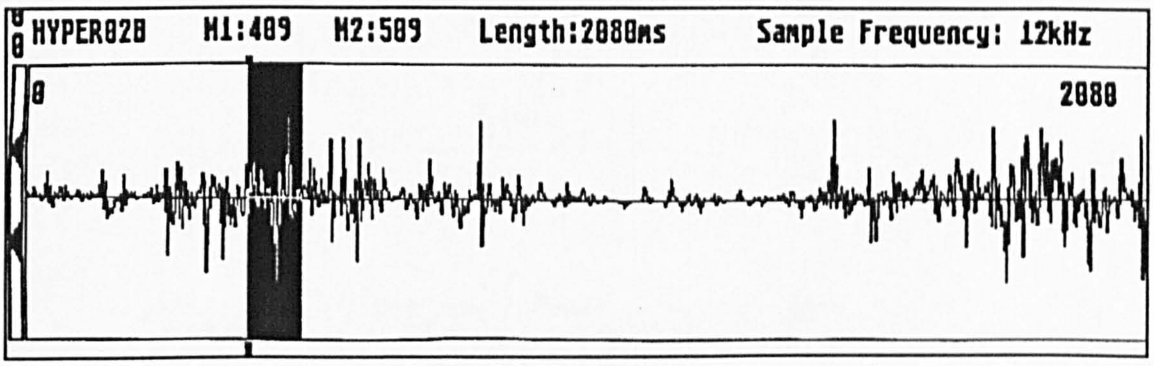


Figure 48. Magnified Waveform Of Hypertonic Subject as in Fig 47.

However, it rarely returns to zero (silence) during the sample which differentiates this sample from the spasm group (fig 52). If the sample is magnified (fig 48), 200ms. can be viewed and is seen to be composed of irregular pulsile activity (fig 49).

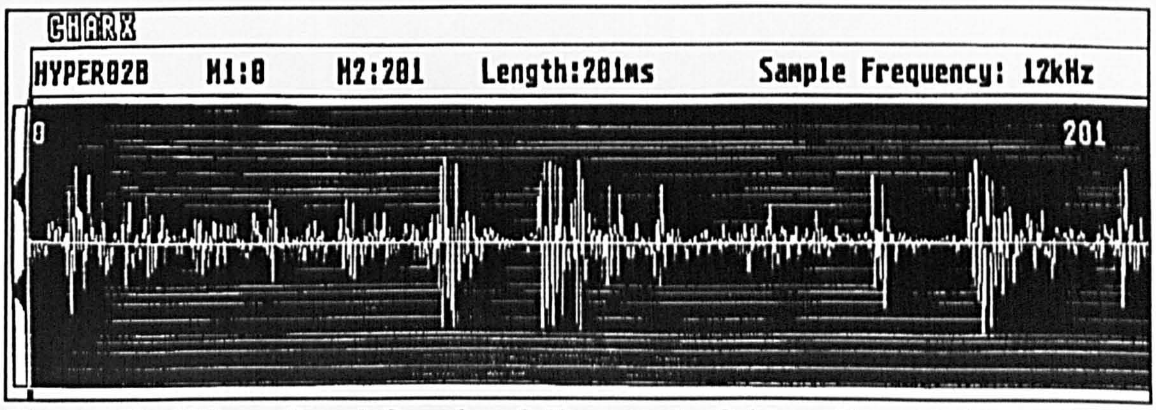


Figure 49. 200ms. Sound Sample of Hypertonic Subject Expanded.

Further amplifying 16 ms of sound shows it to be aperiodic (fig 50).

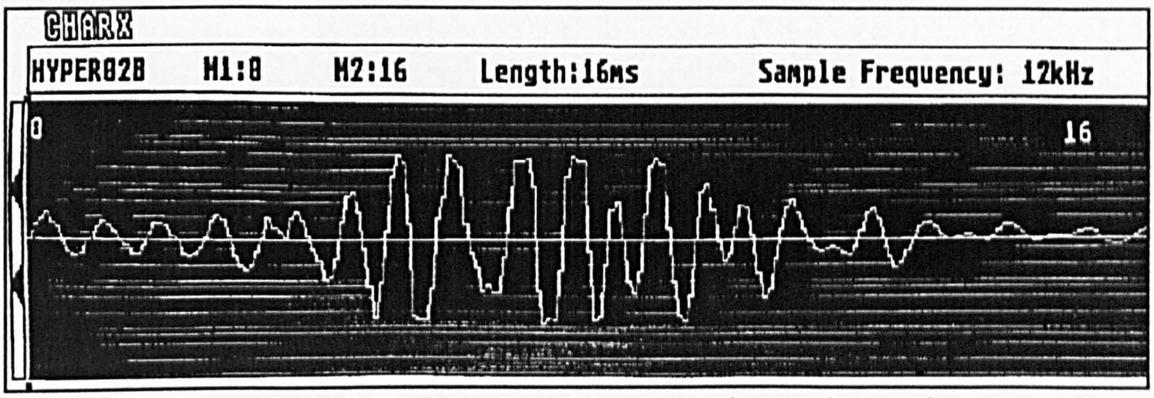


Figure 50. Expansion of Hypertonic subject sound sample to show Aperiodicity.

FFT reveals irregular formants (fig 51) with marked background noise from the subject, although two dominant formants were usually visible in all hypertonic subjects investigated.

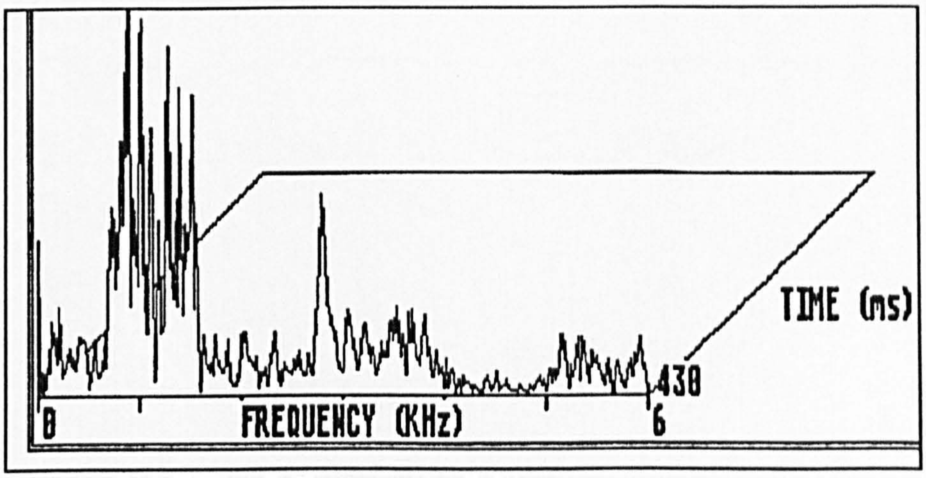


Figure 51. FFT of Hypertonic Subject.

.

4.4.4 Spasm.

There are short bursts of pulsile activity against background silence (fig 52).

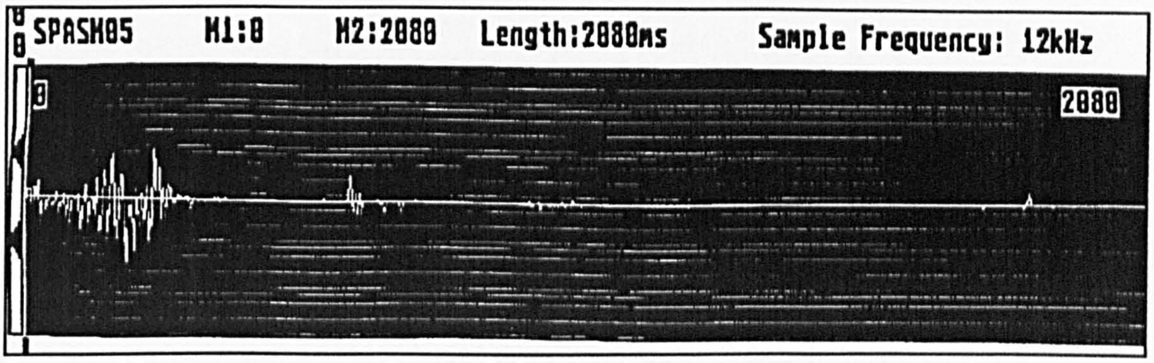


Figure 52. Acoustic waveform of Spasm Subject.

If 200ms of the sample is expanded (fig 53 from subject 53), the irregularity of the signal is visible.

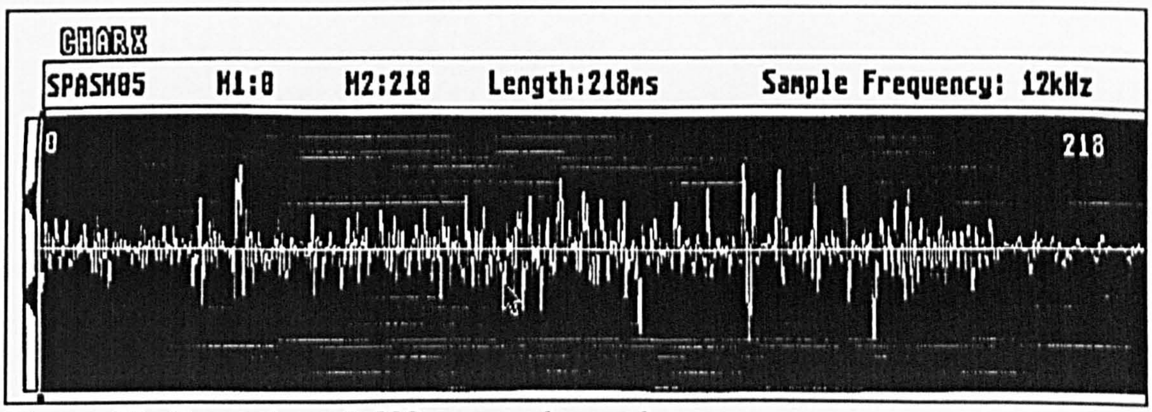


Figure 53. Expansion of 200ms sound sample.

If 9ms of this sample is expanded, this trace can be seen to be aperiodic (fig 54). Regardless of where in the signal the sample is taken, the sound is always aperiodic in appearance.

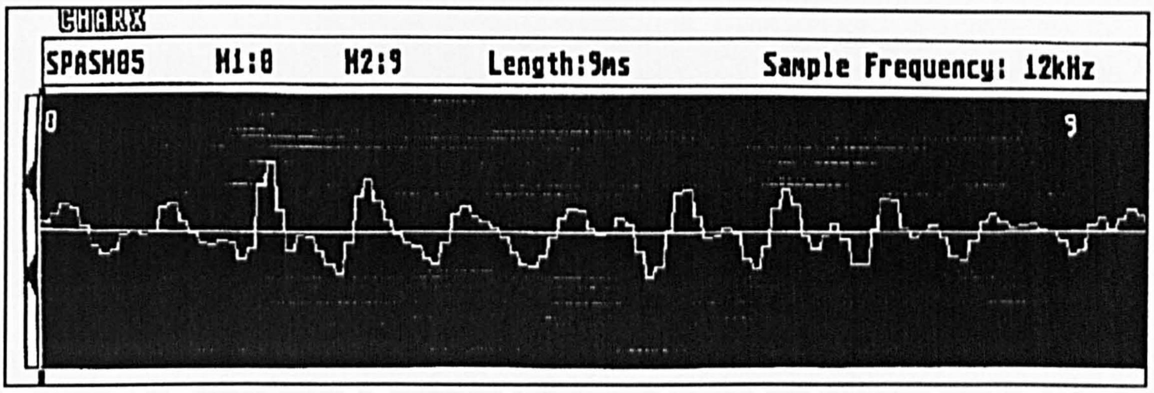


Figure 54. Expansion of Spasm subject sound sample showing Aperiodicity.

FFT demonstrates very dominant formants with little/no background noise (fig 55).

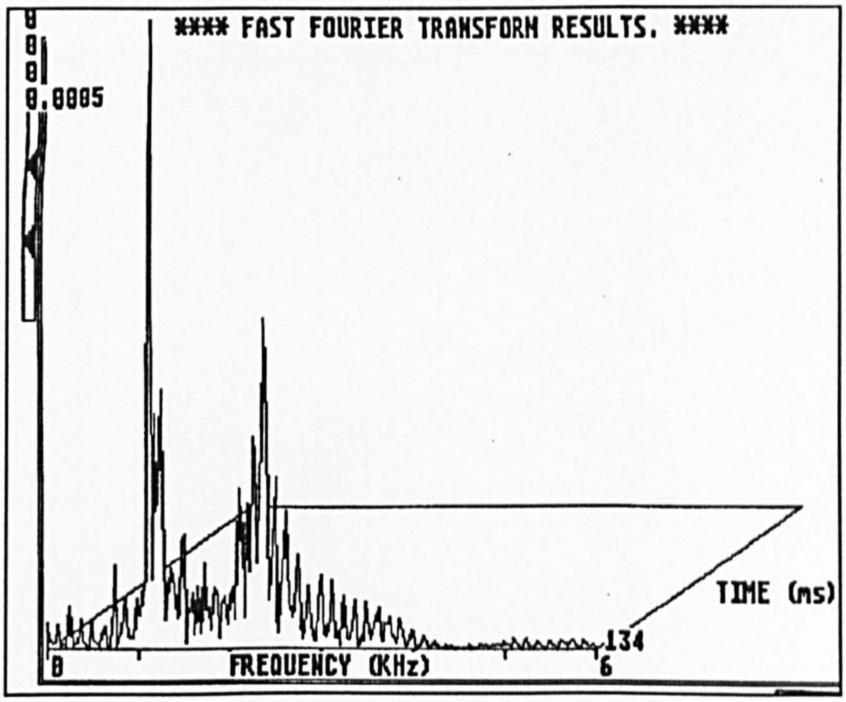


Figure 54. FFT of sample from Spasm Subject.

There is consistency in the position of the formants in all three samples of /a/ recorded for each subject.

CHAPTER 5 : DISCUSSION.

The results of this study are consistent with the experimental hypothesis. Analysis of the pressure and duration results support the hypothesis of a continuum of tonicity in the reconstructed pharynx.
The acoustic results suggest the existence of discreet groups of subjects, according to both analysis of the formant behaviour (FFT) and from the waveform characteristics. In summarising the results:

5.1 Pressure Measurements.

The pressure measurements between the four groups shows a statistically significant difference (see **4.2.5**) which provides evidence for the existence of the four discreet groups whose existence was hypothesised from clinical observations. Hitherto, the existence of the "Hypertonic" group has not been described; Singer & Blom (1981, op cit) describe spasm as if it were an all-or-nothing phenomenon but the results of this study indicate the importance of assessing the *degree* of excess pharyngeal tone as this is critical to the future clinical management of the laryngectomee.

The pressure measures recorded in the 7 Hypotonic subjects with /without digital pressure shows a significant difference for the two sets of circumstances. This was despite the problems described in **4.1.1.** it can be clearly seen in fig.33 that, following the addition of digital pressure, subjects "shift" along the continuum and then overlap with the Tonic (good oesophageal speaker) group.

Despite the small number (4) of subjects in the Spasm group who underwent surgical myotomy, the resulting changes in pressure measurements were obvious (see table 25, chapter 5) and a statistically significant was recorded between the two conditions.

Fig. 33 clearly shows that, after myotomy, the subjects "shift" groups and now have similar pressures recorded to those in the Tonic group. If the Tonic group is regarded as the ideal, then Spasm subjects approach this after the myotomy has been undertaken.

5.2 Duration of Phonation.

The duration of phonation recorded shows a statistically significant difference the four groups. This, again, provides evidence for the existence of the hypothesised continuum.

It can be seen in table 7 that, with the addition of digital pressure, the duration of /a/ sound produced decreases. Unfortunately, this test is variable as it depends on the subject's skill at judging the degree of pressure to apply. With reduction of sensation in the neck tissues, after DXT and surgery, this is often difficult to achieve. This could explain the lack of significant difference for the two conditions. With only 7 subjects recorded in a highly variable situation, perhaps the result is not surprising.

It can be seen in fig. 34 that the overall trend, with added digital pressure, is for the 7 subjects to more resemble the Tonic group, although some Hypotonic subjects could extend the vowel sound /a/ beyond that recorded for Tonic subjects. Presumably this is due to the lower resistance to air flow through the (looser) P-E segment in Hypotonic subjects.

After myotomy, subjects in the Spasm group resemble those in the Tonic group in their ability to extend phonation. This can be seen in fig. 34. the results before and after myotomy show a statistically significant difference in the length of phonation recorded for the two conditions.

There is a negative correlation between pressure and duration recorded. In Spasm subjects, the P-E segment opening pressure is high and the duration of sound is short; conversely, in the Hypotonic subjects, the P-E opening pressure is low and the duration of recorded vowel sound is long.

5.3 Frequency Characteristics.

For Hypotonic subjects, the amplitude of the signal recorded is small, as in **4.4.1**, but the voicing is continuous throughout the sample. The audible quality of voice - quiet, continuous, and breathy - is visible on the screen (fig.33), but no formants are evident on the FFT (fig. 43) until digital pressure is applied (fig.43i) when, audibly, the quality of voice improves in loudness and clarity.

Tonic speakers' voices are characterised by the presence of periodicity in the frequency waveform (figs. 44, 45, 46) and the FFT shows strong dominant formants with little/no background noise.

The Hypertonic subjects, conversely, show irregular waveforms and irregular formants on FFT.

Spasm subjects show "bursts" of sound with irregular frequency. The FFT shows sharp, dominant formants with little/no background noise and a high degree of intra-subject reliability for the position of the formants (fig 56).

There is a highly significant difference in both the pressure and the duration measurements between the groups, which may be related to the physiological characteristics / differences in the reconstructed pharynx. This does not seem to be related to age as there is no correlation between age and the pressure nor the duration of recorded sound. Thus, subjects do not seem to become increasingly hypotonic with age, for example, as they lose muscle tone.

5.4 Clinical Evidence.

Additionally, there is clinical evidence in support of the hypothesis from this study. One subject, 51 (Table 7, Appendix III) was initially assessed as being in the hypertonic group in using videofluoroscopy alone. He received a T-E procedure and it failed. In analysing the pressure and duration characteristics, the subject fitted more appropriately into the Spasm group, having a mean pressure measurement of 40 mmHg and a short duration of 0.6 secs. (Spasm subjects have a mean pressure of 78 mmHg and a duration of 0.84 sec, whereas hypertonic subjects have a mean of 45.56 mmHg but a longer duration of 2.64 secs). After undergoing a surgical myotomy (see 1.3.6.2) this subject achieved good T-E speech with no problem and, indeed, could also then produce oesophageal speech.

5.4.1 Evidence from Hypotonic Subjects.

Further clinical evidence arises from the hypotonic subjects, three of whom achieved only whispery voice after a T-E procedure. However, when added tone was applied (with digital pressure or with a collar and tie), the voice immediately became stronger and clearer. These clinical findings would seem to indicate that the subject 51 (see 5.1) was on the hypertonic/spasm end of the continuum, whereas the other three subjects (4, 9 and 10) may be classified as on the hypotonic/tonic end of the continuum.

5.4.2 Evidence from Spasm Subjects.

Four subjects of the twenty in this group underwent a surgical myotomy (see 1.3.6.2.) to release the area of the pharynx identified as being in spasm. They were re-recorded after surgery and the results were as follows:

TABLE 25

Subj.	Pre-op press. mmHg	Post-op press. mmHg	pre-op dur. secs.	post-op dur. secs.
1	74.67	37.33	0.59	8.33
2	99.33	30.67	0.98	7.00
3	90.00	32.00	0.71	7.97
4	88.67	28.00	0.71	8.93

Despite the small sample, there is a statistically significant difference in both pressure and in duration measures before and after surgery (see 4.2.5.; 4.3.5.) and the post-surgical results, as seen above, are in the direction of the hypothesised continuum. Following myotomy surgery, the pressure recorded drops, and the duration of recorded sound lengthens.

5.5 Relationship of Assessments to Subject Management.

The usefulness of these assessments lies in the fact that it takes account of those subjects who fail to develop oesophageal speech, rather than charting oesophageal speech progress as previous assessments have done (see Section 2.1).

These assessments will help to determine the path of subsequent management for the laryngectomee who wishes to acquire speech, whether this is oesophageal speech or speech after undergoing a surgical/prosthetic procedure. They provide a way of assessing the anatomical/physiological characteristics of the pharynx and P-E segment which will help plan the subsequent clinical management of the subject. They will indicate the subject for whom oesophageal speech training is inappropriate and in whom T-E puncture procedures will not be effective - i.e. those hypertonic and spasm subjects in whom the P-E segment will not release for effortless airflow until further surgery is undertaken.

They will also provide objective, repeatable measures for the future study of laryngectomy speech rehabilitation.

5.5.1 Hypotonic Subjects

If the subject is hypotonic, extra tone (in the form of digital pressure or a collar) will need to be added. This will apply whether oesophageal speech or T-E speech is being planned. By applying digital pressure it can be seen from the FFT (figs 43 and 43i) that the subject does not necessarily improve or increase formant behaviour. Rather, the digital pressure decreases background noise and thus improves the amplitude of speech. Although the waveforms and FFTs do not radically change in character, the voice quality is subjectively perceived as improved due to the increased amplitude.

5.5.2 Tonic Subjects

The tonic subjects will do well without further surgery; they should respond well to a T-E procedure and, if they wish to achieve oesophageal speech, it is simply a question of rehearsal/direction. These are the subjects referred to by Priest (1984) who achieve speech effortlessly, often without any speech therapy intervention.

5.5.3 Spasm Subjects

The subjects in spasm will need a surgical myotomy (see 1.3.6.) to release the P-E segment for more consistent airflow. They will not achieve oesophageal speech, nor do well with a T-E procedure without further surgery. The four subjects in the spasm group, who were re-recorded after myotomy, then had a mean pressure recording of 32.0 mmHg and a mean duration of 8.06 secs. which classifies them nearer the target tonic group - which had a mean pressure recording of 26.35 mmHg and duration of 6.89 sec.

5.5.4 Hypertonic Subjects

The hypertonic subjects may be managed with a T-E procedure, provided a low pressure voice prosthesis is used. This type of prosthesis allows constant airflow with low resistance and a steady flow of air is emitted into the oesophagus. A duck-billed, or higher resistance, prosthesis demands more pressure to open the tip and thus the valved end opens suddenly under pressure, emitting a burst of air into the oesophagus. This has the effect of elevating the tone in the P-E segment, possibly due to the stretch receptors in the upper circular fibres of the oesophagus detecting the extra pressure. Proprioceptive innervation of the P-E segment has been demonstrated by the presence of neuromuscular spindles in the oesophageal wall, which explains the contraction in the lower part of the sphincter and permits the control and precise degree of voluntary contraction needed in oesophageal speech (Diedrich, 1968). It may be that this mechanism reacts to the sudden input of air and raised oesophageal pressure and thus the P-E segment contracts and spasm ensues under raised oesophageal pressure conditions.

All 14 subjects in the hypertonic group who did not wish to undergo myotomy were given the option of using a low pressure prosthesis, although it was fully explained to the subject that this might not be successful and that a later myotomy might be needed. Interestingly, these subjects could not use the duck-billed prosthesis, as the extra pressure needed to open the valved tip resulted in spasm at the P-E segment, which was confirmed on videofluoroscopy. All but one subject was managed successfully without the need for myotomy. The one subject (51) has been discussed earlier (5.1).

If subjects in this group wish to learn oesophageal speech, it may be preferable to teach inhalation as a method of air charging the oesophagus as this demands relaxation of the P-E segment - which in this group of subjects tends to be elevated in tone - whereas injection results in increasing the pressure in the oral cavity, and thus in the P-E segment (see 1.3.4.).

5.6 Consideration of Previous Findings.

There is some support for this hypothesised continuum from the literature (see also Chapter 2).

5.6.1 Simpson, Smith and Gordon (1972).

They described the radiological appearances of subjects after laryngectomy. They use the term "pseudo strictures," stating:

> "These constrictions are due to bulging of the posterior pharyngeal wall into the lumen of the hypopharynx. These posterior bulges are variable and mobile. The barium swallow may appear normal or show only a small posterior filling defect, the bulge becoming more apparent under conditions of small pressure changes involved in injection and phonation. These posterior pharyngeal wall bundles are believed to be due to unsutured bundles of inferior constrictor fibres which have retracted posteriorly."(p.572).

The radiographs illustrating this description exactly correspond to the types of subjects who would be categorised on this classification as being hypertonic or in spasm, depending on the degree of hypertonicity exhibited during air insufflation. In the same paper, two patients described as exhibiting "voluntary contraction" of the pharynx correspond to a description of subjects using pharyngeal speech and these subjects should best be defined as belonging to the spasm group. The authors also describe "inert ballooning of the pharynx" which would be classified as exhibiting hypotonicity on this assessment, and this is further confirmed by examining the radiographs illustrating this paper.

Thus, there is evidence in the literature that the radiographic entities in this study exist and they have been previously observed and described, but it is their interpretation which varies and is unique to this thesis.

5.6.2 Seeman (1967).

As discussed earlier (2.3.2.1.), Seeman used manometric pressure recordings with trans-nasal insufflation and suggested that this had prognostic significance. He stated:

> "The laryngectomised patients in whom the values range between 10-35 mmHg (50%cases) soon learn to speak with an acoustically good oesophageal voice. If the pressure values are between 35-80 mmHg (30%) reeducation is more difficult and takes longer. The new voice is strident, choked and its vibrations are coarse. If the overcoming of the closure of Killian's orifice requires a pressure higher than 80 mmHg (approximately 20%) the results of rehabilitation are unsatisfactory. The patient can only master two or three words of connected speech. Voice is produced with difficulty, it is choked and interrupted by cannulary squeak. If the closure of the oesophageal orifice cannot be overcome by transnasal insufflation, then in most cases esophageal speech cannot be acquired." (p 237).

His description would correspond to the hypotonic/tonic group (10-35 mmHg); hypertonic (35-80 mmHg) and spasm (over 80 mmHg) in this study. In the light of this study, and contrary to Seeman's suggestions, these are not discreet groups, but lie on a continuum with overlapping pressure and acoustic measurements, as this thesis has demonstrated.

5.6.3 Baugh et al (1988)

In their prospective study of subjects who were candidates for T-E puncture procedures, Baugh et al investigated 21 consecutive laryngectomees by air insufflation testing and pressure measurements, correlating their pre-operative intraoesophageal pressures with the post-operative levels of speech fluency achieved.

Three levels of speech fluency were identified and they found that nine patients developed fluent speech; five were non-fluent and seven were non-speakers. These corresponded to patients in whom, pre-operatively, pressures of < 20 mmHg, 21-39 mmHg, and > 40mmHg were respectively recorded. The variation of pressure recordings was greatest for non-speakers and least for good speakers.

Unfortunately, duration of phonation during air insufflation was not reported. However, the pressure measures which were reported in this study correspond to the measures of hypotonicity, tonicity and spasm in this thesis. Also, the degree of spasm was not subdivided, as in this study, into hypertonicity and spasm - depending on the degree of tone present in the reconstructed pharynx. The advantage of this subdivision lies in the fact that the two groups might need to be managed in a different manner. It is not the case that all subjects who exhibit increased tone, for example, need further surgery; a significant proportion may be managed therapeutically, as indicated in 5.2.4.

5.6.4 Salmon (1965)

In her unpublished Ph.D. study, Salmon attempted to examine sound pressure, intrapharyngeal pressure, pressure within the P-E segment and intraoesophageal pressure during cinefluoroscopy under three conditions - at rest; during air intake for speech; and during phonation. She concluded that..

> "using esophageal pressure measures as a diagnostic tool for differentiating between "good" and "poor" speakers is of questionable validity."(p 86, 87)

Her methodology, however, was questionable as a water-filled sensing element enclosed in a bag was used to detect pressure within the P-E segment, and many subjects could not tolerate this as the diameter bore was large (3mm) and somewhat rigid, being constructed of tygon. Additionally, she commented that there was..

> "a possibility of invalid pressure measurements recorded from the bag due to alterations in its position." (p 87).

With the more sophisticated pressure transducer and the flexible radio-opaque catheter used in this study, these problems were overcome.

5.6.5 Diedrich and Youngstrom (1966)

In their textbook "Alaryngeal Speech," as discussed in the literature review, they concluded from their study that:

> "esophageal speech skill following a laryngectomy was not related to morphological function"....and that... "These results would support those who believe that psychological factors are probably more important to the development of good speech than anatomical factors." (p.61).

However, as discussed in 2.3.2., the study did not examine morphological function; rather it described the shape of the articulatory structures and the pharynx and P-E segment during speech manoeuvres in subjects who had acquired oesophageal speech. It is therefore imprudent to conclude, without examining a matched group of "failed" speakers, that the morphological function has no part to play in oesophageal speech acquisition. Undoubtedly, the psychological state of the laryngectomee is important, but it is often difficult to assess in a non-speaking laryngectomee whether his depression/despondency is the cause or the effect of failing to speak. Physiological and anatomical factors should be excluded before assuming that the difficulty in communication is due to psychological factors.

5.7 Evaluation of Methodological Technique.

The use of air insufflation testing to measure the opening pressure of the P-E segment presents some problems.

Firstly, some subjects (1% of the total number examined in this study) cannot tolerate the catheter passed trans-nasally due to hypersensitivity and/or blockages in the nasal passages, which prevent easy transmission of the catheter through the nose. This is an important consideration where time may be of some consideration and where the population of available subjects for study may be limited. Some subjects are able to swallow the catheter instead and this, when checked using videofluoroscopy, proved to be equally acceptable as a method of positioning the catheter and pressure and acoustic measurements could be successfully obtained. The problem becomes less with the increasing expertise of the examiner who passes the catheter. With practice, the trauma to the subjects is minimal.

It is difficult to evaluate the effect that the presence of the catheter has on phonation. By using good oesophageal speakers as a target group, there is a standard for comparison. Surprisingly, perhaps, the presence of the catheter does not seem to have such a marked effect on phonation as one might expect. The good oesophageal speakers are all able to produce voice during air insufflation which is similar in quality to their own oesophageal speech, as subjectively judged by the examining team and reported by the subject. The main difference is, naturally, in duration of achievable sound as air insufflation eliminates the need for the subject to pause to re-inflate the oesophagus prior to phonation.

Undoubtedly, the presence of the catheter has an effect on the vibratory characteristics of the P-E segment; but, as this is consistent across all subjects examined, the difference in pressure measurement achieved cannot be attributed to this alone. There is some variability in successive recordings within the same subject. This may be due to the practice effect of air insufflation testing. Initially, the subject may be tense in anticipating the passage of airflow into the oesophagus, and this might affect performance, but with three successive recordings being taken for each subject, the change in recordings can be assessed and there is a less than 5% variability in successive recordings within the same subject.
There are some factors which could not be controlled in this experiment. The subject's psychological state is difficult to assess. Although reassured continually during assessment, naturally some subjects are more tense during air insufflation testing, in anticipation of what is to occur. However, the more obviously anxious subjects are randomly occurring across all the groups, so

this cannot explain the statistically significant differences between the groups achieved when analysing both pressure and duration characteristics.
An attempt was made to correlate the groups with the type of laryngectomy surgery initially achieved, particularly with the type of pharyngeal closure undertaken, but this proved to be impossible due to the wide variety of sources from whence the subjects came and the paucity of operative notes available.

The amount and complexity of equipment used in this study, together with the technical help, was fortunately available from the resources within a large teaching hospital; particularly from the departments of radiology, medicine and clinical monitoring, but there needs to be a simplified form of pressure-tip transducer developed for easier clinical recordings to render this assessment more accessible to clinicians who are working in settings where such sophisticated apparatus may not be available. (See 6.1.1)
Signal analysis of post-laryngectomy speech is notoriously difficult for the reasons previously given (See section 3.9.2.) and it may be that the Lynex software system could be further improved to take account of the problems in the frequency characteristics of laryngectomy speech. (See following section, 6.1.2).

Despite these limitations, the accurate use of air insufflation testing as a prognostic indicator for the development of post-laryngectomy speech would appear to be of value and, clinically, could save time, resources and frustration on behalf of both subject and therapist by suggesting appropriate rehabilitation procedures for each subject. This could avoid the unrealistic planning of, for example, oesophageal speech therapy or of surgical speech rehabilitation for subjects who are incapable, physiologically, of achieving this. Additionally, the results of this study indicates the importance of directing attention to the type of surgery performed during laryngectomy, particularly the type of closure of the pharynx. (see 6.3).

5.8 Implications of this Assessment.

Grunwell (1983) described therapeutic intervention with clients who have phonological problems as consisting of three levels: premises, principles and procedures.This model may also be applied to laryngectomy therapy. Premises, she states, are:

> "primary, in that they state the fundamental theoretical framework underlying a therapeutic approach; it is from premises that a strategy of intervention is derived."

On the basis of the findings of this thesis, it would seem that the fundamental premise to which clinicians adhere in laryngectomy treatment is incorrect.

This false premise is that all laryngectomees can achieve oesophageal speech. Indeed, it can be demonstrated that, even in centres offering a particular expertise in laryngectomy rehabilitation, a certain percentage of laryngectomees will not succeed.

From this false premise, a number of false principles arise:

Firstly, it is believed that, by prescribing an artificial larynx, the laryngectomee will not develop oesophageal speech. Consequently, many laryngectomees are actively discouraged by clinicians from becoming an effective communication aid user.

Secondly, the belief is often held that any laryngectomee who practises "hard enough" can achieve oesophageal speech. Thus clinicians encourage laryngectomees to continue to strive for oesophageal speech long after this is desirable. This can result in severe frustration and depression.

Thirdly, there is a tendency to assume that psychological factors are the paramount reason for failure to achieve speech. Consequently, clinicians are reluctant to examine anatomical/physiological barriers to speech acquisition and fail to perceive that the psychological problems may overly, and mask, underlying anatomical/ physiological problems.

From the reports of laryngectomees, some of whom took part in this study, the above would appear to be the principles upon which many clinicians operate. It would appear that many clinicians simply undertake exercises to promote speech. Generally, the decision to start/stop therapy does not seem to be based on objective evaluation of progress and factors which may be hindering progress, but rather occurs when the laryngectomee feels a plateau has been reached in communication skills - even when this is not at an acceptable level to client or clinician. Due to inexperience and a reluctance to use (or a lack of awareness of) formal and objective assessments in this field, many clinicians tend to accept an unusually low level of achievement with this population and discharge the laryngectomee from active therapy.
This thesis has suggested the development of a proactive and principled model for laryngectomy treatment. On the basis of this study, the following premises are proposed:

1. There is a continuum of tonicity in the reconstructed pharynx (and P-E segment) of laryngectomees. This tone ranges from hypotonicity (i.e. flaccidity) through tonicity (where good speech is achievable), to spasm (excessive tonicity).

2. The ability to acquire oesophageal and/or T-E speech is directly dependent on the muscle tone in the reconstructed pharynx. It would appear that this is probably related to the type and extent of surgical closure used.

3. Approximately 30% of laryngectomees, through no fault of their own, will not have the requisite physiology to achieve speech. They will need further surgery to the pharynx to remediate this.

4. Psychological problems are not necessarily the CAUSE of failure to develop oesophageal speech post-laryngectomy, but may be the EFFECT of not being able to adequately communicate as a result of anatomical/physiological problems.

From these premises, the principles for therapy emerge:

The first principle should be that intervention is based on a thorough assessment of the subject's existing and potential speech skills, from an examination of the anatomy/physiology of the reconstructed pharynx.Ideally, videofluoroscopy and air insufflation testing should be used.

The second principle is that intervention and treatment goals are defined by the initial assessment. As healing occurs, the reconstructed structures will alter, so regular re-assessment and re-evaluation will be necessary.

From these principles, good procedures emerge:

Either the laryngectomee does not have the prerequisite for oesophageal speech acquisition nor for T-E speech -i.e. the reconstructed pharynx lacks the capacity to pass air and vibrate to produce sound. In this case, the need for an artificial larynx and/or further surgical intervention is needed. The decision for further surgery should be based on the results of the objective testing of the potential of the pharynx to produce sound.

Alternatively, initial assessment indicates good potential for acquiring oesophageal speech and/or T-E speech. In this case, therapy involves directed practice through a hierarchy of skills with oesophageal speech acquisition as the aim. In this type of programme, until sound control has reached a certain level of consistency, it should not be used for conversational speech, or malpractices may emerge. The laryngectomee should thus be instructed in the use of an artificial larynx and encouraged to use this aid whilst he is perfecting his oesophageal speech to an acceptable level of consistency.

With a such a sound theoretically-based framework of underlying premises and principles, decisions as to when to commence and cease therapy become rational and explicit. When a thorough assessment is undertaken, realistic aims and objectives can be stated and expectations can be defined and realised. Once these have been achieved, termination of therapy is easier for both clinician and laryngectomee and the decision becomes mutual, rather than leaving the laryngectomee frustrated and depressed with his skills and the clinician feeling inadequate.

The assessment developed in this thesis and the implications for clinical management may be summarised as in the chart (Table 26).

Firstly, using videofluoroscopy, the subject may be classified into the broad groups: hypo-tonic-hyper-spasm on a continuum, depending on the degree of tonicity in the reconstructed pharynx; or, if a stricture exists, this may be identified. It is possible to do this without video- fluoroscopy, as has been previously described (Perry and Edels, 1986), although this involves subjectivity in interpreting the findings. In order to corroborate this assessment, pressure and duration measures are taken which should confirm the subject's position in the grouping. Further confirmation may then be gained from examining analysis of the acoustic signal as earlier described (4.4). Once the subject is thus evaluated, clinical intervention and management proceeds as indicated in Table. 26.

PRESENTATION OF LARYNGECTOMEE

ASSESSMENT

(VIDEOFLUOROSCOPY)

DIAGNOSIS

MEASUREMENTS	HYPOTONIC	TONIC	HYPERTONIC	SPASM
Mean Pressures recorded (mmHg)	19.30	26.35	45.56	78.00
Mean Duration (secs)	8.70	6.90	2.60	0.85
Acoustic Signal	Small amplitude signal. Continuous sound.	Strong amplitude signal. Continuous sound.	Intermittent signal. Background noise.	Sparse signal. Low ratio of signal to silence.
INTERVENTION	Raise tone in P-E segment.(collar or pressure band)	Oesophageal speech training or TEP (if desired)	Inhalation for oesophageal speech. Low pressure prosthesis for TEP	Surgery (Myotomy)

Table 26. Chart to indicate intervention for post - laryngectomy speech restoration

In order for the hypotonic subject to realise speech, it is necessary to increase the tone in the P-E segment. This may be achieved by using digital pressure on the outside of the neck by pressing on the skin or by the use of a tight collar or pressure band. The tonic subject will be able to achieve speech with practice/rehearsal in a structured manner, or by undergoing a T-E puncture if so required. The hypertonic subject may achieve oesophageal speech with the use of inhalation as a method of air-charging, as this demands reduction in the tone of the P-E segment (see 1.3.4.). It may, however, be the case that the patient will not achieve fluent speech if the tone within the P-E segment is too great to relax and allow good air flow.

If a T-E puncture is chosen, best results will ensue after a surgical myotomy, but if the patient refuses this then it may be possible to manage the problem with a low resistance prosthesis.

The subject who is classified as being in the spasm group will need a surgical myotomy for any speech to result. Once this is achieved, the subject will change group and may be more accurately placed in the tonic group, or possibly on the hypotonic/tonic end of the continuum if the myotomy has been excessively severe(see Table 25). Speech therapy may now be appropriate if oesophageal speech is required, or the subject will now be suitable for a T-E procedure and the management will be as for the earlier - described tonic subjects.

CHAPTER 6 : FUTURE RESEARCH PROPOSALS.

From the results of this thesis, there are implications for future research.

6.1 Improving Methodology.

There are aspects of the methodology of this project that may be improved, although the nature of the required recordings are complex and therefore there are limitations to the simplicity in the methodology which can be achieved. As stated in the Discussion section previously, if there is to be an application of this assessment in Centres where clinicians do not have access to complex apparatus, then there has to be a development of portable, accessible equipment.

6.1.1 Improving Pressure Measurements.

As discussed in 5.4., one of the ways of simplifying the pressure recordings is to use a pressure-tip transducer to measure the opening pressure of the P-E segment at onset to phonation.

This work is currently in progress, in collaboration with researchers from the Royal National Throat, Nose and Ear Hospital, London, and a pilot study into the use of such a transducer gives good intra-subject reliability. If this apparatus becomes commercially available, then it should make the objective assessment of laryngectomees considerably easier. The apparatus will consist of a flexible radio-opaque catheter, at the tip of which is embedded a pressure transducer, and two side connections from the catheter which will link to both a chart recorder and to the compressed air source (i.e. air cylinder) which is used for air insufflation testing.

6.1.2 Improving Acoustic Analysis.

Lynex, although useful, has shortcomings as a tool for analysing the frequency of this type of sound that is so complex.

Advancement of the sound analysis might occur with the use of a computer modelling technique for sound analysis. Cepstrum pitch determination is an extremely complex mathematical technique whereby the whole sound is recorded, and the Fourier analysis (see 3.6.3.) is applied, as in this project. The Cepstrum is the inverse Fourier transform of the logarithm of the (short term) power spectrum of the signal. The spectrum of the voiced speech signal is the complex product of the glottal (or pseudoglottal) source spectrum and the vocal tract transfer function. The Cepstrum method begins by filtering out the highest frequency components of the speech signal. An analog-to-digital converter is then used to convert the amplitude of every point in the (still quite complex) waveform to numerical values at a rate of perhaps 10,000 measurements per second. The wave is then represented by a string of numbers upon which all further operations are done. By the logarithmic operation, the voice source and the vocal tract transfer functions are separated. The next step is to derive the spectrum of the waveform,i.e. discover its frequency components and their amplitudes. The mathematical bases of this technique are extraordinarily complex, but the final outcome is to mathematically remove the overlying harmonics/formants and then the underlying fundamental frequency can be extracted.

Researchers at Cambridge University Department of Electronics and Signals Analysis (where the Lynex was designed) have expressed considerable interest in this project and have now begun a collaboration in analysis of the vowel sounds recorded in this project by using Cepstrum as an analysis tool. The scope of this work is well beyond this project and complex signal analysis will need to be undertaken which may take many years of detailed research to complete.

6.2 Flow During Phonation.

It may be valuable to analyse flow during phonation which was not possible to do with the apparatus available in this project. A mask-microphone, linked to a flow meter, would remove the necessity for subjects to hold a gum shield/flow meter which is unpleasant to use and, in this pilot study, was shown to be unviable as a method of recording. This may also show characteristic patterns across the groups. One might hypothesise that there would be increased flow rates in hypotonic subjects when compared with spasm subjects, due to the difference in resistance in the pharynx, but this remains to be tested.

6.3 Amplitude of Recorded Sound.

This would be a valuable parameter to investigate as, subjectively, the clinicians in this study noted that the subjects in the hypotonic group exhibited quieter sound production on the vowel /a/ than those in the spasm group, and this could be confirmed using the Lynex samples where the amplitude of recorded sound was smaller in the hypotonic subjects than in that of the spasm subjects - who exhibited shorter sound of greater amplitude. The available apparatus did not permit the measurement of the amplitude in this study but this could theoretically be possible if appropriate software was available.

6.4 Improving Surgical Closure of the Pharynx.

For the four subjects in this study who underwent surgical myotomy as a secondary procedure, a "long" myotomy was undertaken. i.e. the muscle cut was made from the base of tongue to the oesophageal inlet, as the videofluoroscopy studies showed an extensive length of the pharyngo-oesophagus in spasm. The post-operative pressure and duration measurements have already been discussed (4.2.5 and 4.3.4.) and, as has been stated, these were on the hypotonic/ tonic end of the continuum of recordings; perhaps because the myotomy was extensive. From these results, the aim of surgically producing tone without constriction in the reconstructed pharynx at the time of initial laryngectomy was identified.

As a result of this project and its outcome, the surgical technique for laryngectomy has considerably altered at Charing Cross Hospital. All subjects who undergo primary laryngectomy surgery at Charing Cross Hospital currently undergo a myotomy (see 1.3.6.2.) after the surgical closure of the pharynx. Experimentation with different methods of myotomy (i.e. altering the site of muscle cutting) has shown that optimum results occur with a "short" myotomy i.e. only severing the upper circular fibres of the oesophagus. After meticulous closure of the pharynx, the surgeon inserts his finger into the reconstructed pharyngo-oesophageal junction and, where he feels the circular muscles of the upper oesophagus grip tightly around his finger, he severs the fibres to a level of 1-2 cms. This has the result, clinically, of creating tone without spasm or constriction. The subjects (now 50 primary laryngectomees) all began to produce effortless oesophageal speech by the tenth post-operative day, successfully used T-E speech, and there have been no fistulae formation and remarkably few complications to date (Milford et al, 1988).

The question remains whether it is the surgical myotomy which is producing such good results or whether it is the early acquisition of speech (through the use of primary T-E puncture) which prevents hypertonicity or spasm from developing. It may be that the early use of the P-E segment for speech "re-educates" the muscles of the pharynx for speech and thus prevents the situation (which is normal for the laryngeal speaker) where the muscles resist air flow and are used extensively for swallowing. Prior to the use of T-E puncture as a primary procedure, laryngectomees often had many months of silence in which the muscles of the reconstructed pharynx were used for the vegetative functions of eating and swallowing only. This may have resulted in the patterns being re-established for swallowing but, simultaneously, air flow for speech being restricted.

It has been suggested (McGarvey and Weinberg, 1984) that the oesophagus of normal (laryngeal) speakers resists airflow; thus, McGarvey and Weinberg concluded, it is the abnormal status of the laryngectomee that allows him to produce oesophageal speech. However, this study has indicated that the majority of laryngectomised subjects lack this skill. Perhaps they have a "normal" reconstructed pharynx - i.e. one which resists air flow. Singer and Blom (1985) state that the two populations are not directly comparable, but in their reply, McGarvey and Weinberg indicate their awareness of this fact but that they continue to view the "functionally normal esophageal response" as the that which resists air flow during insufflation testing.

It is possible to test the theory that early speech acquisition prevents muscle spasm from developing by undertaking a laryngectomy without a simultaneous myotomy and then assessing whether the same good speech result ensue.
It is planned that this will be done in the next series of six subjects who undergo laryngectomy surgery in our Unit and it is hoped that this might result in a better understanding of the phenomenon of "spasm" and the reason for its occurrence.
The further research stimulated by this thesis will enable us to investigate the occurance of spasm, as well as to asssess the optimum surgical treatment of this condition. This may result in our ability to prevent its occurance in newly-laryngectomised subjects, and to treat this problem effectively in pre-exisiting subjects who are experiencing poor/no speech acquisition. The future for speech restoration after laryngectomy has never seemed so promising.

APPENDIX I.

INTAKE DATA SHEET.

The enclosed intake data sheet is used to gather background information on each laryngectomised subject who was accepted for this study. It was sent to each subject to complete and bring with hm to the videofluoroscopy clinic.

Instructions:

Please complete all questions.

Circle a yes or no answer.

Please give complete details if requested.

If extra space is needed, please use reverse side of sheets.

1. Name:

2. Address:

   Tel.

3. Sex:
4. Age:
5. Education:
6. Occupation:
7. Spouse:
8. Children:
9. What are your interests/hobbies?
10. Surgeon:
11. General Practitioner:
12. Date of Surgery:
13. Where:
14. How long hospitalised:
15. Any complications:
16. Extent of surgery, if known:

17. Did you receive any training in learning to speak before your surgery? YES NO
    By whom?
18. What means of communication did you use immediately after your operation, before you learned to speak:
    Artificial larynx/whispered speech/gesture/writing/ other(explain)
19. How soon after your surgery did you begin speech therapy:
20. Who taught you during your first sessions:
    (a) Self-taught (b) Speech therapist (c) Nurse (e) Another laryngectomee
    (f) Other (explain)
21. If self-taught, what materials did you use?
22. How many hours of therapy have you had to date?
23. How frequently were the sessions held?
    By whon?
    Where?
24. How would you rate your speech?
    (a) Excellent (b) Good (c) Fair (d) Bad
    (e) Need more training
25. Do you have a hearing loss? YES NO
26. Do you have a reduced sense of taste? YES NO
27. Do you have a reduced sense of smell? YES NO
28. Do you have difficulty in swallowing? YES NO
29. Did you have X-ray therapy? YES NO
30. If yes, how many sessions?
31. Is your tongue stiff? YES NO
32. If you use a speech aid, what type is it?
33. How soon after voice training wre you able to
    Say your first sound?
    Say your first word?
    Say your first phrase?

APPENDIX I

Say your first sentence?

Hold a conversation?

Any comments on difficulty in the above.

34. Under what conditions do you speak best?
35. Under what conditions do you speak poorest?
36. Please circle the rating level which corresponds to your speech: (See Wepman Scale, Table 2)

Please tick appropriate statement(s)

I wish to attend the assessment clinic ....

I do not wish to attend the assessment clinic ....

Spouse wishes to attend with me ....

No financial help needed to attend ....

Financial help is needed to attend ....

Please fill out this form and return to: Alison Perry,
Chief speech therapist, Charing Cross Hospital, Fulham Palace Road, London W6 8RF.

APPENDIX II.

THE LYNEX : DESCRIPTION OF SOFTWARE.

The Lynex is an 18 bit stereo sampling system. It is 8 voice polyphonic at full bandwidth and, when connected to an Atari 520ST and a compatible sequencing package, it will form the heart of a studio recording system. It was designed in Cambridge University Electronics Department for use in the recording music business.

Unlike other software - based sampling systems, the Lynex is configured as an Atari GEM desktop accessory (see fig 25). This makes upgrading the system easy to undertake, simply by obtaining the latest Lynex system disk.

Once the Lynex disc is loaded, the screen shows as in fig. 56.

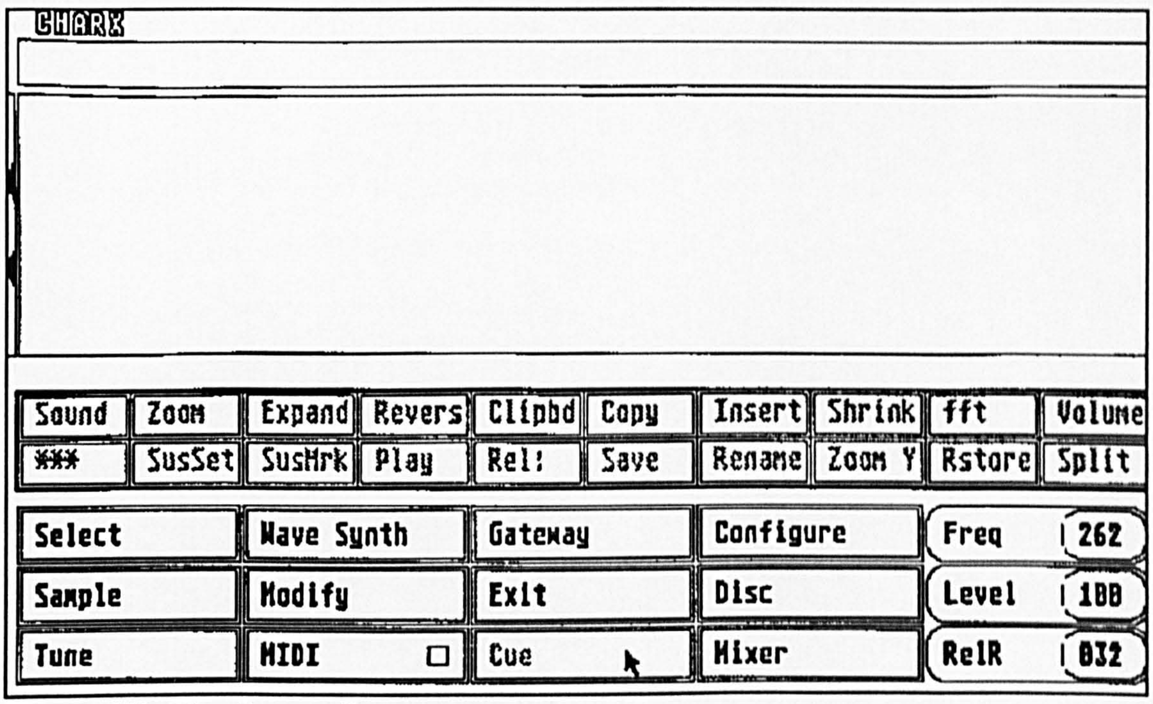


Figure 56. The Lynex Screen.

The screen is divided into five main areas:

1. The information bar at the top of the screen.
2. The scope with the scroll bars on either side.
3. The EDIT page functions.

4. The PAGE select boxes (black).
5. FREQUENCY and LEVEL displays.

Only functions written in black are available for selection.

Initially, one uses the mouse pointer over the "select" box and produces the select page (fig 57).

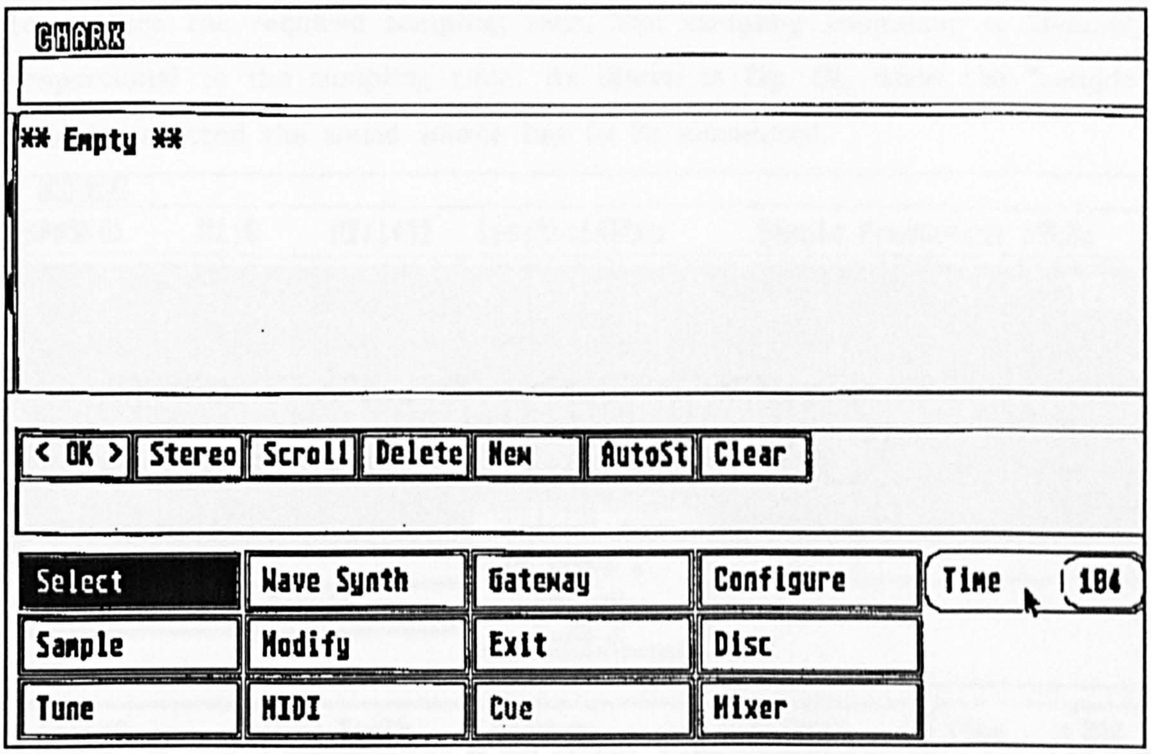


Figure 57. The <u>Select</u> mode of the Lynex Programme.

This is where the Lynex is told how much memory to allocate the sample, and the name to allocate for file purposes. Selecting the "NEW" icon produces the following message:

"Please set length in 1/10s at 50 Hz and then press <OK> Anything else cancels."

One can select <OK> and set the sample by using the TIME box. This takes the display to the maximum amount of free time left in the memory and can be set at 1/4 of the desired length of sound to be sampled. e.g. setting the "time" box at 005 means 2 sec of sound can be sampled.

The Sample page (fig 58) is where the sound is sampled into the memory of the Lynex unit. At the maximum sampling rate of 50 KHz, 10.4 seconds of mono (5.2 sec of stereo) sound can be sampled. At a sampling rate of 12 KHz, this allows 43.2 sec of mono or 21.6 sec of stereo sound to be sampled. Lynex always samples at 50 KHz and then downsamples in the digital domain to produce the required sampling rate. The sampling frequency is inversely proportional to the sampling time. As shown in fig. 58, when the "sample" page is selected the sound source has to be connected.

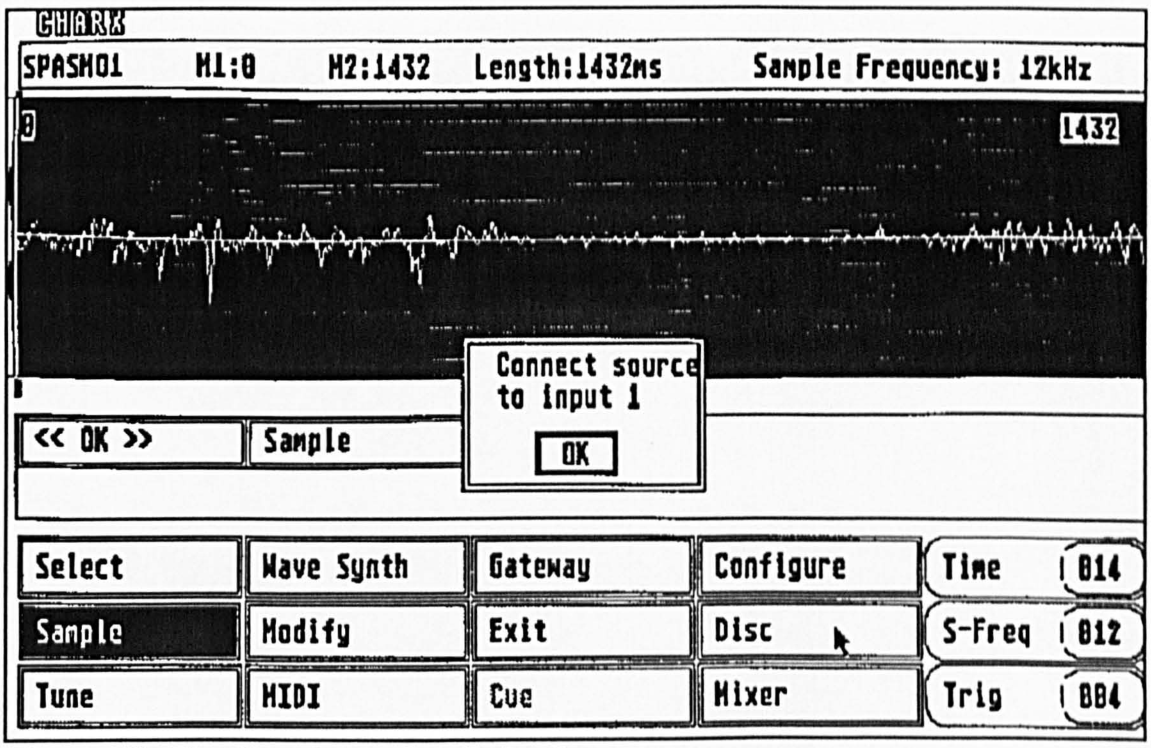


Figure 58. The Sample mode of the Lynex programme.

With this analysis system, the sound replayed from the Racal recorder could be heard (via the amplifier), sampled, replayed and then stored on a floppy disk with hard disc back - up, once a satisfactory sample was obtained. This could then be replayed at any time from the Lynex system and a hard copy (via the Laser printer) obtained.

APPENDIX II

Once in the memory, the sound may be edited. With the Lynex system, the duration of sound could be accurately recorded as an oscilloscope trace. Sampling at 12 KHz, 43.2 seconds of mono sound could be stored on each file.

The programme has a FFT facility on the EDIT page. In fig. 59, Spasm 01 is the name of the sound. M1 and M2 are cursors which mark the area of the scope to be analysed which is 649 (M2) - 9 (M1) ms i.e. 640ms. Length 1432 is the length of the sound. Sample frequency: 12KHz is the frequency at which the sound was sampled.

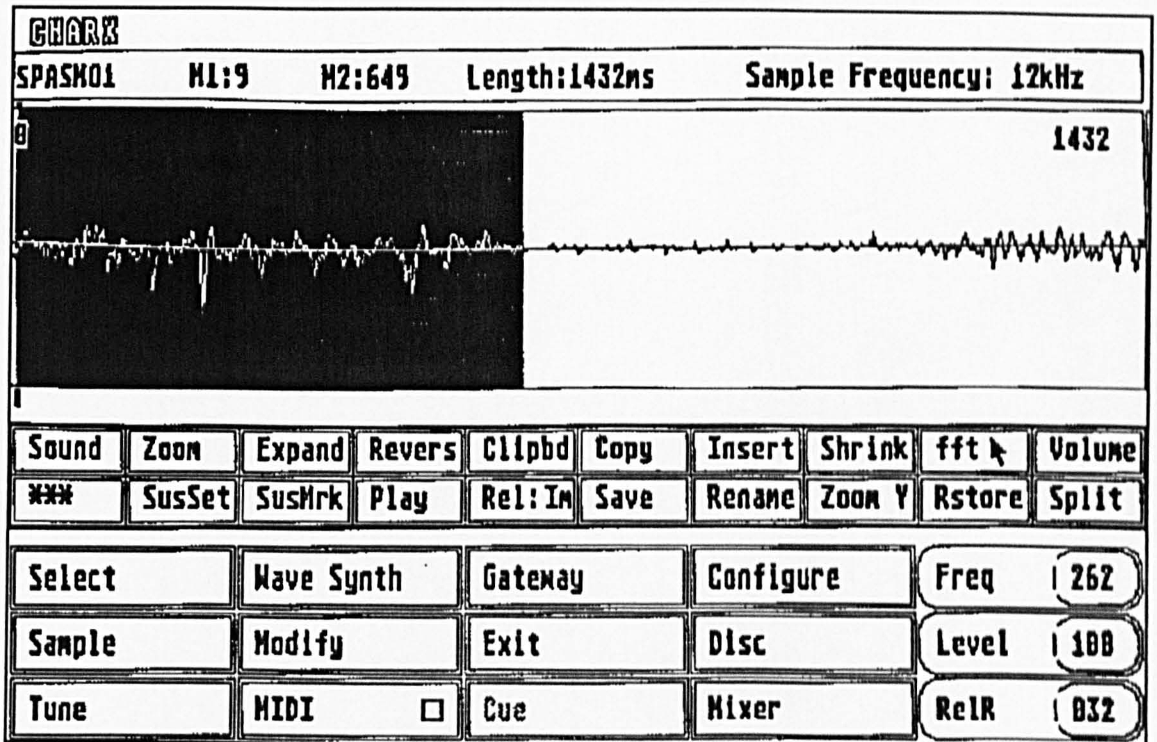


Figure 59. M1 & M2 are the cursors marking the black portion of the screen which is the sample of sound to be analysed.

After experimentation, it was decided to store 2080 ms (i.e. just over 2 sec) of possible sound for FFT analysis. As FFT analysis relies on steady state sound, a sample of 100ms was then used as a maximum unit of sound on which FFT could be deployed.

Beyond this time, there was a possibility of the sound no longer being steady state as the articulators may have moved, thus altering the formants of the sound. The 100ms marked area of sound may then be treated to FFT analysis as shown in Fig 60 which demonstrates a completed FFT.

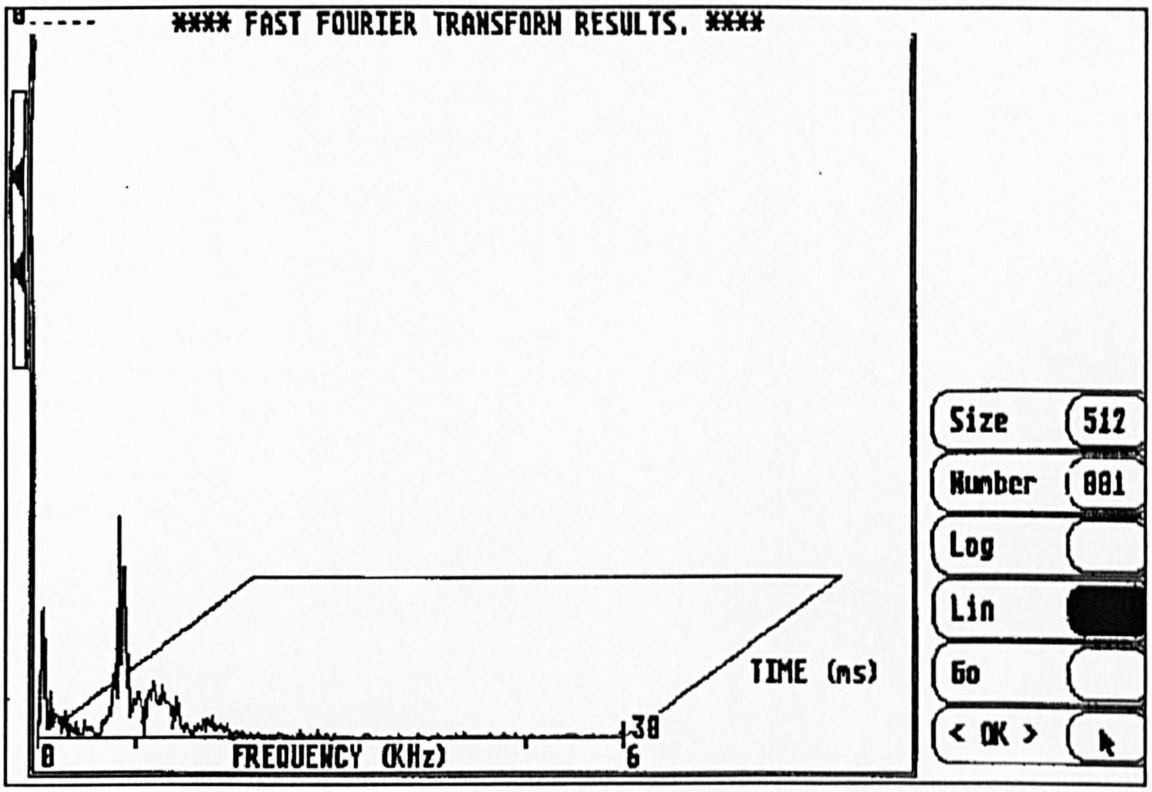


Figure 60. Completed FFT.

SIZE determines the resolution of the analysis, i.e. the number of frequency bands into which the sound is divided. The higher the resolution, the more accurate the plot. There are eight possible resolution settings from 8 to 1020 and, after experimentation, the setting of 512 was used as the optimum level for the sounds sampled. There was not any significant advantage over the setting at a higher level of 1020 and the computer time needed to undertake the calculation was considerably greater for the higher number of frequency bands.

APPENDIX II

NUMBER determines the number of times that the sound is analysed during the time scale of the marked area. The maximum number is 10 but, in order to assess the dominant formant frequencies, a plot of 1 was used. LIN shows that it is a linear, rather than logarithmic, plot that is used.

GO starts the analysis.

In its original form, sampling for the FFT was 0-25 KHz. However, for reasons explained earlier (see 3.9.2.2.), the software was re-written in order to sample and display formants in the range 0-6 KHz.

APPENDIX III

TABLE 7

Subj No	Group	Sex M/F	Age	Press.1 mmHg	Mean Press.1	Press.2 mmHg	Mean Press.2	Dur.1 Sec.	Mean Dur.1	Dur.2 Sec.	Mean Dur.2
1	1	M	77	26	24.67	50	48.00	6.00	6.33	4.00	4.33
1	1	M		30		48		7.00		3.00	
1	1	M		18		46		6.00		6.00	
2	1	F	69	10	18.33	50	55.00	7.00	13.67	7.00	14.00
2	1	F		20		60		18.00		15.00	
2	1	F		25		55		16.00		20.00	
3	1	M	53	28	24.00	32	35.33	10.00	8.67	4.00	5.33
3	1	M		16		38		7.00		6.00	
3	1	M		28		36		9.00		6.00	
4	1	M	71	14	14.67	28	35.33	20.00	21.00	10.00	8.00
4	1	M		12		38		22.00		8.00	
4	1	M		18		40		21.00		6.00	
5	1	F	61	16	18.00	60	64.00	6.00	7.00		
5	1	F		20		64		7.00			
5	1	F		18		68		8.00			
6	1	M	65	40	29.33	48	44.00	14.00	18.00	8.00	8.00
6	1	M		20		50		21.00		8.00	
6	1	M		28		34		19.00		8.00	
7	1	M	56	12	17.33	34	32.00	1.64	2.75		
7	1	M		20		32		3.00			
7	1	M		20		30		3.60			
8	1	M	67	26	22.33			6.00	6.73		
8	1	M		21				6.20			
8	1	M		20				8.00			
9	1	M	60	8	12.00			5.20	5.30		
9	1	M		10				5.50			
9	1	M		18				5.20			
10	1	F	69	10	12.00	28	22.67	3.00	3.07	2.20	1.93
10	1	F		12		18		3.20		2.10	
10	1	F		14		22		3.00		1.50	
11	1	M	52	18	17.33	28	31.33	0.60	0.60	1.50	2.00
11	1	M		16		26				2.00	
11	1	M		18		40				2.50	
12	1	F	69	18	16.00			12.00	10.60		
12	1	F		14				10.00			
12	1	F		16				9.80			
13	1	M	79	18	19.67	28	28.67	10.00	10.67		
13	1	M		16		26		13.00			
13	1	M		25		32		9.00			
14	1	M	80	20	18.67	60	49.33	5.00	4.90		
14	1	M		12		48		4.50			
14	1	M		24		40		5.20			
15	1	M	68	28	24.67			7.20	6.30		
15	1	M		26				4.70			

APPENDIX III

TABLE 7

Subj No	Group	Sex M/F	Age	Press.1 mmHg	Mean Press.1	Press.2 mmHg	Mean Press.2	Dur.1 Sec.	Mean Dur.1	Dur.2 Sec.	Mean Dur.2
15	1	M		20				7.00			
16	2	M	65	22	24.00			9.00	9.07		
16	2	M		26				5.20			
16	2	M		24				13.00			
17	2	M	49	30	30.67			7.00	7.00		
17	2	M		30				5.00			
17	2	M		32				9.00			
18	2	M	82	24	22.67			7.00	7.67		
18	2	M		22				8.00			
18	2	M		22				8.00			
19	2	F	70	30	37.33			11.00	8.33		
19	2	F		34				6.00			
19	2	F		48				8.00			
20	2	M	66	16	19.33			4.50	5.17		
20	2	M		20				5.00			
20	2	M		22				6.00			
21	2	M	61	18	24.67			11.00	9.67		
21	2	M		32				10.00			
21	2	M		24				8.00			
22	2	M	62	28	28.00			8.50	8.83		
22	2	M		30				10.00			
22	2	M		26				8.00			
23	2	M	62	32	30.67			3.50	3.13		
23	2	M		30				3.00			
23	2	M		30				2.90			
24	2	M	69	26	25.33			7.00	7.17		
24	2	M		24				8.00			
24	2	M		26				6.50			
25	2	M	66	14	16.00			6.00	5.83		
25	2			18				7.00			
25	2			16				4.50			
26	2	M	40	26	20.67			8.50	7.33		
26	2			16				5.00			
26	2			20				8.50			
27	2	M	36	14	15.33			4.00	4.07		
27	2			12				4.20			
27	2			20				4.00			
28	2	M	67	30	28.00			5.80	8.93		
28	2			28				9.00			
28	2			26				12.00			
29	2	M	61	36	32.00			7.50	7.97		
29	2			40				8.00			
29	2			20				8.40			
30	2	F	69	40	40.00			14.00	13.33		
30	2			38				13.00			

APPENDIX III

TABLE 7

Subj No	Group	Sex M/F	Age	Press.1 mmHg	Mean Press.1	Press.2 mmHg	Mean Press.2	Dur.1 Sec.	Mean Dur.1	Dur.2 Sec.	Mean Dur.2
30	2			42				13.00			
31	2	M	66	30	28.67			4.00	4.13		
31	2			38				4.50			
31	2			18				3.90			
32	2	M	54	15	24.33			2.00	2.03		
32	2			28				2.10			
32	2			30				2.00			
33	2	M	67	26	28.00			2.50	2.77		
33	2			22				2.80			
33	2			36				3.00			
34	2	M	52	32	29.00			8.20	8.57		
34	2			30				9.50			
34	2			25				8.00			
35	2	M	63	22	22.00			9.00	8.83		
35	2			20				8.50			
35	2			24				9.00			
36	3	M	60	34	38.00			3.50	3.83		
36	3			38				5.00			
36	3			42				3.00			
37	3	M	54	42	53.33			5.00	4.17		
37	3			46				4.00			
37	3			72				3.50			
38	3	M	48	38	40.00			0.62	2.74		
38	3			32				1.60			
38	3			50				6.00			
39	3	F	67	46	54.67			0.68	1.03		
39	3			66				0.55			
39	3			52				1.85			
40	3	F	67	46	51.33			1.30	1.38		
40	3			48				1.53			
40	3			60				1.32			
41	3	M	50	55	50.33			3.00	3.67		
41	3			46				3.00			
41	3			50				5.00			
42	3	M	68	50	50.67			2.00	1.67		
42	3			50				0.93			
42	3			52				2.08			
43	3	M	67	38	43.33			1.78	1.76		
43	3			48				2.00			
43	3			44				1.50			
44	3	M	75	42	40.00			0.97	1.41		
44	3			38				1.83			
44	3			40				1.42			
45	3	M	82	36	43.33			2.05	1.52		
45	3			56				1.50			

APPENDIX III

TABLE 7

Subj No	Group	Sex M/F	Age	Press.1 mmHg	Mean Press.1	Press.2 mmHg	Mean Press.2	Dur.1 Sec.	Mean Dur.1	Dur.2 Sec.	Mean Dur.2
45	3			38				1.02			
46	3	M	72	54	49.33			1.08	1.29		
46	3			46				1.44			
46	3			48				1.34			
47	3	F	57	36	40.00			0.79	2.62		
47	3			38				2.08			
47	3			46				5.00			
48	3	M	61	34	38.00			8.00	7.73		
48	3			36				7.00			
48	3			44				8.20			
49	4	M	63	64	74.67			0.54	0.75		
49	4			100				0.50			
49	4			60				1.20			
50	4	F	70	80	74.67	30	37.33	0.59	0.59	11.00	8.33
50	4			74		34		0.68		6.00	
50	4			70		48		0.50		8.00	
51	4	M	67	40	40.00			0.50	0.63		
51	4			38				0.90			
51	4			42				0.50			
52	4	M	66	100	83.33			1.50	0.66		
52	4			50				0.30			
52	4			100				0.19			
53	4	M	61	82	74.00			0.89	0.72		
53	4			70				0.76			
53	4			70				0.50			
54	4	F	76	82	87.33			1.50	1.23		
54	4			100				1.50			
54	4			80				0.70			
55	4	M	49	100	99.33	30	30.67	0.88	0.98	7.00	7.00
55	4			98		30		1.04		5.00	
55	4			100		32		1.03		9.00	
56	4	M	85	100	99.33			0.19	0.37		
56	4			98				0.22			
56	4			100				0.69			
57	4	M	55	68	66.00			1.26	0.88		
57	4			70				0.72			
57	4			60				0.66			
58	4	M	62	90	90.00	36	32.00	1.39	0.71	7.50	7.97
58	4			80		40		0.34		8.00	
58	4			100		20		0.40		8.40	
59	4	M	67	100	88.67	30	28.00	0.43	0.71	5.80	8.93
59	4			100		28		0.25		9.00	
59	4			66		26		1.45		12.00	
60	4	M	67	84	82.67			0.22	0.33		
60	4			84				0.36			

APPENDIX III

TABLE 7

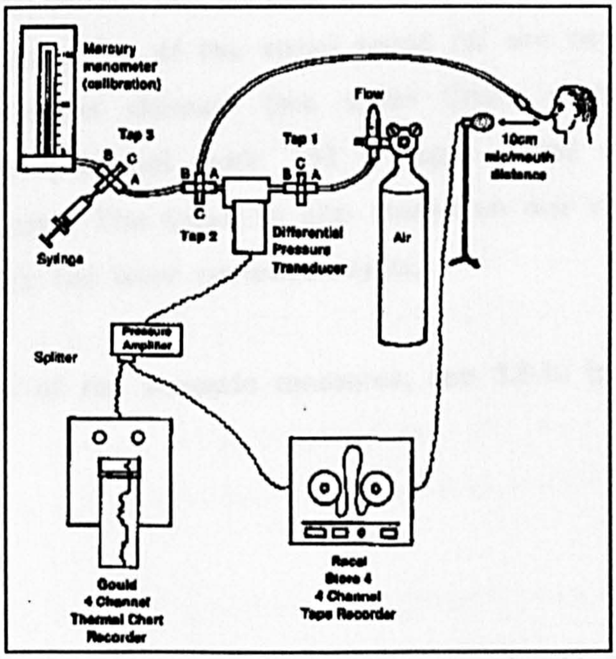
Subj No	Group	Sex M/F	Age	Press.1 mmHg	Mean Press.1	Press.2 mmHg	Mean Press.2	Dur.1 Sec.	Mean Dur.1	Dur.2 Sec.	Mean Dur.2
60	4			80				0.41			
61	4	F	59	80	72.67			2.75	1.87		
61	4			68				1.02			
61	4			70				1.85			
62	4	M	45	100	92.00			0.50	0.95		
62	4			100				1.28			
62	4			76				1.08			
63	4	M	66	40				0.51			
63	4			38	47.33			0.45	0.85		
63	4			40				0.74			
64	4	M	70	64	64.00			1.35	1.67		
64	4			62				2.00			
64	4			66				1.66			
65	4	M	60	60	53.33						
65	4			40							
65	4			60							
66	4	F	53	94	95.33				0.74		
66	4			98							
66	4			94				0.74			
67	4	M	65	98	95.33						
67	4			98							
67	4			90							
68	4	M	67	100	88.00						
68	4			80							
68	4			84							

APPENDIX IV

Additional Information on Pressure and Flow Apparatus.

See Section 3.6.4. (p 82) and fig 20 (p 90).

Figure 20.



Each tap, 1, 2 and 3, can have any one route closed or all open.

A. For calibration:

Tap 1 path B is closed

Tap 2 path C is closed

Tap 3 path A, B and C is open.

The pressures from the manometer are adjusted with the syringe and recording levels of 100, 80, 60, 40, 20 and 0 mmHg are used to calibrate the pressure transducer.

B. In order to take recordings:

Tap 1, path C is closed.

Tap 2, path B is closed.

Tap 3 does not matter - it can be open or closed.

Air flows from the cylinder into pressure transducer chamber, then through the catheter into the subject. The apparatus used is an S.E. laboratory 4.86 pressure transducer, sometimes used as a blood pressure transducer.

The catheter - a red rubber 14 FG modified bronchographic tube, with a radio-opaque tip - is positioned, via the subject's nose, with the tip under the P-E segment. This is achieved by visualising the tract using Videofluoroscopy.

As the subjects phonate, the opening and sustaining pressures of the P-E segment during production of the vowel sound /a/ are recorded on the Gould 8188-440-09 4 channel thermal (hot wire) Chart recorder, which has a sensitivity set at 2.5v full scale and a paper speed of 10 cm/sec. This provides a hard copy. The signal is also stored on one channel of the Racal 4 channel recorder for later review/analysis.

For a description of the acoustic measures, see 3.6.5. (p.82).

APPENDIX V.

Additional Surgical Details On Subjects.

See also 4.1. (p.100) and Table 7, Appendix III (p.165).

HYPOTONIC GROUP

In Group 1, Hypotonic Subjects, there were 15 subjects recorded. Details are on p.100.

Three of the subjects in this Group had a laryngo-pharyngo-laryngectomy with gastric pull-up. These are Subjects 5, 9, and 10, in Table 7. As might be expected, Subjects 9 and 10 exhibit the lowest Mean P-E opening pressures in this Group. Subject 4, also with a low Mean opening pressure, was however, not a subject who had a gastric pull-up. He had received a standard laryngectomy, but had a remarkably low P-E segment (below level C6 cervical vertebra).

In the Hypotonic subjects, the time since surgery varied between nine months (Subject 7) and eight years (Subject 4).

TONIC GROUP

All subjects in this Group had a standard laryngectomy. Additionally, both Subjects 16 and 32 had a surgical myotomy at the time of laryngectomy. Two subjects (24 and 32) had additional unilateral radical block neck dissections, both on the Left side.

The time since surgery varied in this Group between 1 year (Subject 16) and 19 years (Subject 35).

HYPERTONIC GROUP

All subjects in this Group had standard laryngectomy surgery. None had additional surgery in the form of Radical neck dissection nor surgical myotomy as far as could be ascertained. However, some subjects had developed a post-operative fistula which had delayed healing.

The time since surgery varied between 1 year (Subject 46) and 10 years (Subject 47).

SPASM GROUP

All Subjects in this Group had standard laryngectomy surgery. None had any additional surgery in the form of Radical neck dissection, nor a surgical myotomy.

Some subjects (50, 65) in this Group reported developing a post-operative fistula which delayed healing.

The time since surgery varied between 18 months (Subject 65) and 7 years (Subject 64).

It was extremely difficult to obtain detailed operative notes on Subjects other than those who had received their initial surgery at Charing Cross, as the researcher was dependent on the Intake Data Sheet (Appendix I, p.156). There was often, therefore, a paucity of surgical details available as Subjects' memory was often unreliable and/or they had a poor understanding of the procedures undergone.

Attempts to obtain operative details from the surgeon who undertook the original surgery proved impossible as the standard response, when this was attempted during the Pilot Study, was.."A Standard laryngectomy was performed..." and .." The pharynx was closed in the usual manner." Thus, the decision was taken to rely on the Intake Data Sheet for operative information, in as far as it was available.

REFERENCES

Algaba, J., Camacho, J.J., Zuluetta, A. et al.(1983) Fistuloplastie phonatoire et laryngectomises. Acte Otorhinolaringol Esp 34: 3-8.

Algaba, J., Zuluetta, A., Camacho, J.J. et al.(1984) Fistuloplastic phonatoire et laryngectomie totale. Acte Otorhinolaringol Esp 35: 201-205.

Algaba, J., Vea, J.C., Perello, J. et al.(1985) Rehabilitation chirurgicale de la voix apres une laryngectomie totale. Review of Logopaedics Phoniatrics and Audiology 4: 200-205.

Algaba, J. (1987) Recuperation de la voix pour les patients laryngectomises. Review of Laryngology and Rhinology (Bordeaux) 2: 139-142.

Amatsu, M. (1978) A one-stage surgical technique for postlaryngectomy voice rehabilitation. Laryngoscope 90: 1378-1386.

Amatsu, M. (1980) A new one stage surgical procedure for post laryngectomy speech. Archives of Otorhinolaryngology (New York) 220: 149-152.

Amatsu, M. (1985) Tracheoesophageal speaking fistula. In Myers, E.N. (ed) New Dimensions in Otorhinolaryngology, Head and Neck Surgery. Vol I. Elsevier Science Publication BV, Amsterdam.

Amatsu, M., Makino, K., Kinishi, M. et al (1986) Primary tracheoesophageal shunt operation for postlaryngectomy speech with sphincter mechanism. Annals of Otology, Rhinology and Laryngology 95: 373-376.

Applebaum, E.L. and Levine, H.L. (1977) Pharyngeal reconstruction after laryngectomy. Laryngoscope, 87: 1884-1890.

Arslan, M. (1972) Reconstructive Laryngectomy. Report on the first 25 cases. Annals of Otology, Rhinology and Laryngology 81: 1349-1360.

Asai, R. (1966) Asai's new voice production method substitution for human speech. In Ono Y (ed) Proceedings of VIII International Congress Otorhinolaryngology, Tokyo 1965. Pub Exerpta Medica, Amsterdam.

Baken, R.J. (1987) Clinical Management of Speech and Voice. Taylor and Francis Ltd., London.

Baugh, R.F., Lewin, J.S., and Baker, S.R. (1987) Preoperative assessment of tracheosophageal speech. Laryngoscope, 97: 461-466.

Baugh, R.F., Baker, S.R. and Lewin, J.S. (1988) Surgical treatment of pharyngoesophageal spasm. Laryngoscope, 98: 1124-1127.

Bentzen, N., Guld, A. and Rasmussen, H. (1976) X-ray videotaped studies of laryngectomised patients. Journal of Laryngology and Otology 90: 655-666.

Berg, Jw. van den, Moolenaar-Bijl, A.J. and Damste, P.H. (1958). Oesophageal Speech. Folia Phoniatrica 10: 65-84.

Berg, Jw van den (1958) Myoelastic-aerodynamic theory of voice production. Journal of Speech and Hearing Research 1: 227-244.

Berg, Jw.van den, Moolenaar-Biji, A.J. (1959) Cricopharyngeal sphincter, pitch, intensity and fluency in oesophageal speech. Pract Oto-rhino-laryngology 21: 298-315.

Berkowitz, J.F., and Lucente, F.E. (1985) Counselling before laryngectomy. Laryngoscope 95: 1332-1336.

Berlin, C.I. (1963a) Clinical measurement of esophageal speech: I.Methodology and curves of skill acquisition. Journal of Speech and Hearing Disorders, 28: 42-51.

Berlin, C.I. (1963b) Clinical measurement of esophageal speech: II. An unexpected dividend. Journal of Speech and Hearing Disorders 28: 389-392.

Berlin, C.I. (1965) Clinical measurement of esophageal speech: III. Performance of non-biased groups. Journal of Speech and Hearing Disorders 30: 174-183.

Berry, R.J. (1983) Chemotherapy and Radiotherapy. In Edels, Y. (ed) Laryngectomy. Diagnosis to Rehabilitation. Croom Helm. London-Canberra.

Bleach, N.R., Perry, A., Cheesman, A.D. (1989). Surgical voice restoration with the Blom-Singer prosthesis following pharyngo-laryngo oesophagectomy and pharyngogastric anastomosis: a report of four cases. Annals of Otology, Rhinology and Laryngology (In press).

Blom, E.D. (1979) Radiographic and manometric assessment of the patient who fails to acquire esophageal voice. In Keith, R.L. and Darley, F.L. (eds) Laryngectomee Rehabilitation. College Hill Press, Houston

REFERENCES

Blom, E.D. and Singer, M.I. (1979) Surgical-prosthetic approaches to post-laryngectomy speech rehabilitation. In Keith, R.L. and Darley, F.L. (eds) Laryngectomee Rehabilitation. College Hill Press, Houston.

Blom, E.D. (1985) An improved esophageal insufflation test. Archives of Otolaryngology 111: 211-212.

Bockler, R., Wein, B., Klajman S. et al (1988) Die ultra schalluntesuchung der pseudoglottis bei kehlkopflosen. HNO 36: 115-118.

Bozymski, E.M. and Pharr, S.Y. (1972) Esophageal manometry and speech proficiency in postlaryngectomy patients, abstracted. Gastroenterology 62: 627.

Brandenberg, J.H., Bless, D.M. and Salinsky, B. (1985) Speech results of Brandenberg's neoglottic procedure. In Myers, E.N. (ed) New Dimensions in Otorhinolaryngology, Head and Neck Surgery. Vol II. Pub. Elsevier science BV, Amsterdam.

Briani, A.A. (1952) Riabilitazione fonetica di laringectomizzati a mezzo della corrente. English abstract. Archives of Italian Otology 63: 469-475.

Bronwer, B., Snow, G.B. and Dam F.A.S.M. van (1979) Experiences of patients who undergo laryngectomy. Clinical Otolaryngology 4 : 109-118.

Cheesman, A.D. (1983) Surgical management of the patient. In Edels, Y. (ed) Laryngectomy. Diagnosis to Rehabilitation. Croom Helm, London Canberra.

Cheesman, A.D., Knight, G., McIvor, J. and Perry, A.R. (1986) Assessment procedures for post-laryngectomy patients who desire surgical voice restoration. Journal of Laryngology and Otology 2: 191-199.

Christensen, J.M. and Weinberg, B. (1976) Vowel characteristics of esophageal speech. Journal of Speech and Hearing Research 19: 678-689.

Clerf, L.H. and Putney, F.G. (1942) Cricopharyngeal spasm. Laryngoscope 52: 944-953.

Conley, J.J. (1959) Vocal rehabilitation by autogenous vein graft. Annals of Otology, Rhinology and Laryngology 68: 990-995

Cooley, J.W. and Turkey, J.W.(1965) An algorithm for the machine calculation of the complex Fourier series. Mathematics of Computation, 19: 297-301.

Curry, E.T. and Snidecor, J.C. (1961) Physical measurement and speech perception in esophageal speech. Laryngoscope 71: 415-424.

Damste, P.H. (1956) Why are some patients unable to learn esophageal speech? Annals of Otolology, Rhinology and Laryngology 65: 998-1005.

Damste, P.H. (1958) Oesophageal speech after laryngectomy. Ph.D thesis, Groningen State University. Published Gebr Hoitsema.

Damste, P.H. (1979) Rehabilitation score list for the laryngectomised. In Keith, R.L. and Darley, F.L. (eds) Laryngectomee Rehabilitation. College Hill Press, Houston.

Darley, F.L., Aronson, A. and Brown, J.R. (1975) Motor Speech Disorders. W.B. Saunders & Co.

Darvill, G. (1983) Rehabilitation, not just voice. In Edels, Y. (ed) Laryngectomy. Diagnosis to Rehabilitation. Croom Helm London-Canberra.

Davis, S.B. (1979) Acoustic characteristics of normal and pathological voices. In Lass, N.J. (ed) Speech and Language: Basic Advances in Research and Practice Vol I. New York Academic Press.

Davis, P.J. (1981) Rehabilitation after total laryngectomy. Medical Journal of Australia 1: 396-400.

Decroix, G., Libersa, C.L. and Lattard, R. (1958) Bases anatomiques et physiologiques de la reeducation vocale des laryngectomises. French Journal of Oto-Rhinolaryngology 7: 549-573.

Dhillon, R.S., Palmer, B.V., Pittam, M.R. et al (1982). Rehabilitation after major head and neck surgery - a patient's view. Clinical Otolaryngology 7: 319-324.

Diedrich, W.M. (1968) The mechanism of esophageal speech. Annals of the New York Academy of Sciences 155: 303-317.

Diedrich, W.M. and Youngstrom, K.A. (1966) Alaryngeal Speech. CC Thomas publication, Springfield Illinois.

Duguay, M. (1979) Special problems of the alaryngeal speaker. In Keith, R.L. and Darley, F.L.(eds) Laryngectomee Rehabilitation. College Hill Press, Houston.

REFERENCES

Edels, Y. (1983) Pseudo voice - its theory and practice. In Edels, Y.(ed) Laryngectomy. Diagnosis to Rehabilitation. Croom - Helm, London-Canberra.

Edwards, N. (1983) The surgical approach to speech rehabilitation. Chapter in Edels, Y. (ed) Laryngectomy. Diagnosis to Rehabilitation. Croom Helm, London Canberra.

Evans, B.S. and Drummond, S.S. (1985) Surgical voice restoration procedures for laryngectomees. A review. Folia Phoniatrica 37: 163-194.

Gardener, W.H. (1971) Laryngectomee Speech and Rehabilitation. Charles C. Thomas, Springfield, III.

Gardener, W.H. and Harris, H.E. (1961) Aids and devices. Archives of Otolaryngology. 73: 145-152.

Gates, G.A., Ryan, W., Cantu, E. et al (1982) Current status of laryngectomee rehabilitation II: Causes of failure. American Journal of Otolaryngology 3: 8-14.

Ghosh, P. (1977) Transverse tracheo-esophagoplasty. A new one-stage operation for construction of a "neo-larynx." Journal of Laryngology and Otology 91: 1077-1083.

Ghosh, P. (1985) Neolarynx (a new operation). In Myers, E.N. (ed) New Dimensions in Otorhinolaryngology, Head and Neck Surgery Vol II. Elsevier science publ. BV, Amsterdam.

Gilchrist, A.G. (1973) Rehabilitation after laryngectomy. Acta Otolaryngology 75: 511-518.

Graham, J. (1983) The course of the patient from presentation to diagnosis. In Edels, Y. (ed) Laryngectomy. Diagnosis to Rehabilitation pub. Croom Helm, London Canberra.

Grunwell, P. (1983) Phonological therapy: premises, principles and procedures. Proceedings of XIX Congress of International Association of Logopaedicas and Phoniatrics. 14-18 August. University of Edinburgh.

Hall, J.G., Dahl, T. and Arnesen, A.R. (1985) Speech prostheses: failures, problems and success. In Myers, E. N. (ed) New Dimensions in Otorhinolaryngology, Head and Neck Surgery. Vol I. Elsevier science publ. BV, Amsterdam.

Hirose, H. (1983) A fibrescopic study of articulatory adjustments in esophageal speech. Proceedings of XIX Congress of International Association of Logopaedics and Phoniatrics. August 14-18, University of Edinburgh.

Hoops, H.R. and Noll, J.D. (1969) Relationship of acoustic variables to judgements of esophageal speech. Journal of Communication Disorders 2: 1-13.

Inglefinger, F.J. (1958) Esophageal motility . Physiology Review 38: 533-584.

Isshiki, N. (1964) Regulatory mechanism of voice intensity variation. Journal of Speech and Hearing Research 7: 17-29.

Isshiki, N. (1969) Airflow in esophageal speech. In Snidecor, J.C. (ed) Speech Rehabilitation of the Laryngectomised. CC Thomas publication. Springfield Ill.

Jakobi, H. and Muller, O. (1958) Die korperliche belastungsfahigkeit glottisschlussloser patienten. Archives Ohren Nasen Kehlkopfheilkd 173: 312-316.

Johns, M.E., and Cantrell, R.W. (1981) Voice restoration of the total laryngectomy patient: The Blom-Singer technique. Otolaryngology, Head and Neck Surgery 89: 82-86.

Kawasaki, H., Mihashi, K., Aso, M. et al (1983) Conservation, restoration, and rehabilitation of voice in treating patients with a carcinoma of the larynx. A 10 year review of 244 patients. Auris Nasus Larynx Suppliment 10: 111-114.

King, P.S., Fowlks, E.W. and Pierson, G.A. (1968) Rehabilitation and adaptation of laryngectomy patients. American Journal of Physical Medicine 47: 192-203.

Kytta, J. (1964) Finnish oesophageal speech after laryngectomy. Sound spectrographic and cineradiographic studies. Acta Otolaryngology Supplement 195: 1-94.

Lauder, E. (1969) A laryngectomee's point of view on the intelligibility of esophageal speech. Journal of Speech and Hearing Disorders 34: 355-360.

Lecluse, F.L., Tiwari, R.M. and Snow, G.B. (1981) Electroglottographic studies of the Staffieri neoglottis. Laryngoscope 91: 971-975.

Lerman, J. (1977) Rehabilitation of the laryngectomee: whose role? Presented at the Laryngectomee Rehabilitation Seminar, Mayo Clinic.

Levin, N.M. (1962) Voice and Speech Disorders. Medical Aspects. CC Thomas publ. Springfield, Ill.

REFERENCES

Locke, B. (1966) Psychology of the laryngectomee (IALC survey). Military Medicine, July.

Logemann, J.A., Fisher, H.B. and Becker, N.G. (1980) Northwestern otolaryngology communication profile. In Shedd, D. P. and Weinberg, B. (eds) Surgical and Prosthetic Approaches to Speech Rehabilitation. G.K. Hall & Co Medical Publishers, Boston.

Logemann, J.A. (1983) Vocal Rehabilitation after Extensive surgery for post-cricoid carcinoma. In Edels, Y. (ed) Laryngectomy. Diagnosis to Rehabilitation. Croom Helm, London and Canberra.

Mahieu, H.F. (1988) Voice and Speech Rehabilitation. Ph.D. thesis. University of Groningen, Kemper, Holland. October.

Mahieu, H.F., Annyas, A.A., Schutte, H.K. and Van der Jagt, E.J. (1987) Pharyngeal myotomy for vocal rehabilitation of laryngectomees. Laryngoscope 97: 451-457.

Martin, D. (1979) Pre and Post - operative anatomical/ physiological observations in laryngectomy. In Keith, R and Darley F.L. (eds) Laryngectomee Rehabilitation. College Hill Press.

Metfessel, M. (1927) Techniques for the objective study of vocal articulation. Psychology Monograph 36: 1-40.

McConnel, F.M.,and Whitmore, R.(1980) Neoglottic reconstructive laryngectomy: a preliminary report. In: Shedd, D.P. and Weinberg, B. (eds). Surgical and Prosthetic Approaches to Speech Rehabilitation. G.R. Hall & Co. Medical publishers.

McGarvey, S.D. and Weinberg, B.(1984) Esophageal insufflation testing in non-laryngectomised subjects. Journal of Speech and Hearing Disorders 49: 272-277.

Micheli-Pellegrini, V. (1957) On the so-called pseudoglottis in laryngectomised persons. Journal of Laryngology, Otology and Rhinology 71: 405-410

Milford, C.A., Perry, A., Mugliston, T.A. and Cheesman, A.D. (1988) A British experience of surgical voice restoration as a primary procedure. Archives of Otolaryngology, Head and Neck Surgery 114 (12): 1419-1421.

Miller, A.H. (1967) First experiences with the Asai technique for vocal rehabilitation after laryngectomy. American Journal of Otology, Rhinology and Laryngology 76: 829-833.

Miller, A.H. (1968) Further experiences with the Asai technique for vocal rehabilitation after laryngectomy. Trans American Academy of Opthalmology and Otolaryngology 72: 779-781.

Miller, A.H. (1971) Four years experience with the Asai technique for vocal rehabilitation for the laryngectomised patient. Journal of Laryngology and Otology 85: 567-576.

Miller, A.H. (1974) Asai technique for vocal rehabilitation. American Laryngology Association 95: 71-74.

Miller, A.H. (1975) Experiences with the Asai technique. Canadian Journal of Laryngology 4: 549-551.

Motta, G.A., Cesari, U., Iengo, M. et al (1985) Comparative study by means of electroglottography and spectrography in voice patients operated by cordectomy or total laryngectomy. In Myers, E.N. (ed) New Dimensions in Otorhinolaryngology. Head and Neck Surgery Vol II. Elsevier Science publ. Amsterdam.

Moolenaar-Bijl, A.J. (1952) Some data on speech without a larynx. Folia Phoniatrica 3 : 20-24.

Mueller, P.B. and Kupperman, G.L. (1973) Post laryngectomee speech: an evaluation of a Japanese pneumatic speech aid. Eye, Ear, Nose and Throat Monthly 51: 48-52.

Natvig, K. (1983) Laryngectomees in Norway. Study no. 3. Pre-and post-operative factors of significance to esophageal speech acquisition. Journal of Otolaryngology 12: 249-254.

Negus, V.E. (1938) Affections of the cricopharyngeal fold. Laryngoscope 48: 847-858.

Nichols, A.C. (1969) Loudness and quality in esophageal speech. In Snidecor, J.C. (ed) Speech Rehabilitation of the Laryngectomised. CC Thomas publ. Springfield, Ill.

Novak, A., Hlava, A., Hybasek, I. et al (1982) Phonation with the neoglottis phonation according to Staffieri. In Novak, A. and Sram, F. (eds) Congressus Decimus Societatis Phoniatricae Europaeae, Abstract UEP, Prague. Nov 10-12.

Panje, W.R. (1981) Prosthetic vocal rehabilitation following laryngectomy; the voice button. Annals Otology, Rhinology, Laryngology 90: 116-120.

Perry, A. and Edels, Y. (1985) Recent advances in the assessment of failed oesophageal speakers. British Journal of Disorders of Communication 20: 229-236

REFERENCES

Perry, A., Cheesman, A.D., McIvor, J. and Chalton R. (1987) A British experience of surgical voice restoration techniques as a secondary procedure following total laryngectomy. Journal of Laryngology and Otology. 101. 2: 155-163.

Perry, P.S. and Tikofsky, R.S. (1965) The occurrence of low valued, weak intensity frequencies in normal and esophageal phonation. Appendix VI In Tikofsky, R.S.(ed) Phonetic Characteristics of Esophageal Speech. ORA Project 05539. Office of research administration, University of Michigan.

Pfrang, H. (1986) Social and psycho-social aspects of vocal rehabilitation in laryngectomised patients - preliminary results. In Herrmann, I.F. (ed) Speech Restoration Via a Voice Prosthesis. Springer Verlag, Berlin-Heidelberg.

Pfrang, H., de Maddelena, H., Schohe, R. et al (1986) Intelligibility and psycho-social adjustment. A comparison between patients with oesophageal voice and patients with a voice prosthesis. In Herrmann, I.F. (ed) Speech Restoration Via a Voice Prosthesis. Springer Verlag, Berlin-Heidelberg.

Pinzola, R.H. and Cain, B.C. (1988) Acceptibility ratings of T-E speech. Laryngoscope 98: 394-398.

Pitkin, N. (1953) Factors affecting psychological adjustment in the laryngectomized patient. Archives of Otolaryngology 58: 38-49.

Priest, A. (1984) Speech after laryngectomy. Journal of the Royal Society of Medicine. Dec: 1020.

Pronovast, W. (1942) An experimental study of methods of determining natural and habitual pitch. Sp Monog 9: 111-123.

Putney, F.J. (1958) Rehabilitation of the postlaryngectomized patient. Annals of Otology, Rhinology and Laryngology 67: 544-549.

Robe, E.Y., Moore, P. Andrews, A.H. and Hollinger, P. H. (1956) A study of the role of certain factors in the development of speech after laryngectomy. I: Type of operation. Laryngoscope 66: 173-186.

Salmon, S. (1965) Pressure Variation in the Esophagus, P-E Constriction and Pharynx Associated with Esophageal Speech Production. Unpublished Ph.D thesis. University of Kansas.

Salmon, S. and Goldstein, L.P. (1978) The Artificial Larynx Handbook. Grune and Stratton, New York.

Salmon, S. (1979) Methods of air intake for esophageal speech and their associate problems. In Keith, R.L. and Darley, F.L. (eds) Laryngectomee Rehabilitation. College Hill Press.

Salmon, S. (1983) Artificial larynx speech: a viable means of alaryngeal communication. In Edels, Y. (ed) Laryngectomy: Diagnosis to Rehabilitation. Croom Helm Ltd. London-Canberra.

Sanchez-Salazar, V. and Stark, A. (1972) The use of crisis intervention in the rehabilitation of laryngectomees. Journal of Speech and Hearing Disorders 37: 323-328.

Sandberg, N. (1970) Motility of the pharynx and oesophagus after laryngectomy. Acta Otolaryngology 263: 124-127.

Schaefer, S.D. and Johns, D.F. (1982) Attaining functional esophageal speech. Archives of Otolaryngology 108: 647-649.

Schall, L.A. (1938) Psychology of laryngectomised subjects. Archives of Otolaryngology 28: 581-584.

Schultz-Coulon, H.J. (1984) Rehabilitationskonzept fur kehlkopflose - herorientierung fur die praxis HNO 32: 3-12.

Schutte, H.K. (1980) The Efficiency of Voice Production. Ph.D thesis, University of Groningen, Kempe, Groningen.

Seeman, M. (1922) Speech and voice without larynx. Cas Lek Cesk 61: 369-372.

Seeman, M. (1961) Nouveaux resultants des recherches experimentales dans la reeducation de la voix oesophagienne. Journal of French Otorhinolaryngology 10: 47-56.

Seeman, M. (1967) Rehabilitation of laryngectomised subjects. Acta Otolaryngology 64: 235-241.

Serafini, I. (1969) Restoration of laryngeal function after laryngectomy. Experimental research in animals. Advances in Otorhinolaryngology 16: 95-122.

Serafini, I. (1980) Reconstructive laryngectomy. In Shedd, D.P. and Weinberg, B. (eds) Surgical and Prosthetic Approaches to Speech Rehabilitation. G.K. Hall & Co. Medical Publishers, Boston. 67-76.

Shames, G.H., Font, J. and Mathews, J. (1963) Factors relating to speech proficiency of the laryngectomised. Journal of Speech and Hearing Disorders 28: 273-287.

REFERENCES

Shipp, T. (1967) Frequency, duration and preceptual measures in relation to judgements of alaryngeal speech acceptability. Journal of Speech and Hearing Research 10: 417-427.

Shipp, T. (1970) EMG of pharyngoesophageal musculature during alaryngeal voice production. Journal of Speech and Hearing Research 13: 184-192.

Simpson, I.C., Smith J.C.S. and Gordon, M.T. (1972) Laryngectomy: the influence of muscle reconstruction on the mechanism of oesophageal voice production. Journal of Laryngology and Otology 86: 961-990.

Singer, M.I. and Blom, E.D. (1980) An endoscopic technique for restoration of voice after laryngectomy. Annals of Otology Rhinology and Laryngology 89: 529-533.

Singer, M.I. and Blom, E.D. (1981). Selective myotomy for voice restoration after total laryngectomy. Archives of Otolaryngology 107: 670-673.

Singer, M.I. and Blom, E.D. (1985) Esophageal insufflation testing in non-laryngectomised adults: a response to McGarvey and Weinberg. Journal of Speech and Hearing Disorders 50: 287.

Singer, M.I., Blom, E.D. and Hamaker, R.C. (1986) Pharyngeal plexus neurectomy for alaryngeal speech rehabilitation. Laryngoscope 96: 50-53.

Singer, M.I. (1988) The upper esophageal sphincter: role in alaryngeal speech acquisition. Suppliment II Head and Neck Surgery 10: 118-123.

Sisson, G.A., Bytell, D.E., Becker, S.P. et al (1978) Total laryngectomy and reconstruction of a pseudoglottis: problems and complications. Laryngoscope 88: 639-650.

Sisson, G.A. and Goldman, M.E. (1980) Pseudoglottis procedure: update and secondary reconstruction techniques. Laryngoscope 90: 1120-1129.

Snidecor, J.C. and Curry, E.T. (1959) Temporal and pitch aspects of superior esophageal speech. Annals of Otology Rhinology and Laryngology 68: 1-14

Snidecor, J.C. and Curry, E.T. (1960) How effectively can the laryngectomee expect to speak? Norms for effective esophageal speech. Laryngoscope 70: 62-67.

Snidecor, J.C. and Isshiki, N. (1965) Air volume and air flow relationships of six male esophageal speakers. Journal of Speech and Hearing Disorders 30: 205-216.

Snidecor, J.C. (1969) The charging and expulsion of esophageal air. In Snidecor, J.C. (ed) Speech Rehabilitation of the Laryngectomised. CC Thomas publ. Springfield, Ill.

Staffieri, M. (1980) New surgical approaches after total laryngectomy. In Shedd, D.P. and Weinberg, B. (ed) Surgical and Prosthetic Approaches to Speech Rehabilitation. G.K. Hall & co. Med Publ., Boston.

Stevens, S.S. and Volkman, J. (1940) The relationship of pitch to frequency. American Journal of Pyschology 53: 239-253.

Tarnowska, C., Jach, K. and Mozolewski, E. (1982) Zur frage des phonationsmechanismus bei sprachfisteln nach der laryngektomie. In Novak, A. and Sram, F. (ed) Congresses Decimus Societatis Phoniatricae Europaeae, Abstract. UEP, Prague.

Tiwari, R.M., Snow, G.B., Lecluse, F.L.E. et al (1982) Surgical rehabilitation of speech after total laryngectomy. (Dutch) Ned Tijdschr Geneeskd 126: 373-378.

Vuyk, H.D. (1985) Surgical Voice Rehabilitation after Total Laryngectomy and Staffieri's Procedure. Ph.D. Thesis, Free University of Amsterdam.

Vrticka, K. and Svoboda, M. (1963) Time changes in the X-ray picture of the hypopharynx, pseudoglottis and esophagus in the course of vocal rehabilitation of 70 laryngectomised speakers. Folia Phoniatrica 15: 1-12.

Watts, R.F. (1975) Total rehabilitation of Laryngectomees. Laryngoscope 85: 671-673.

Weinberg, B. and Bennett, S. (1972) Selected acoustic characteristics of esophageal speech by female laryngectomees. Journal of Speech and Hearing Research 15: 211-216.

Wepman, J.M., MacGahan, J.A., Rickard, J.C. and Shelton, N.W. (1953) The objective measurement of progressive esophageal speech development. Journal of Speech and Hearing Disorders 18: 147-251.

Winans, C.S., Reichbach, E.J. and Waldrop, W.F. (1974) Esophageal determinants of alaryngeal speech. Archives of Otology 99: 10-14.

Wirth, G. (1972) Moglichkeiten der rehabilitation von laryngektomierten. Z. Laryngology Otology 51: 102-110

REFERENCES

Wochnick, M. (1972) Die Leistungsminderung der Kehlkopfexstirpierten und ihre psychische widerrspiegelung in verschiedenen Fragebogentests. Folia Phoniatrica 24: 195-213.

Wolfe, R.D., Olsen, J.E. and Goldenberg, D.B. (1971) Rehabilitation of the laryngectomee: the role of the distal esophageal sphincter. Laryngoscope 81: 1971-1978.

Zaino, C., Jacobson, H.G., Lepow, H. and Oztuck, C. (1967) The pharyngoesophageal sphincter. Radiology 89: 639-645.

Zinner, E.M. and Fleshler, B. (1972) Intraesophageal pressures during phonation in laryngectomised patients. Journal of Laryngology and Otology 86: 129-140.

Zwitman, D.H. (1979) The function of the pharyngo-esophageal segment during speech. In Keith, R.L. and Darley, F.L. (eds) Laryngectomee Rehabilitation. College Hill Press, Houston.

Zwitman, D.H. and Calcaterra, T.C. (1973) Phonation using the tracheo-esophageal shunt after total laryngectomy. Journal of Speech and Hearing Disorders 38: 369-373.