IMPORTANCE OF HYOID REGION IN VOICE QUALITY

by

Volkan Adem Bilgin

M.D., Yeditepe University, 2003

Submitted to the Institute of Biomedical Engineering in
partial fulfillment of the requirements
for the degree of
Master of Science
in
Biomedical Science

Boğaziçi University
June, 2007

IMPORTANCE OF HYOID REGION IN VOICE QUALITY

by

Volkan Adem Bilgin

APPROVED BY:

Assoc. Prof. Dr. (Thesis Supervisor)	Halil Özcan Gülçür
Assoc. Prof. Dr. (Thesis Co-advisor)	Yasemin P. Kahya
Asst. Prof. Dr. (Thesis Co-advisor)	İsmail Koçak

DATE OF APPROVAL: June 15, 2007

ACKNOWLEDGMENTS

I would like to thank my advisor Assoc. Prof. Dr. Halil Özcan Gülçür for his cordial kindness; his patience and certainly support at every step of my study. I also thank him for introducing me to the world's leading biomedical scientists of during Biyomut 2006.

I would like to thank Dr. İsmail Koçak, who is my co-advisor. He supported me throughout my thesis study by opening his office to me and giving me a chance of being together with the patients, examining them and having definite ideas about my topic. He also showed me the way of thinking by deduction.

I would like to thank the professors of Biomedical Engineering who shared their knowledge with me. I would like to thank my experimental subjects who are my family members and my friends for their patience.

I would like to thank my mother, my father and my sister Sezin for their tolerance during the thesis period. Finally I would like to thank my special Sinem Serap, who always was with me and supported me at least with her wishes.

IMPORTANCE OF HYOID REGION IN VOICE QUALITY

ABSTRACT

Speech is a combined form of phonation and articulation in human beings. Both phonation and articulation implies the acoustic formation, which is the result of mechanical and aerodynamic forces. The acoustic system as a tube analogy is shaped with constrictions and regularities, which represents the vowels in the phonetic system. The effects of the anatomic sites such as mouth, tongue, and pharynx have been well studied in acoustic science. However, the deeper structures in the hypopharynx below the hyoid constriction have not been well defined in the quality and the vowel production.

This study aims to identify the effect of hyoid level and below on voice quality. 20 normal subjects, 7 females and 13 males are included in the present study. The mean age of the males is 30.8 and the mean age for females is 32.8. During data acquisitions, the subjects are instructed to produce Turkish phonetic vowels and the resulting sounds are recorded at a sampling rate of 44100 Hz. The recordings are repeated by applying backward pressure on the main corpus of the hyoid bone at its center.

The data of pre and post constriction is analyzed by wide band spectrograms with a bandwidth of 100 Hz to obtain the formants F1-F4. The frequencies of the peak formantic levels which form the main data are compared using paired samples T-test. Although no significant changes were observed within the sex groups and the whole samples, voice changes are perceived by the listeners in all vocal tasks. As a conclusion, the hyoid region does have profound effect on the resonance system and thus on voice quality, but no considerable influence in the acoustic structure of vowels, i.e., articulation.

Keywords: Voice quality, hyoid, Turkish vowels, resonance, articulation

HYOİD BÖLGESİNİN SES KALİTESİNDEKİ ÖNEMİ

ÖZET

Konuşma, insanlarda fonasyon ile artikülâsyonun birleşmiş biçimidir. Fonasyon ve artikülasyon mekanik ve aerodinamik güçlerin sonucu oluşan akustik oluşumu ifade eder. Tüp benzeri akustik sistem fonetik sistemdeki sesli harfleri ifade eden daralmalar ve düzen ile şekillenmiştir. Ağız, dil ve farinks gibi anatomik alanların etkileri akustik bilim çağında iyice araştırılmıştır. Hyoid kemiğin altında kalan hipofarinksteki daha derin yapılar ses kalitesi ve sesli harf üretimi bakımından yeterince araştırılmamışlardır.

Bu çalışma, hyoid bölgesinin ve altındaki alanın etkilerini belirlemeyi amaçlamaktadır. Bu çalışmaya 20 normal, 7 bayan ve 13 bay denek alınmıştır. Erkeklerin ortalama yaşı 30.8 ve bayanlarınki 32.8 dır. Denekler kendilerine verilen doğrultuda Türkçe sesli harfleri okumuş ve bu 44100 Hz lik okuma hızıyla bilgisayara kaydedilmiştir. Daha sonra hyoid bölgesine ana korpsundan geriye doğru bastırılarak aynı kayıtlar alınmış ve çözümlenmiştir.

Sıkıştırmadan önceki ve sonraki veriler alınıp F1 ile F4 arasındaki "formant" frekanslarını bulmak için 100 Hz bant genişliğindeki geniş bant spektrogramda çözümlenmiştir. Ana veriyi oluşturan tepe formant düzeyleri çift örnekler T-testi ile karşılaştırılmıştır. Tüm örneklerde ve gruplarda anlamlı değişiklikler gözlenmemesine rağmen, tüm vokal çalışmalar sırasında dinleyiciler ses farklılıklarını algılamışlardır. Sonuç olarak hyoid bölgesinin rezonans sistemi üzerinde ve dolayısıyla da ses niteliği üzerinde önemli bir etkisi vardır fakat sesli harflerin akustik yapısı ve dolayısıyla artikülasyon üzerinde önemli bir etkisi yoktur.

Anahtar Sözcükler: Ses niteliği, hyoid, Türkçe sesli harfler, rezonans, artikülasyon

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS

Symbol	Name/Definition
d	width of the vocal fold
E_k	the kinetic energy of the vocal fold
E_p	the potential energy
k	a positive constant that represent the stiffness coefficient of the vocal fold,
l_g	length of the vocal fold
m	a positive constant that represent the mass coefficient of the vocal fold,
$P_g(x)$	the glottal air pressure
r	a positive constant that represent the damping coefficient of the vocal fold
τ	the time delay for the mucosal wave in traveling along the glottis
x	displacement of the vocal fold

1. INTRODUCTION

1.1 Background and Motivation

Speech is a complex process that results from interaction between numerous physical structures within the body, social expectations, emotional influences, and our conscious decisions. The larynx (voice box) is essential to normal voice production, but voice production is not limited to the larynx. The total vocal mechanism includes the abdominal and back musculature, rib cage, lungs, and the pharynx (throat), oral cavity and nose. The physiology of voice production is exceedingly complex. Production of voice begins in the cerebral cortex of the brain. Many other brain centers are involved in sending appropriate impulses to the nerves and muscles required for phonation. The brain also receives tactile and auditory feedback information and makes adjustments in order to control the voice sounds produced. Although it is possible to produce voice even without a larynx, for example in patients who have undergone laryngectomy for cancer, the function of each of the aforementioned body parts is important in voice production and in overall voice quality [1].

Speech involves coordinated use of voice, articulation, and language skills. Although many animals are physiologically able to use the voice for communicating a wide range of simple messages to others of their species, only humans are able to produce true speech. Distinguishing features of voice are its pitch, quality, and intensity, or loudness. Optimum pitch, which means the most appropriate pitch for speaking, varies with each individual. Both optimum pitch and range of pitch are fundamentally determined by the length and mass of the vocal cords; within these limits, pitch may be varied by changing the combination of air pressure and tension of the vocal cords. This combination determines the frequency at which the vocal cords vibrate; the greater the frequency of vibration, the higher the pitch.

The second element in speech is articulation. Articulation refers to the speech sounds that are produced to form the words of language. The articulating mechanism consists of lips, tongue, teeth, jaw, and palate. Speech is articulated by interrupting or shaping airstreams through movement of the tongue, lips, lower jaw, and soft palate. The teeth are used to produce some specific speech sounds. The last is the language. Language is an arbitrary system of abstract symbols agreed upon by any group of people to communicate their thoughts and feelings.

Virtually all parts of the body play some role in voice production and may be responsible for voice dysfunction or bad voice quality. Even something as remote as a strained ankle may alter posture, thereby impair abdominal muscle function and result in vocal inefficiency, weakness and hoarseness. Respiratory, neurological, gastrointestinal, psychological, endocrine, or some other medical disorders can in turn cause disorders in the systems that are responsible in voice production. For example, the vocal tract is physically affected by stress, smoking, drinking, and age, etc.; emotions are affected by hormones and by environmental stresses, etc., which in turn have some effect on voice quality. Thus, any disorder in any of these systems will result in poor voice quality or voice disorder.

Treatment of voice disorders depends upon the etiology. Although many voice disorders can be treated by drug therapy, there are many situations that require laryngeal surgery. Until recently, surgical treatments were avoided in favor of non-surgical therapy if possible; and surgical techniques were used only on people who have lost their voices through injury or disease. However, today the demands of the society is greater; there are many people who have already had plastic surgery and want to sound as well as look younger than they are. Surgeons can make the voice sound younger by bringing the vocal cords closer together. This is done by injecting collagen or similar substances to plump the vocal cords up or by making a little incision in the neck and implanting a little piece of graft to bring all of the vocal fold tissues closer together. Thus, voice that is soft and breathy will be given some strength and solidity that will make it sound more believable and younger. Clearly, such techniques are particularly useful for people who use their voice in their careers; it can benefit people, who may be getting toward the end of their singing careers, and it can benefit people like politicians and teachers who need to have a strong voice.

Just as good looks are important for social status, a “good” voice is also very important. Moreover, just as social rules determine norms for which clothes are considered appropriate for a particular gender; social rules determine norms about what speech patterns are considered “feminine” or “masculine”. In many languages and cultures there are gendered norms and stereotypes relating to how the voice sounds: e.g., pitch, loudness, breathiness, the way particular sounds are produced. There is no universal norm for how women and men are expected to speak. Norms vary depending on many factors, including the specific norms of the language being spoken and the norms of a geographic region, personal characteristics (age, ethnicity/culture, class, etc.) of both the speaker and the listener, the relationship between the people who are talking, the environmental context (work, social setting, etc.), the intended purpose of the discussion, and the mode of communication (telephone vs. face-to-face). Also, many social beliefs about differences between men’s and women’s speech are inaccurate stereotypes.

Today the trend is not just to have “voice lift” surgeries, but more complex voice surgeries to improve voice quality, or for changing voice gender (for trans-sexuals). In this trail a pioneering study was accomplished in our Institute [2]. This thesis study is a complimentary study; the objective is discussed in the next section.

1.2 Objectives

The objective of this thesis is to evaluate the characteristics of phonetic structural changes of Turkish vowels under adjustment at the low pharyngeal (hyoid region) airway and its effects on the voice quality. Turkish phonetic structure of vowels has not been deeply analyzed. In a previous thesis work carried out in our Institute, Koçak showed that laryngeal resonance is one of the most important measures of timbre and quality of the speaking voice and that to treat pitch disorders, resonance characteristics of the patient should be taken into consideration. He reasoned that the hyoid region has a key role in modifying the laryngeal outlet resonance, and that the hyoid region can be surgically manipulated to change the quality of voice [2]. In this work we want to see whether this is justifiable.

1.3 Outline of the Thesis

The remaining chapters are organized as follows. Anatomy of the speech organs, mainly the larynx are discussed in Chapter 2. Vocal fold dynamics and acoustic properties of the vocal tract are discussed in Chapter 3. The data acquisition setup used for acquiring the sound signals, the software used in data analysis and for formant computations are presented in Chapter 4. The results are given in Chapter 5. Conclusions and the recommendations for future work are presented in Chapter 6. A brief summary of the Multi-Speech Model 3700 speech analysis software is given in Appendix A.

2. ANATOMY OF THE SPEECH ORGANS

2.1 Anatomy of the Lips

The Lips are two fleshy folds that surround the orifice of the mouth, are formed externally of integument and internally of mucous membrane, between which are found the Orbicularis oris muscle, the labial vessels, some nerves, areolar tissue, fat and small labial glands.

The lips function to provide competence to the oral cavity during mastication and at rest. They affect uttered sounds that facilitate spoken language and provide changes of facial expression that facilitate unspoken language. They provide sensory information about food prior to its placement in the oral cavity. To accomplish the multitude of functions, lips require a complex system of muscles and supporting structures.

The sphincter muscle of the lip is orbicularis oris and the dilator muscles consist of a series of small muscles that radiate out from the lips.

The fibers of orbicularis oris encircle the oral orifice within the substance of the lips. Some of the fibers arise near the midline and the others arise from the deep surface of the skin. Many of the fibers are derived from the buccinators muscle. The muscle is supplied by the buccal and mandibular branches of facial nerve and its main function is to compress the lips together [3].

The dilator muscles radiate out from the lips, and their action is to separate the lips in combination with the separation of the jaws. The muscles are supplied by the buccal and mandibular branches of the facial nerve.

2.1.1 Articulation

The lips function for creating different sounds mainly the labial, bilabial and labiodental consonant sounds so they form an important part of the speech apparatus. They are also enable whistling and the performing of wind instruments like the trumpet, clarinet, and flute.

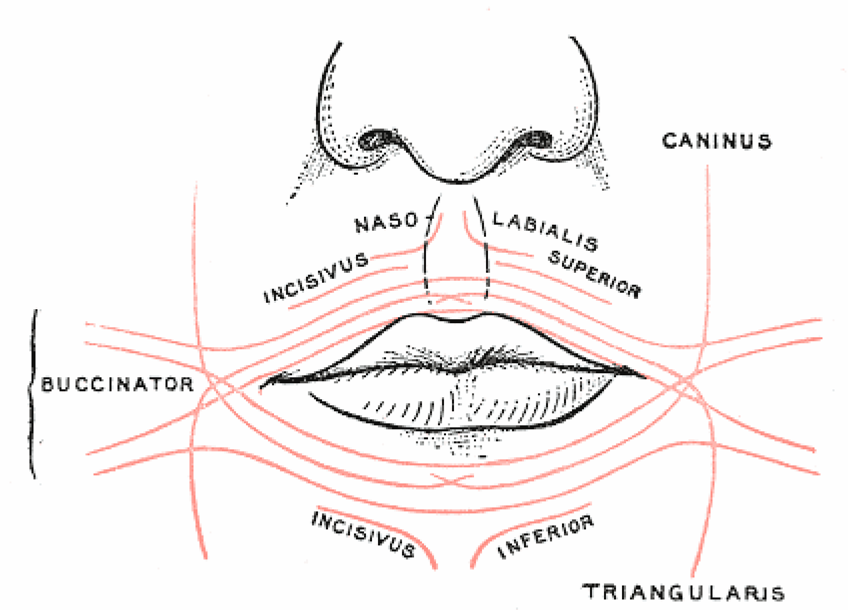


Figure 2.1 Scheme showing arrangement of fibers of orbicularis oris [4].

2.2 Anatomy of the Tongue

The tongue is anchored to the floor of the mouth and slung at the rear from muscles attached to an outgrowth at the base of the skull. The tongue is a strong muscle that is covered by the lingual membrane, which has special areas which detect the flavor of food. The tongue is made up of muscles covered by mucous membranes. These muscles are attached to the lower jaw and to the hyoid bone above the larynx. There are very small nodules, called papillae, from the top surface of the tongue, which give it its rough texture. Between the papillae at the sides and base of the tongue are small, bulblike structures that are sensory organs, called "taste buds,". The muscle fibers are heavily supplied with nerves, so it can manipulate food in the mouth and place it between the teeth for chewing without being bitten in the process. Babies have many more taste buds than an adult, and they have these almost everywhere in the mouth, including the cheeks. Nevertheless, adults enjoy more flavors than babies, who dislike bitter tastes and prefer bland food. The tongue also aids in the formation of sounds of speech and coordinates its movements to aid in swallowing. It is especially helpful when we are forced to "eat" our words [5].

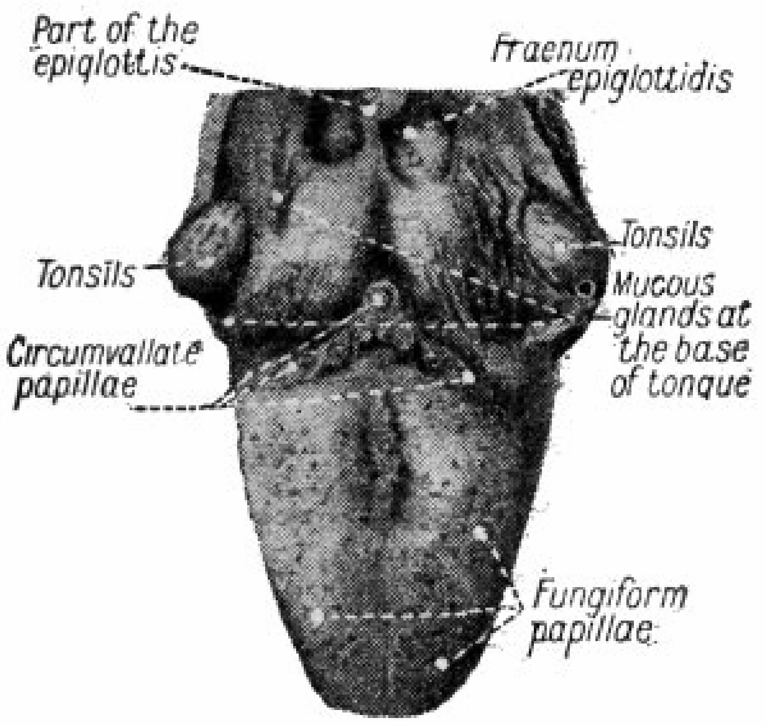


Figure 2.2 Anatomy of the tongue [6].

2.3 Anatomy of the Alveolar ridge

Alveolar ridge, a short distance behind the upper teeth, is a change in the angle of the roof of the mouth. In some people the change is quite abrupt, in other very slight. Sounds which involve the area between the upper teeth and the alveolar ridge are called alveolars.

2.4 Anatomy of the Hard palate

The hard palate is a thin horizontal bony plate of the skull, located in the roof of the mouth that spans the arch formed by the upper teeth. It is formed by the palatine process of the maxilla and horizontal plate of palatine bone. It forms a partition between the nasal passages and the mouth. This partition is continued deeper into the mouth by a fleshy extension called the soft palate.

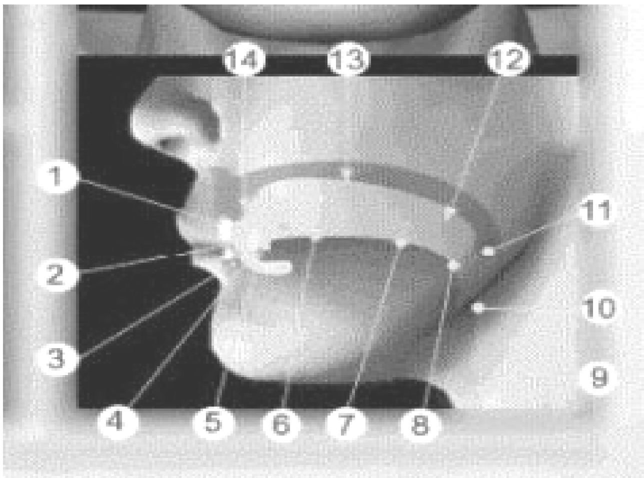


Figure 2.3 Number 3 representing Alveolar Ridge [7].

The interaction between the tongue and the hard palate is essential in the formation of certain speech sounds, notably. Also surrounds the teeth to protect them.

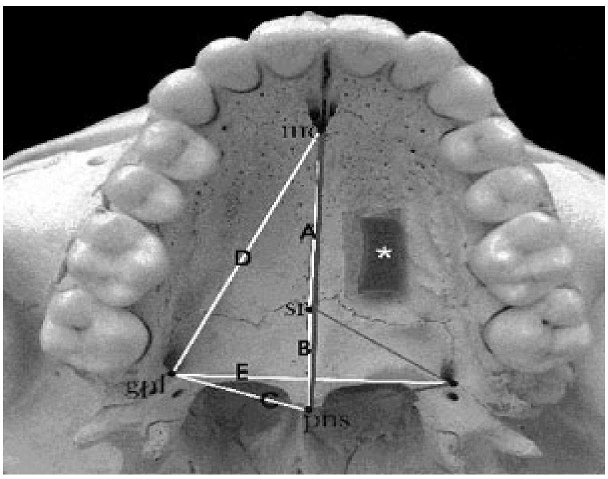


Figure 2.4 Measurement of the Hard Palate [8].

2.5 Anatomy of the Soft palate (velum)

The soft palate is the soft portion of the roof of the mouth, lying behind the hard palate and it performs two important roles in speech : The tongue body hits it in order to make the sounds k, g. It acts as the "gatekeeper" to the nasal cavity. Normally during speech, the velum is in its raised position, blocking off airflow through the nose. But during some sounds (the nasal sounds, like m and n) it lowers and allows air to flow through the nose [9].

Sounds that are made with the velum are called velar sounds. The soft palate's motion during breathing is responsible for the sound of snoring.

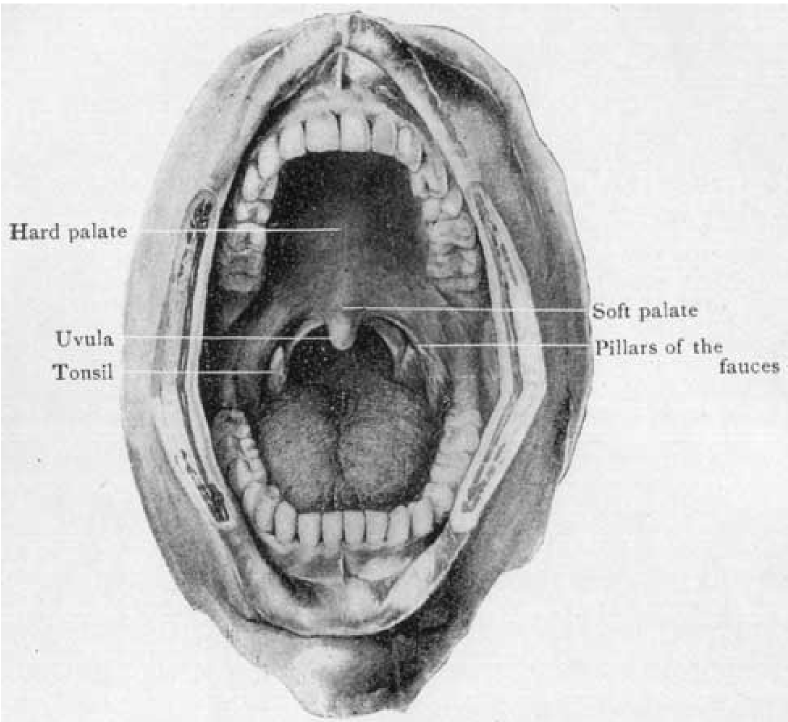


Figure 2.5 Soft Palate and Uvula [10].

2.6 Anatomy of the Uvula

The uvula plays an important role in the articulation of the sound of the human voice to form the sounds of speech. It functions in tandem with the back of the throat, the palate, and air coming up from the lungs to create a number of guttural and other sounds. Consonants pronounced with the uvula are not found in English; however, languages such as Arabic, French, German, Hebrew, Ubykh, and Hmong use uvular consonants to varying degrees. Certain African languages use the uvula to produce click consonants as well. In English (as well as many other languages), it closes to prevent air escaping through the nose when making some sounds [11].

The uvula is the small dangly thing at the end of the soft palate; it is used in many important activities, such as snoring.

2.7 The Larynx

The larynx extends from the root of the tongue to the trachea. It is situated from the level of the third cervical vertebra to approximately the level of the sixth or the seventh cervical vertebra. Anteriorly it is covered by the skin, the fascia, and the infrahyoid muscles. The thyroid gland is attached to and covers the side of the larynx from the cricoid cartilage to the level of the oblique line of the thyroid cartilage.

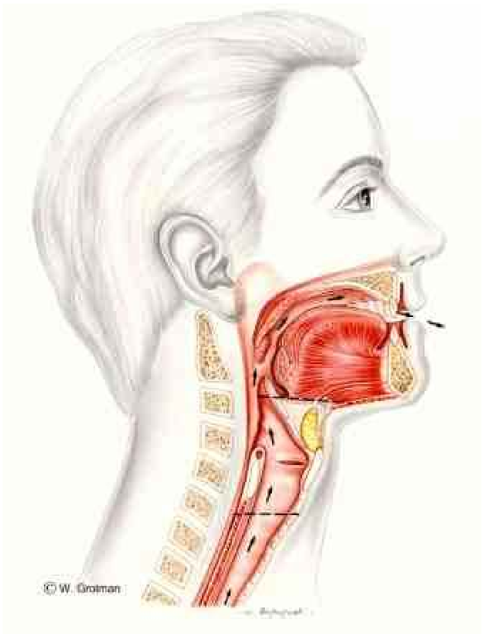


Figure 2.6 The larynx lateral view [13].

Laterally, it gives an attachment to the inferior constrictor muscle, which also encloses the larynx posteriorly. Continuous with the trachea from below, the superior part of the larynx opens into the pharynx above. In the adult, the length of the larynx is 44mm in men and 36mm in women, the transverse diameter is 43mm and 41mm respectively, whereas the anteroposterior diameter is 36mm and 26mm [12].

2.7.1 Bones of the Larynx

Hyoid Bone

The hyoid bone is situated above the thyroid cartilage in the anterior wall of the hypopharynx. It is a sesamoid bone in that it articulates with no other cartilagenous or bony structure but is suspended between the supra and infrahyoid musculature.

Not strictly considered part of the larynx, but as a point of attachment for muscles and ligaments it is essential for laryngeal functions. It is intimately attached to the larynx by the thyrohyoid and the thyroepiglottic ligaments and the extrinsic muscles of the larynx. Since the inferior and posterior borders of the hyoid give attachment to the thyrohyoid and hyoepiglottic membranes, the hyoid forms the anterior and superior boundaries of the preepiglottic space.

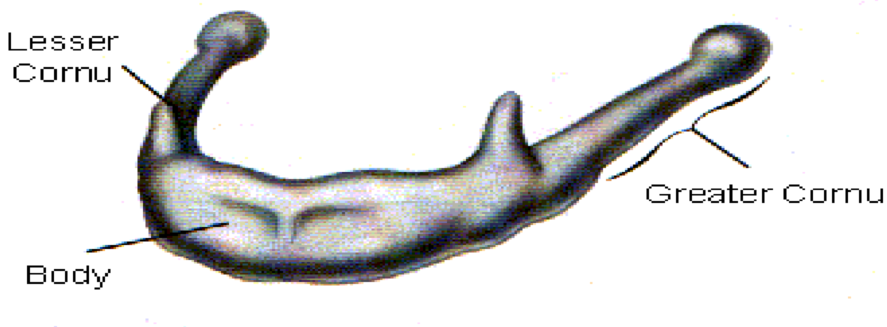


Figure 2.7 Hyoid bone [14].

2.7.2 Cartilages of the Larynx

The cartilages of the larynx are listed below:

- The thyroid
- The cricoid
- The epiglottic
- The arytenoid
- The corniculate
- The cuneiform

The arytenoid, the corniculate, and the cuneiform cartilages form a "group" cartilagenous structure, usually referred to as the arytenoids.

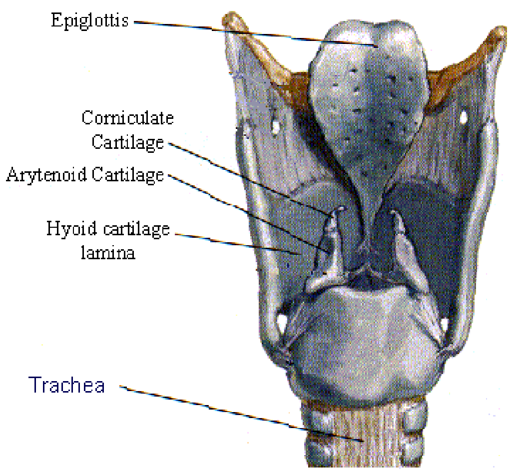


Figure 2.8 Cartilages of the Larynx [15].

Thyroid Cartilage

This is the largest cartilage of the larynx and it is composed of two flat wings, the laminas, and two posterior processes or horns, the superior and inferior thyroid cornua. The lower 2/3's of the flattened laminas are fused anteriorly in the midline forming an angle of 90 degrees in the male (forming the Adam's apple) and 120 degrees in the female. The upper 1/3's of the laminas are not fused and form the thyroid notch. Posteriorly, the laminas do not come in contact giving the thyroid cartilage a V-shape.

The inferior cornu articulates with a facet of the cricoid cartilage, the cricothyroid joint. Anteriorly and laterally, the thyroid cartilage is connected to the hyoid bone by the thyrohyoid membrane and the thyrohyoid ligament (from the superior cornu). The inferior border of the thyroid lamina gives attachment to the circothyroid ligament and CT muscle. The anterior inner surface is used as an attachment to the thyroepiglottic, vestibular, and vocal ligaments, as well as the thyroarytenoid and the vocalis muscles. The oblique line on the lateral side of the laminas serves as an attachment for the inferior constrictor, thyrohyoid and sternothyroid muscles.

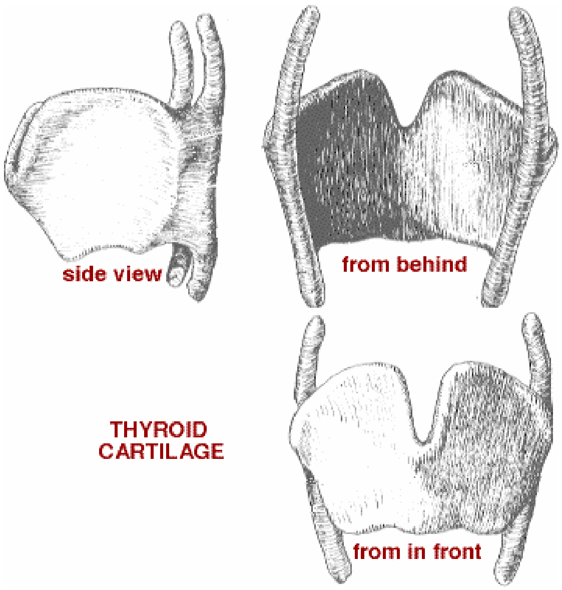


Figure 2.9 Thyroid cartilage [16].

Cricoid Cartilage

This cartilage is shaped like a signet ring, the narrow part of ring faces anteriorly (the cricoid arch) while the broad signet part faces posteriorly (cricoid lamina). It is the only complete cartilagenous ring in the upper airway.

It articulates with the thyroid cartilage through the cricothyroid joints (on the posterior lateral aspect) and with the arytenoids through the cricoarytenoid joints (on the superior lateral aspect). Inferiorly, it is related to the first tracheal ring via ligamentous and musculature attachments only. The cricothyroid ligament and lateral cricoarytenoid muscles are attached anteriorly, along the superior surface of the cricoid arch. The posterior aspect gives an attachment to the posterior cricoarytenoid muscles.

The external and lateral surface of the arch serves as an attachment to the cricothyroid muscle. Laterally, the cricoid cartilage receives the lowermost fibers of the inferior constrictor muscle and the semicircular fibers of the cricopharyngeus muscle.

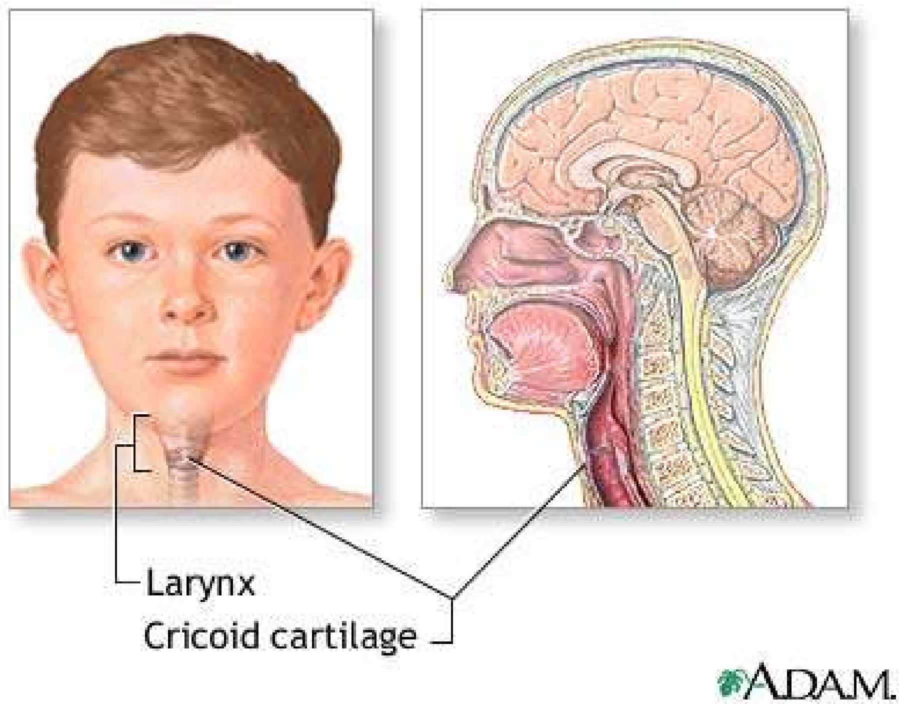


Figure 2.10 Cricoid cartilage [17].

Epiglottic Cartilage

This leaf shaped cartilage is composed mainly of elastic cartilage (unlike the remaining laryngeal cartilages, most of which are hyaline and thus it is very flexible.

At its inferior end, the epiglottis is attached to the inner surface of the thyroid cartilage by the strong thyroepiglottic ligament. Anterosuperiorly, it is attached to the inner surface of the body of the hyoid bone by the midline hyoepiglottic ligament. Its sides are attached to the arytenoid cartilages by the aryepiglottic folds.

The upper part and the anterior surface of the epiglottis are free.

Arytenoid Cartilages

These are two pyramidal-shaped structures which sit on the lateral part of the superior border of the lamina of the cricoid cartilage (the cricoarytenoid joints).

The posterior curved surface receives the fibers of the interarytenoid muscles (transverse and oblique). The posterolateral surface is referred to as the muscular process and it is the point of attachment of the thyroarytenoid and the cricoarytenoid (lateral and posterior) muscles. The anteromedial pointed portion which is called the vocal process, is the attachment to the vocal ligament, the thyroarytenoid and the vocalis muscles. The medial surface of the arytenoids if flat, smooth and it is covered with mucous membrane.

Corniculate Cartilages

Corniculate cartilages are small, paired, nodular structures found in the posterior part of the aryepiglottic folds. They articulate with the summit of the arytenoid cartilages.

Cuneiform Cartilages

These are small, paired, rod like structures embedded in the margin of the aryepiglottic folds. They are found superio-anteriorly to the corniculate cartilages.

2.7.3 Joints

The cricoid cartilage may be regarded as the base and support for the entire larynx. It articulates with the thyroid cartilage via the cricothyroid joints and with the arytenoids via the cricoarytenoid joints. Both of these are true synovial joints.

Cricothyroid Joint

Cricothyroid joint is located between the facet on the posterolateral sides of the cricoid cartilage and the facet on the medial surface of the inferior cornua of the thyroid cartilage. It permits the thyroid cartilage to rotate anteriorly or posteriorly on the cricoid cartilage. When the thyroid cartilage is rotated anteriorly, the anterior arch of the cricoid will come in closer proximity to the lower border of the thyroid cartilage. In so doing, the distance between the cricoarytenoid joint and the inner surface of the thyroid cartilage is lengthened and, assuming that the arytenoids are fixed at this time, stretches and increases the tension on the vocal ligaments. The Cricoid cartilage itself may also rotate to some degree. The cricothyroid joint are called bilateral because the movement of one joint affects opposite joint.

Cricoarytenoid Joint

The Cricoarytenoid joint is the articulation point between the inferior surface of the arytenoid cartilages and the postero-superior surface of the cricoid lamina. This joint permits two types of movements:

1. Rotation of the arytenoid around an axis that runs obliquely from the dorsomediocranial to the ventrolaterocaudal position, permitting movement of the vocal process medially or laterally. The rotation in which the anterior vocal process deviates medially causes adduction of the vocal folds, whereas a lateral rotation causes abduction.

2. Medial and lateral sliding of the arytenoid cartilage toward or away from midline, which brings the arytenoids toward or away from each others. Medial sliding causes adduction of the vocal folds, whereas lateral sliding causes abduction.

The cricoarytenoid joints are bilateral but operate independent of one another.

2.7.4 Ligaments and Membranes

The laryngeal membranes and ligaments can be divided into two subgroups: the extrinsic ligaments and the intrinsic ligaments.

Extrinsic Ligaments

a. *Thyrohyoid Membrane*

This broad fibroelastic sheet is attached from the superior border and the superior horn of the thyroid cartilage to the posterior surface of the body and greater cornua of the hyoid bone. The outer surface of this membrane is covered by the infrahyoid muscles. The middle part of the membrane is thicker and forms the median thyrohyoid ligament, which is pierced on each side by the superior laryngeal artery and vein and the internal laryngeal nerve. The tips of the superior horns of the thyroid cartilage are connected to the greater cornua of the hyoid bone by the lateral thyrohyoid ligaments, which are also thickenings of the thyrohyoid membrane.

b. *Thyroepiglottic Ligament*

The Thyroepiglottic Ligament attaches the anterior lower part of the epiglottis to the thyroid cartilage.

a. *Hyoepiglottic Ligament*

These ligaments attach the anterior upper part of the epiglottis to the hyoid bone.

b. *Cricotracheal Ligament*

The inferior border of the cricoid cartilage is joined to the first ring of the trachea by this ligament.

Intrinsic Ligaments

1. *Quadrangular Membrane*

It extends from the lateral margins of the epiglottis within the aryepiglottic fold and attaches to the arytenoid and corniculate cartilages. The inferior free edge is thickened to form the vestibular ligament (false vocal cord). The superior edge is also free and it is covered with aryepiglottic fold of mucosa.

2. *Conus Elasticus* (also called the Cricothyroid or the Cricovocal Membrane)

This membrane arises from the inner surface of the cricoid arch and it consists of two distinct parts: The anterior or superficial part, also called the medial cricothyroid ligament, extends from the upper border of the cricoid cartilage to the lower border of the thyroid cartilage. The lateral part of this membrane extends from the superior inner border of the cricoid cartilage to the inner surface of the thyroid angle and posteriorly to the tip of the vocal process of the arytenoid cartilage. The superior edge of this part of the Conus Elasticus is free and thickened between its two attachments, the angle of the thyroid lamina and vocal process of the arytenoid cartilage, and forms the vocal ligament (true vocal cord).

2.7.5 Spaces

The internal cavity of the larynx can be divided into four parts. The supraglottic space (also called the vestibule which is surrounded by the piriform fossa), the preepiglottic space, the paraglottic space (which contains the ventricles) and the subglottic space (which is the area below the true vocal folds).

1. *Supraglottic Space* (Vestibule)

It's superior border is free margin of the epiglottis and aryepiglottic folds where the inferior border forms lower margin of the ventricular or false vocal folds. On either side of this cavity there is a small recess called the piriform fossa. This fossa is bound medially by the aryepiglottic fold and laterally by the thyroid cartilage and thyrohyoid membrane. Within the piriform fossa lie branches of the internal laryngeal and the recurrent laryngeal nerves. The piriform fossa, together with the space between the dorsal aspect of the laryngeal cavity and the pharynx, composes the hypopharynx.

2. *Preepiglottic Space*

It's superior border forms the hyoepiglottic ligament, the anterior border forms the thyrohyoid membrane and ligament and the posterior border forms the anterior surface of the epiglottis and thyroepiglottic ligament. This space, which is anterior to the vestibule, is shaped like an inverted pyramid and it contains fat and loose areolar tissue.

3. *Paraglottic Space*

Paraglottic space is superiorly surrounded with the quadrangular membrane, inferiorly conus elasticus, laterally inner surface of the thyroid cartilage and medially ventricle. The ventricle is a laterally directed sac between the undersurface of the ventricular bands and the upper surface of the true vocal folds. The space between the vestibular folds and the true vocal folds is a compartment called the fusiform recess. At the level of the vocal folds and the ventricles, the paraglottic space can be effectively divided by muscular and ligamentous attachments into supraglottic and infraglottic parts.

4. *Subglottic Space*

Superior border of the space is undersurface of the vocal ligament at the midline (the true vocal folds). Inferior border forms lower border of the cricoid cartilage, lateral border is formed by the medial surface of the conus elasticus and medial border is formed by subglottic mucosa.

2.7.6 Muscles

The laryngeal musculature can be divided into two groups:

a. The *intrinsic muscles* include cricoarytenoid, cricothyroid, interarytenoid, thyroarytenoid, and thyroepiglottic muscles. Most of the intrinsic muscles are innervated by the recurrent laryngeal nerve. The cricothyroid is the exception because it is innervated by the external branch of the superior laryngeal nerve.

b. The *extrinsic muscles* can also be divided into two subgroups : the infrahyoid and the suprahyoid. The infrahyoid subgroup contains omohyoid, sternohyoid, sternothyroid, thyrohyoid muscles, whereas the suprahyoid subgroup (which is not strictly laryngeal) contains digastric, geniohyoid, mylohyoid, stylohyoid, stylopharyngeus, and thyrohyoid muscles. The extrinsic muscles, also called the strap muscles, are capable of moving the larynx up or down during deglutition, respiration and phonation. The middle and inferior constrictor muscles are also extrinsic laryngeal muscles and play their most important role in swallowing.

Intrinsic Muscles

Cricoarytenoid Muscles (CA)

There are two cricoarytenoid muscles:

1. *Posterior Cricoarytenoid Muscle*

It attaches to the posterior aspect of the cricoid lamina, posterior surface of the muscular process of the arytenoid cartilages. It is innervated by recurrent laryngeal nerve. It functions as abductor of the vocal cords, serves to open the glottis by a rotary motion imparted to the arytenoid cartilages around an apparent axis of the cricoarytenoid joints.

2. *Lateral Cricoarytenoid Muscle*

It attaches to the upper border of the arch of the cricoid cartilage , anterior aspect of the muscular process of the arytenoids. It is innervated by recurrent laryngeal nerve. It's main function is to close the glottis by adducting the vocal folds, which it does by rotating the arytenoids cartilages medially.

Cricothyroid Muscles (CT)

It attaches to the anterior and lateral part of the external surface of the arch of the cricoid cartilage, anterior fibers (the straight portion or pars recta) run to the lower border of the lamina of the thyroid cartilage, lateral fibers (the oblique portion or pars oblique) insert into the anterior border of the inferior horn of the thyroid. It is innervated by external branch of the superior laryngeal nerve. It functions as the elongation and tension of the vocal fold ligament by elevating the arch of the cricoid cartilage upward toward the lowermost aspect of the thyroid ala. Contraction of the CT also rotates the arytenoids medially, adducting the vocal folds.

Interarytenoid Muscles (Ia)

There are two interarytenoid muscles

1. *Transverse Arytenoid Muscles*

It attaches to the muscular process and lateral border of one arytenoid cartilage, corresponding surfaces of the arytenoid of the opposite side. It is innervated by recurrent laryngeal nerve. It's main function is adduction of the vocal folds by approximating the arytenoid cartilages.

2. *Oblique Arytenoid Muscles*

It attaches to the muscular process of the arytenoid cartilage, apex of the arytenoid on the opposite side. It is innervated by recurrent laryngeal nerve. It functions as a sphincter of the inlet of the larynx during the act of swallowing and also closes laryngeal inlet by approximating arytenoid cartilages. This X-shaped muscle is superficial to the transverse arytenoid muscle. The uppermost fibers continue along the aryepiglottic fold, forming the aryepiglottic muscle

Thyroarytenoid Muscles (Ta)

It attaches to the inner surface of the lower half of the angle of the thyroid cartilage anterolateral surface and vocal process of the arytenoid cartilage. It is innervated by recurrent laryngeal nerve. This muscle can be divided into three parts: Thyroarytenoideus internus or vocalis muscle is the major tensor of the free edge of the vocal fold, thyroarytenoideus external contraction draws the arytenoid cartilages forward toward the thyroid, thus shortening the vocal ligament and thyroepiglotticus widens the inlet of the larynx.

Thyroepiglottic Muscle

It attaches to the internal surface of lamina of thyroid cartilage, lateral margin of epiglottic cartilage. It is innervated by recurrent laryngeal nerve. It widens inlet of the larynx.

2.7.7 Nerves

The nerve supply to the larynx is derived from the motor nuclei in the medulla oblongata in the brainstem. The vagus (cranial nerve X) is the main nerve innervating the larynx and it arises from the nucleus ambiguous (branchiomeric nucleus) and the dorsal motor nucleus of vagus (autonomic, parasympathetic nucleus). The vagus reaches the larynx via the internal and external branches of the superior laryngeal nerve, as well as by the recurrent laryngeal nerve. It is important that, the innervation of the larynx is mainly from the parasympathetic nervous system, however, the sympathetic fibers arising from the superior cervical ganglion also innervate the larynx.

Superior Laryngeal Nerve

The superior laryngeal nerve separates from the main trunk of the vagus just outside the jugular foramen. It passes anteromedially on the thyrohyoid membrane where it is joined by the superior thyroid artery and vein. At approximately this level, the external laryngeal nerve leaves the main trunk. The main internal laryngeal nerve enters the thyrohyoid membrane through a hiatus. It then divides into three set of branches (ascending, transverse and descending), which communicate with the recurrent laryngeal nerve posterior to the cricoid cartilage; this is referred to as the ansa galeni. The external laryngeal nerve supplies the inferior constrictor muscle and the cricothyroid muscle. The main internal laryngeal nerve supplies sensory innervation to the epiglottis, the pyriform sinus, and the larynx as far down as the vocal folds.

Recurrent Laryngeal Nerve

The recurrent laryngeal nerve is derived as a branch of the vagus nerve, on the left side as it passes the arch of the aorta and on the right side as it passes the subclavian artery. Upon reaching the larynx, it becomes the left and the right inferior laryngeal nerve and it passes at the posterior aspect of the cricothyroid joint. These branches communicate with the internal branch of the superior laryngeal nerve posterior to the cricoid cartilage; this is referred to as the ansa galeni. With the exception of the cricothyroid muscle, the inferior laryngeal nerve supplies all the intrinsic muscles of the same side and the transverse arytenoid muscles bilaterally. Moreover, both the left and the right branches supply the trachea and the esophagus while traveling superiorly. Infraglottic portion of mucosa supplies sensory innervation.

2.7.8 Blood Supply and Drainage

The blood reaches the larynx via the arterial blood supply, and it leaves the larynx either by the venous drainage or the lymphatic drainage.

Arterial Blood Supply

The arterial blood supply of the larynx is derived from the laryngeal branches of superior and inferior thyroid arteries and to a small extent from the cricothyroid. Most of the arteries anastomose freely with each other.

1. *Superior Laryngeal Artery*

   It is a branch of the superior thyroid artery, which comes from the external carotid artery. It enters the larynx through the thyrohyoid membrane, along with the internal branch of the superior laryngeal nerve, and supplies the muscles and mucous membranes in the superior portion of the larynx.

2. *Inferior Laryngeal Artery*

   It is a branch of the inferior thyroid artery, which comes from the thyrocervical trunk of the subclavian artery. It ascends on the trachea, together with the recurrent laryngeal nerve, and enters the larynx beneath the lower border of the inferior constrictor muscle. It supplies the muscles and mucous membranes in the lower part of the larynx.

3. *Cricothyroid Artery*

It is a branch of the superior thyroid artery. It passes across the superior portion of the cricothyroid ligament.

Venous Drainage

The venous drainage is supplied by the superior and inferior laryngeal veins, which essentially follow the arteries in their course. The superior drainage joins the superior and middle thyroid veins and then the internal jugular. The inferior drainage joins the middle thyroid vein and the inferior thyroid vein, which empties into the superior vena cava.

Lymphatic Drainage

The larynx is very well supplied with lymphatics, with the exception of the free margins of the vocal folds themselves. The lymphatics of the larynx are divided into a superior and an inferior group. The area of the larynx above the vocal cords is drained into the superior and middle jugular nodes. Drainage is through the cricothyroid membrane to the middle and inferior jugular nodes as well as to the paratracheal lymph nodes.

2.7.9 Histology of the larynx and the vocal folds

Laryngeal histology is very similar to that of the rest of the respiratory tract. It is lined mainly by a pseudostratified, ciliated, columnar epithelium. It also contains a mucosa with laryngeal glands and a few taste buds. The true vocal folds have a specialized histology different from the rest of the larynx. Virtually all the laryngeal membranes and ligaments consist of elastic and collagenous fibers. All the laryngeal muscle is cross-striated muscle. Laryngeal tissues are richly supplied with nerves [18].

Early in life all laryngeal cartilages are hyaline cartilages. Later on, most of the cuneiforms, corniculates, epiglottis and the apices and vocal processes of the arytenoids become converted to elastic cartilage. The mucous membrane of the larynx is continuous with that of the mouth and pharynx and the trachea below. With some exceptions, all of the larynx is lined by respiratory, pseudostratified, columnar, ciliated epithelium with goblet cells. The normal ciliated epithelium of the larynx has an innermost layer of small, round cells, the basal or reserve cell layer. This single cell layer of basal cells is overlaid by a second row of ciliated columnar cells. This ciliated layer may vary considerably in thickness. In the average individual, the upper half of the posterior surface and the anterior surface of the epiglottis, the superior portion of the aryepiglottic folds, and the true vocal folds, the mucous membrane is stratified squamous epithelium. Numerous laryngeal mucous glands, which are small, branching, tubuloalveolar invaginations of the epithelium, are found in the lamina propria of the mucous membrane of the larynx. They are particularly plentiful on the epiglottis, preepiglottic space, posterior wall, and subglottic region. Their secretions, which condition the inspired air and add a protective coating of mucus spread over the epithelial surface, usually flow down over the vocal folds. There are no mucous glands on the free edges of the vocal folds. Taste buds, similar to those found in the tongue, are found on the posterior surface of the epiglottis and in the aryepiglottic folds. Beneath the epithelium this tissue is a rather loose lamina propria, but this becomes denser around solid structures. In this connective tissue all cell types and all fibers are represented, and mast cells are unusually rich.

Vocal Folds

From the histological point of view, the area consists of five layers, which are generally homogenous along the entire vocal fold length. The vocal folds are covered with stratified squamous epithelium which is glabrous. Along with the microvilli, the superficial cell layer of this epithelium forms ridges or plicae. Moreover, there appear to exist contradictory views about the keratinisation of the squamous epithelium on the true vocal folds. The junction between the ciliated, columnar epithelium inferior and superior to the squamous epithelium of the true vocal cords is usually a transitional zone that varies from several cells to a width of 1 to 2 mm.

Immediately beneath the mucosa of the vocal fold is a superficial layer of the lamina propria of the mucosa, which primarily consists of loose fibers and matrix. This superficial gelatin like layer is referred to as Reinke's space because it represents a potential space. It is essential to the free flow of the loose mucosa over the vocal folds during phonation. Reinke's space contains a few capillaries but lacks lymphatics. Due to this limited vascular access, carcinomas confined to the true vocal cords tend to remain localized and are amenable to curative radiation or surgical therapy. The poor lymphatic drainage of Reinke's space also probably contributes to the development of vocal cord nodules and polyps when abnormal amounts of edema like fluid collect in this region.

The intermediate layer of the lamina propria lies underneath the Reinke's space. It consists of elastic fibers which run almost parallel to the edge of the vocal fold.

Underneath this layer is the deep layer of the lamina propria, which is composed chiefly of collagenous fibers. The structure consisting of the intermediate and deep layers of the lamina propria is generally known as the vocal ligament. It is the uppermost portion of the conus elasticus. The vocalis muscle is found beneath the deep lamina propria layer and it constitutes the main body of the vocal fold.

There are gradual changes in stiffness from the very pliable superficial layer of the lamina propria to the rather stiff vocalis muscle. From the mechanical point of view the five layers can be reclassified into three sections: the cover, consisting of the epithelium and the superficial layer of the lamina propria; the transition, consisting of the vocal ligament; and the body, consisting of the vocalis muscle.

3. DYNAMICS OF THE VOCAL FOLD OSCILLATION

3.1 Introduction

The vocal folds, the glottis and vocal tract constitute a self-excited biomechanical oscillator that acts as the sound source during voice production. Under certain conditions such as air pressure, vocal fold tension, and glottal area, the air flow through the glottis causes oscillations. These oscillations produce air pressure waves that we perceive as voice [1, 19].

This oscillator has a relatively complex dynamical structure, as consequence of nonlinear viscoelastic characteristics of its tissues, collisions between the opposite vocal folds, and nonlinear interaction between the airflow and the glottal area. Using mathematical models of that structure, past works have shown the existence of several nonlinear phenomena, such as multiple equilibrium positions and limit cycles [20,21] , several types of bifurcations [20-22], and chaotic behavior and chaotic behavior [21, 23]. The existence of an oscillation hysteresis phenomenon, which is commonly perceived as different laryngeal configurations at voice onset and offset. Voice onset requires a subglottal pressure above certain positive threshold level. However, after voice has started, the subglottal pressure may be decreased below the initial threshold, without the interruption of voice. The pressure level at which voice stops is lower than the level at which it starts. These results are also applied to the analysis of phonation control strategies of men vs. women [24], and the development of motor control in children [25].

3.2 The One-Mass Model

This was the first mathematical model for the vocal fold oscillation dynamics, proposed by Flanagan and Landgraff [26]. As shown in Figure 3.1, each fold is represented by a mass-damper-spring system. Both folds are assumed symmetrical, and motion is allowed only in the horizontal direction. To simplify, the effect of the vocal tract load on the larynx may be neglected , and the subglottal pressure Ps is assumed constant and the supraglottal pressure Po is equal to the atmospheric one. From experimental studies on excised larynges, it is known that the vocal folds oscillate under such conditions, with similar characteristics as those observed during speech [27]. The glottal aerodynamics may be described by Bernoulli's equation, with modifications introduced by the boundary layer model of Pelorson et al. [28] for high Reynolds numbers. The resultant equation of motion is;

$$m\ddot{x} + r\dot{x} + kx = dl_g P_g(x) \tag{3.1}$$

where *m, r, e k* are positive constants that represent the mass, damping coefficient, and stiffness coefficient, respectively, of the vocal fold, x is its displacement, d and l_g are their width and length, respectively, and $P_g(x)$ is the glottal air pressure. This pressure depends on the glottal cross-sectional area and hence is a function of the displacement x. It reaches its maximum when the glottis is fully closed, and decreases as the glottal area increases.

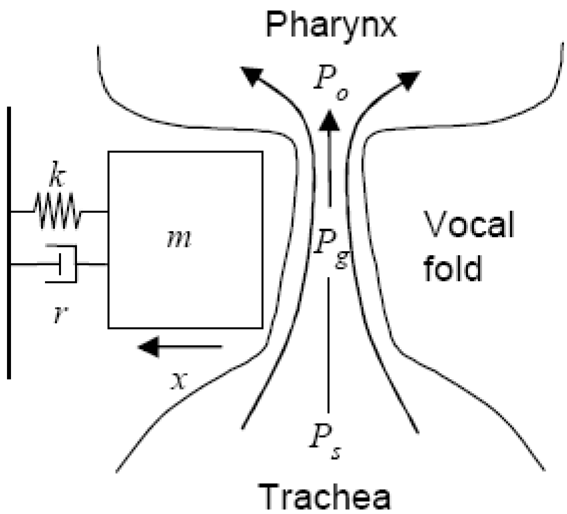


Figure 3.1 The one-mass model [29].

3.2.1 Vocal fold oscillation

It may be easily shown that there is no limit cycle in a system represented by (3.1), which means that it is impossible to obtain a self-excited oscillation with this model. From (3.1) it is obtained that;

$$\frac{d}{dt}(E_k + E_p) = -r\dot{x}^2 < 0 \tag{3.2}$$

where $E_k = m\dot{x}^2/2$ is the kinetic energy of the fold and $E_p = kx^2/2 + V(x)$, with $dV(x)/dx = -dl_g P_g(x)$, is its potential energy. This relation tells that the total energy of the system decreases along trajectories, and so no oscillation of constant amplitude may exist. However, it must be noted that it is possible to obtain a self-excited oscillation when the vocal tract is added to the model [30, 31]. Although this model does not capture the main oscillation mechanism of the vocal folds, nevertheless it may be used as a simple sound source in voice synthesis systems. It may be used also as a model for the falsetto register, where the vocal tract load is significant due to the high fundamental frequency of the oscillation [30].

3.3 The Mucosal Wave Model

This model by Titze [30] improves the previous mass-damper-spring system by adding a surface wave which propagates in the airflow direction. It reproduces observations of the vocal fold oscillation, which show a wavelike motion pattern of the superficial mucosal tissues [27]. Its equation of motion is [20];

$$m\ddot{x} + r\dot{x} + kx = dl_g \frac{2\tau P_s \dot{x}}{x_0 + x + \tau\dot{x}} \tag{3.3}$$

where τ is the time delay for the mucosal wave in traveling along the glottis. The stability of the equilibrium position at x = 0 may be analyzed by taking the linear part of the equation of motion in its neighborhood [32]. Linearizing (2.1), it is obtained that;

$$m\ddot{x} + (r - 2dl_g P_s / x_0)\dot{x} + k x = 0. \quad (3.4)$$

This equation shows that the airflow acts on the vocal folds as an equivalent negative damping. When P_s is zero or very small, the total damping is positive and the equilibrium position is stable. When P_s is large, the total damping becomes negative, which implies a net transfer of energy from the airflow to the vocal folds. In this case, an oscillation of increasing amplitude is produced. The amplitude will be limited eventually, due to collision between the opposite vocal folds and other nonlinear effects which are not included in the model. The critical value of the subglottal pressure is the phonation threshold pressure, given by;

$$Pth = \frac{rx_0}{2dl_g} \quad (3.5)$$

Therefore it is concluded that it is the mucosal wave which allows the vocal folds to absorb energy from the airflow, permitting then their oscillation.

3.4 Oscillation Hysteresis

How an oscillation may be generated from an equilibrium position? In the theory of nonlinear dynamical systems [32], the qualitative change of dynamical behavior at a critical value of a parameter is called a bifurcation. At a Hopf bifurcation, an equilibrium position changes its stability and an oscillation (limit cycle) is generated. Two types of Hopf bifurcations are possible, as illustrated in Figure 3.2. In the figure, a solid line represents stable equilibrium (a position or a limit cycle), and a dashed line represents unstable equilibrium. At the supercritical Hopf bifurcation (top), as the parameter increases a stable equilibrium position bifurcates into an unstable position and a stable limit cycle. This is the simplest case, and corresponds, e.g., to the well-known van der Pol oscillator. In the subcritical Hopf bifurcation (bottom), as the parameter increases a stable equilibrium position and an unstable limit cycle coalesce into an unstable equilibrium position.

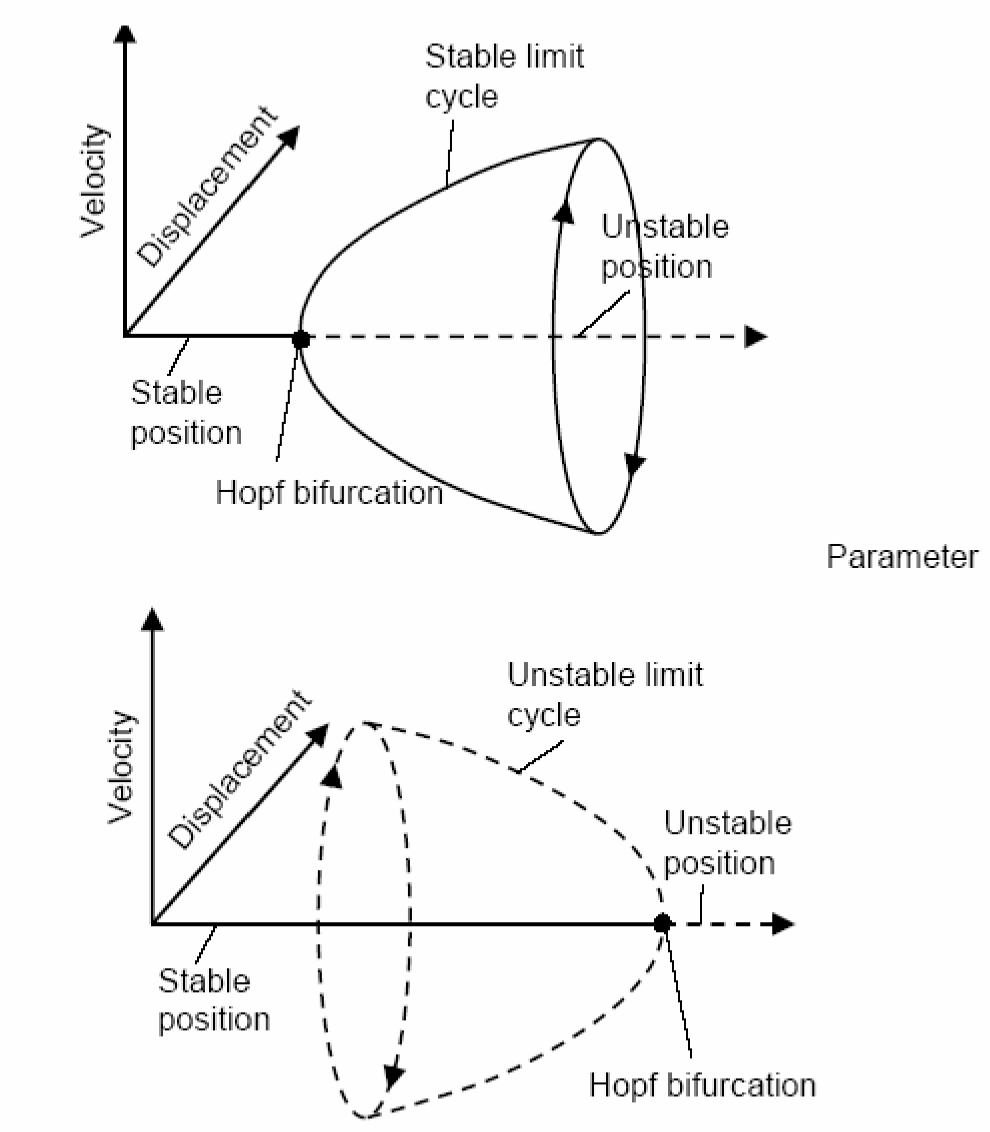


Figure 3.2 The Hopf bifurcation. Top: supercritical, bottom: subcritical [29].

Analyzing (3.1), it can be shown that, at the phonation threshold pressure given by (3.2), two complex eigenvalues cross the imaginary axis transversally from left to right and the equilibrium position becomes unstable. At this pressure value, the equilibrium position is a weak focus and its Lyapunov number (the first nonzero derivative $d^{(k)}(0) \neq 0$, where $d(s) = P(s) - s$ and $P(s)$ is the Poincaré map for the focus) is [33].

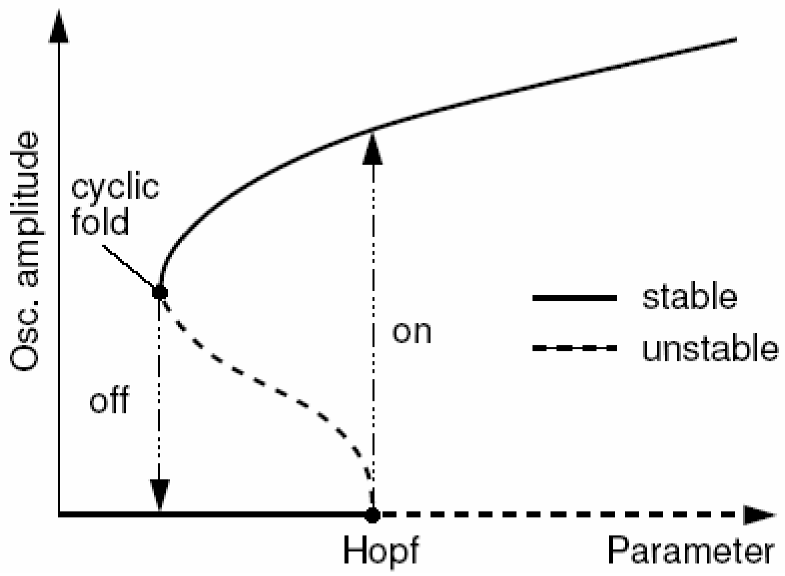


Figure 3.3 Oscillation hysteresis phenomenon. The curves in full and broken lines represent a stable and an unstable limit cycle, respectively. Along the horizontal axis, the full and broken line regions represent the stable and unstable regions of an equilibrium position [29].

$$\sigma = \frac{3\pi r}{2\sqrt{mk}}\left(\frac{r\tau}{m}+\frac{3k\tau^2}{m}+1\right) > 0. \tag{3.6}$$

Since $\sigma > 0$, and according to Hopf Bifurcation Theorem [32], a Hopf bifurcation of the subcritical type occurs. This type of bifurcation often appears in combination with a cyclic fold between limit cycles, as shown in Figure 3.3. At the cyclic fold bifurcation, the unstable limit cycle generated by the Hopf bifurcation, and a stable second limit cycle coalesce and cancel each other.

Suppose that the control parameter is increased from zero. At the Hopf bifurcation, the oscillation will start and will increase its amplitude rapidly to the stable limit cycle. If the parameter is now decreased, the oscillation will follow the curve corresponding to the stable limit cycle, until reaching the cyclic fold bifurcation. At this point, it will vanish abruptly. Thus, oscillation onset and offset occur at different threshold values of the parameter with a hysteresis effect. Note that between the onset and offset thresholds, two stable states co-exist: an equilibrium position and a stable limit cycle. This phenomenon appears commonly in cases of flow-induced oscillation [34]. A large experimental evidence shows hysteresis at the onset-offset of phonation.

For example, studies of excised larynges [27] have shown that the subglottal pressure is lower at oscillation offset than at oscillation onset. Similar results have been found in subjects producing speech [35]. In all cases, the results always show that the conditions to start the vocal fold oscillation are more restricted than those to maintain it, as described above.

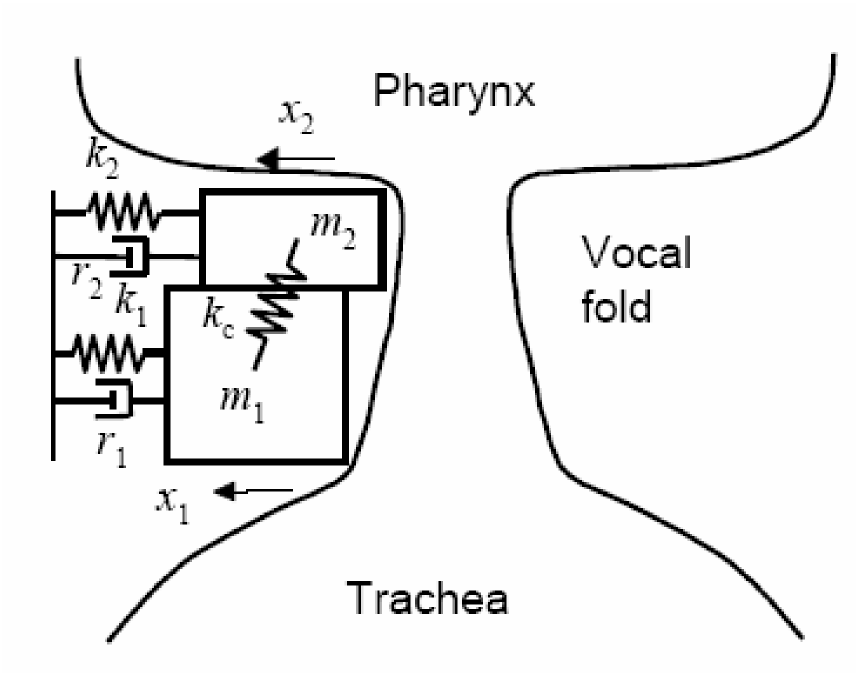


Figure 3.4 Two-mass model of the vocal folds [29].

3.5 The Two-Mass Model

The previous models are simple enough to permit an analytical study of their dynamics. However, to capture more details of the oscillation through computer simulations, more elaborated models are required. In increasing complexity, the next model is the popular two-mass model of Ishizaka and Flanagan [33], shown in Figure 3.4.

Each vocal fold is represented by a coupled pair of mass-damper-spring oscillators. When the upper mass oscillates with a phase delay in relation to the lower one, a wavelike motion in the airflow direction is reproduced. The equations of motion have the general form;

$$\begin{cases} m_1\ddot{x}_1 + b_1(x_1, \dot{x}_1) + s_1(x_1) + kc(x_1 - x_2) = f_1(x_1, x_2), \\ m_2\ddot{x}_2 + b_2(x_2, \dot{x}_2) + s_2(x_2) + kc(x_2 - x_1) = f_2(x_1, x_2), \end{cases} \tag{3.7}$$

Details of these equations may be found in the cited references [21, 23, 25, 24, 28, 33]. Figure 3.5 shows simulation results of glottal airflow when varying the subglottal pressure from 0 to a maximum value, and back to zero. Identification of the onset threshold is in general an easy task, because the oscillation builds up quickly at that point. The offset threshold, on the other hand, is more difficult and imprecise, because the oscillation amplitude tends to vanish slowly, and it is not clear at which point the rest position has become a stable equilibrium point. However, a clear hysteresis effect may be noted: oscillation stops at a lower value of the subglottal pressure than the value at which it starts.

Figure 3.6 shows plots of simulated oral airflow, as an example of the model's output. For this case, a two-tube approximation of the vocal tract in configuration for vowel /a/ [1] was added to the two-mass model. The simulations were obtained by varying the glottal half-width from 0.02 cm to 0.1 cm, and then back to the original value, following a sinusoidal pattern. This variation pattern imitates the glottal abduction-adduction gesture during the production of utterance /aha/ in running speech [25]. Oral airflow of adult man, woman and 5-year-old female child were simulated by adjusting the dimensions of the models to their respective anatomy.

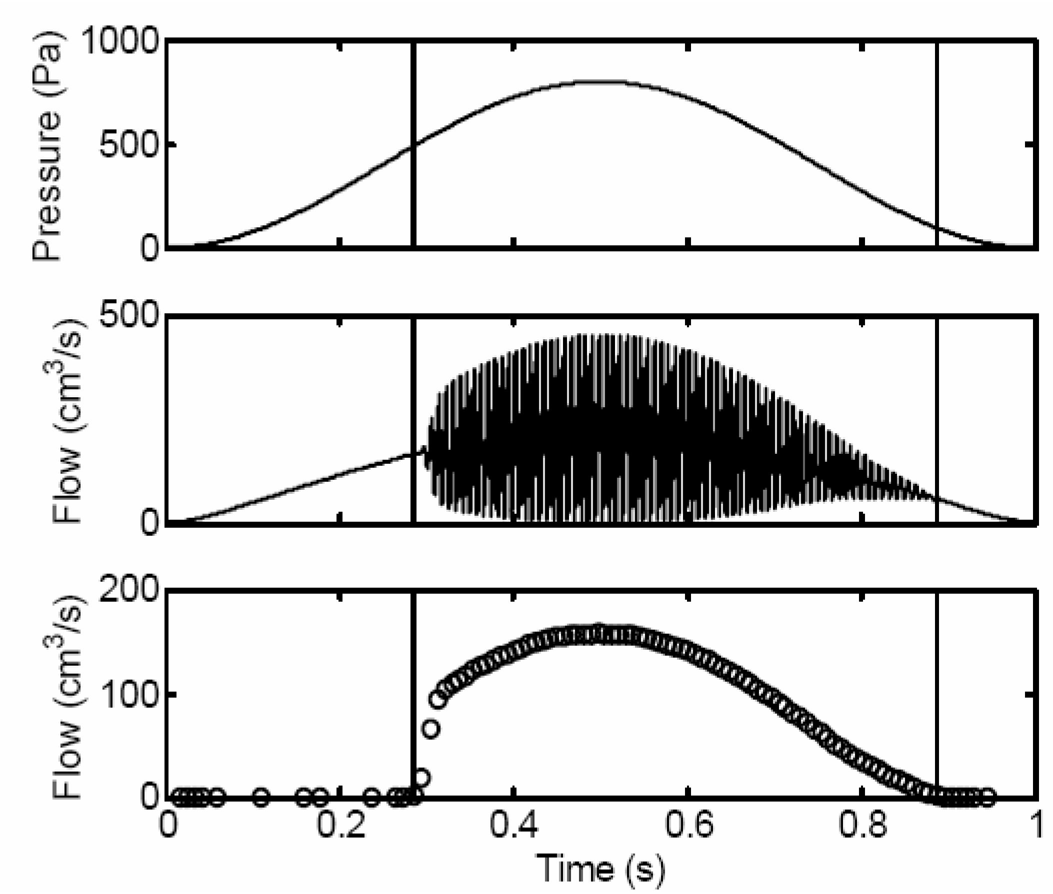


Figure 3.5 Simulation of glottal airflow when varying the subglottal pressure. Top: subglottal pressure, middle: glottal airflow, bottom: rms value of the AC glottal airflow. The left and right vertical lines mark the position of the oscillation onset and offset, respectively [29].

Comparing the plots, it was seen that the male flow has larger amplitude and lower fundamental frequency, as expected. In the female case, the glottal pulses stop at the peak abduction and restart at the end of the following adduction, with a hysteresis effect. In the child case, the glottal pulses stop even earlier than the female case, at a lower value of the glottal width. The plots show that the oscillation conditions become more restricted as the laryngeal size decreases. Smaller larynges have more restricted phonation regions because the medial surface of the vocal folds is smaller.

It is on this surface where the energy transfer from the airflow to the vocal fold oscillation takes place. In fact, (3.2) shows that the amount of energy transferred is proportional to the fold medial surface area. This result would explain the higher incidence of devoicing during glottal abduction-adduction for /h/ in running speech in women as compared to men [36]. As the vocal folds abduct, they easily reach the oscillation offset threshold in women, whereas men would require more extreme degrees of abduction. It would also agree with the observation of higher values of subglottal pressures in children during phonation [36], consequence of a larger value of the phonation threshold pressure.

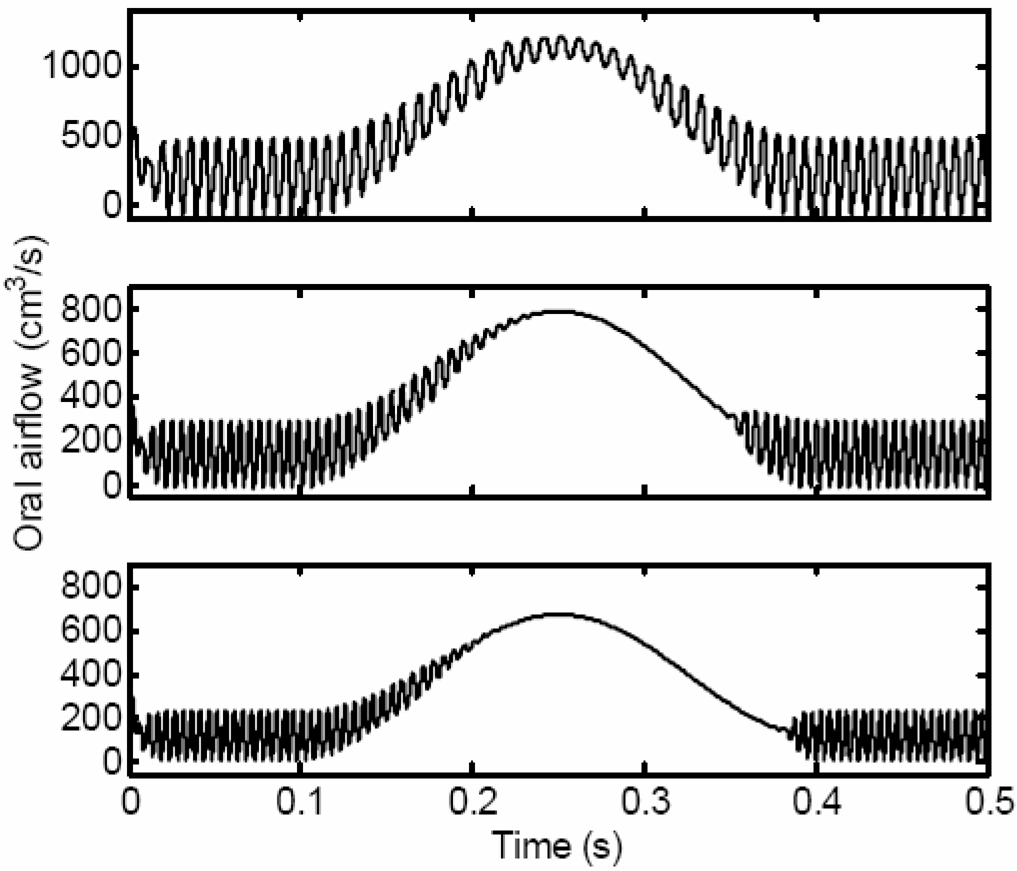


Figure 3.6 Oral airflow patterns during a vocal fold abduction-adduction gesture. Top: male adult, middle: female adult, bottom: 5-year-old child [29].

3.6 Acoustic Properties of the Vocal Tract

The vocal tract is effectively an acoustical tube, which has certain natural resonance and antiresonance frequencies depending on its shape. The resonance frequencies of the vocal tract are known as the formants for speech (represent the frequencies that pass the most acoustic energy from the source to the output), which are the key features used to distinguish between speech sounds. Formants are denoted by F1, F2, … . Due to the acoustic length of the vocal tract, there is on average 1 formant for each 1 kHz frequency band. The speech signal is a slowly time varying signal. A way of labeling events in speech is via the vocal cords.

Convention accepted: a three-state representation

- silence - no speech produced;
- unvoiced - the vocal cords are not vibrating;
- voiced - the vocal cords vibrate when air flows are coming from the lungs.

Different speech sounds are produced by varying the shape of the vocal tract using the speech articulators: vocal cords, tongue, lips, etc. Speech in any given language is composed of a set of phonemes, which are the basic units of speech and carry lexical information.

The phonemes can be grouped according to the method used in their production:

- *Vowels*: voiced sounds with an open vocal tract. Can be classified based on the formant frequencies (especially F1, F2), which depend mostly on the highest position of the tongue in the oral cavity and the roundedness of the lips.

- *Dipthongs*: a combination of two vowels: the articulators begin at a position corresponding to one vowel and smoothly move to a position corresponding to another (e.g. /ay/ as in ”buy”).

- *Semivowels*: voiced, vowel-like sounds but are not as steady and depend strongly on the adjacent phonemes, e.g. /w/,/l/.

- *Nasal Consonants*: formed by a complete restriction in the oral cavity, and the velum are lowered so that the sound radiated from the nostrils, e.g. /m/,/n/.

- *Fricatives*: determined by the place in the vocal tract where there is a constriction. This constriction causes the air flow to become turbulent at the constriction, which acts as acoustic source of noise. Fricatives can be voiced (/v/,/z/) or unvoiced (/f/,/s/).

- *Stops*: These are sounds produced by building up pressure behind a constriction somewhere in the oral tract and then suddenly releasing the pressure. May be voiced /d/,/b/,/g/ or unvoiced /t/,/p/,/k/.

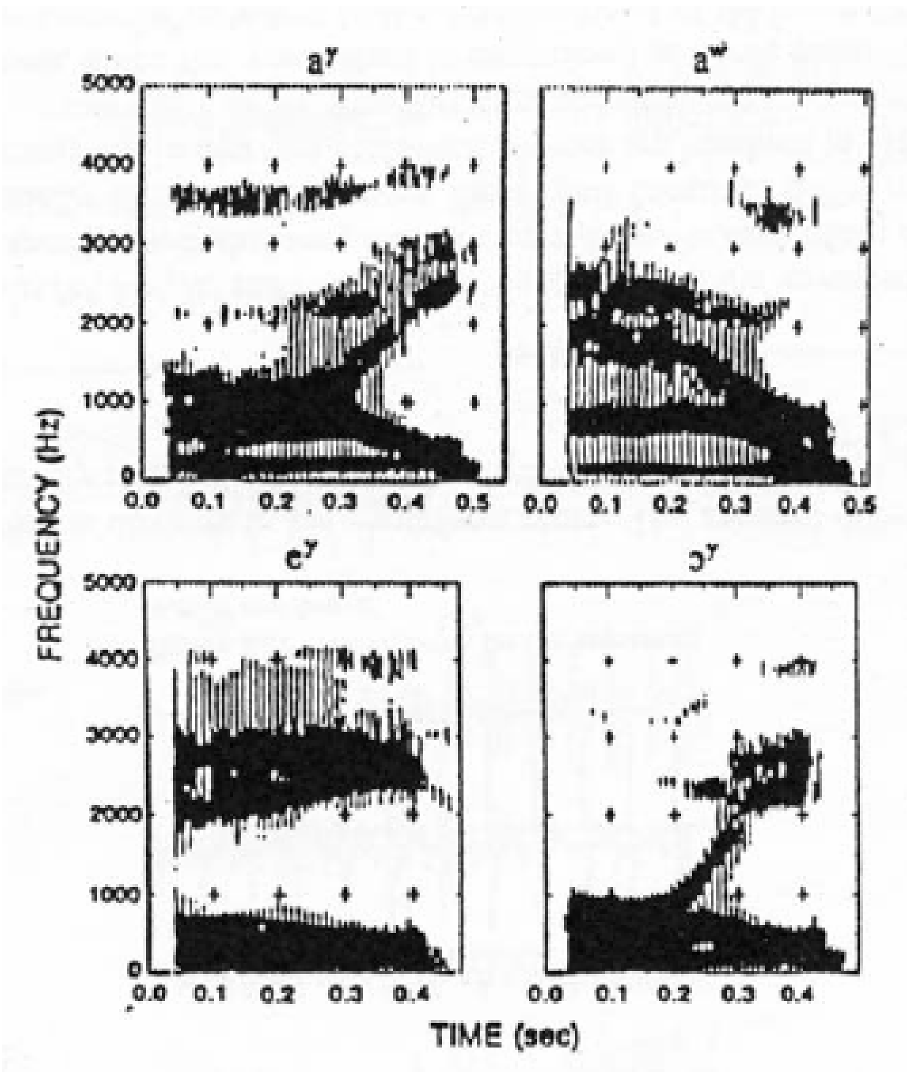


Figure 3.7 Dipthongs and the vowels [37].

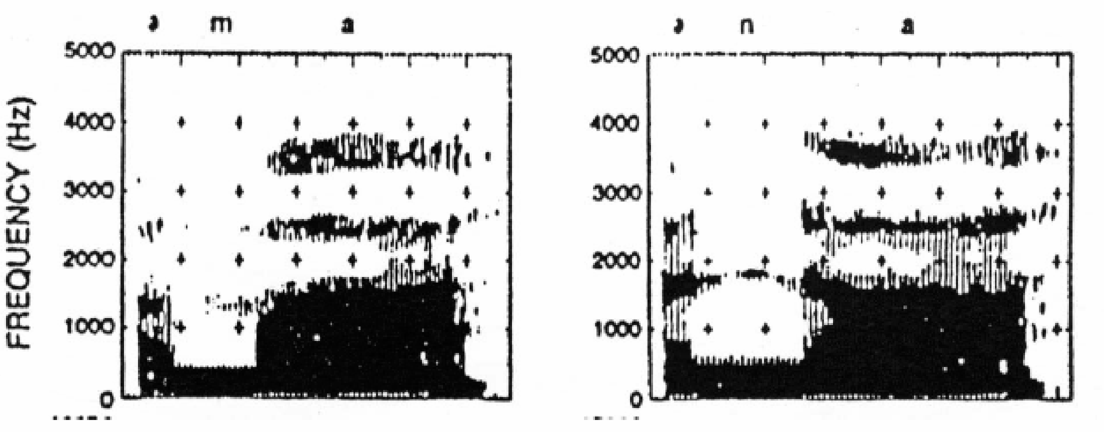


Figure 3.8 Spectrogram of two nasal consonants [37].

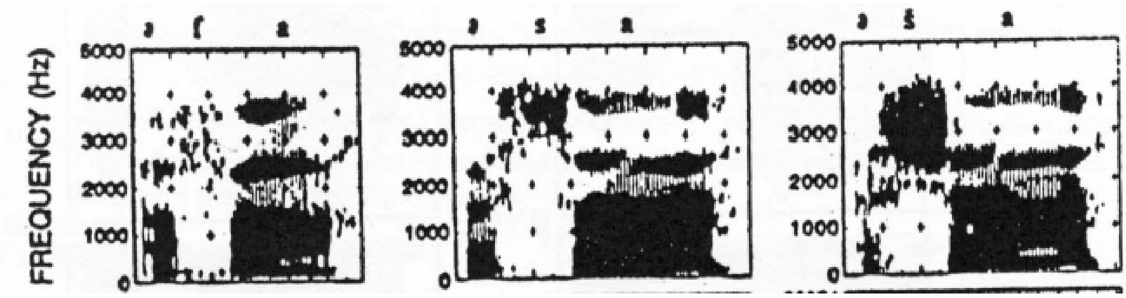


Figure 3.9 Spectrogram of fricatives (thaw, saw, shaw, chaw) [37].

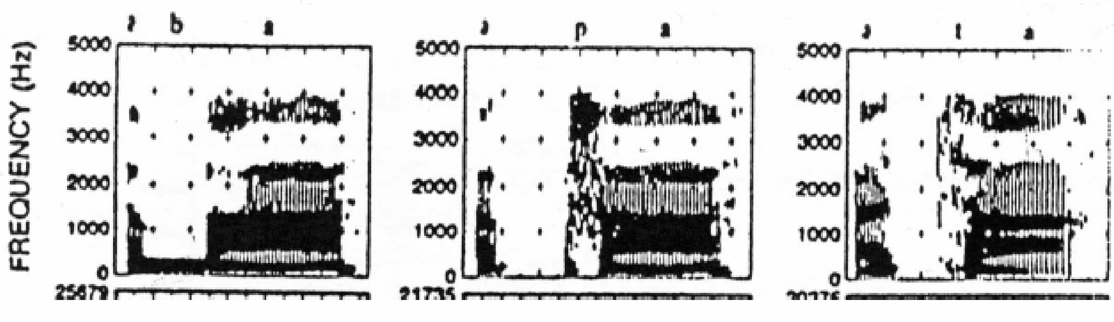


Figure 3.10 Spectograms of three stop consonant s(b,d,g) [37].

4. METHODOLOGY

4.1 Setup and Data Acquisition

The data acquisition setup used for acquiring the sound signals is shown in Figure 4.1. The setup consists of a Laptop PC, a special microphone, two stereo loudspeakers and a professional external sound card.

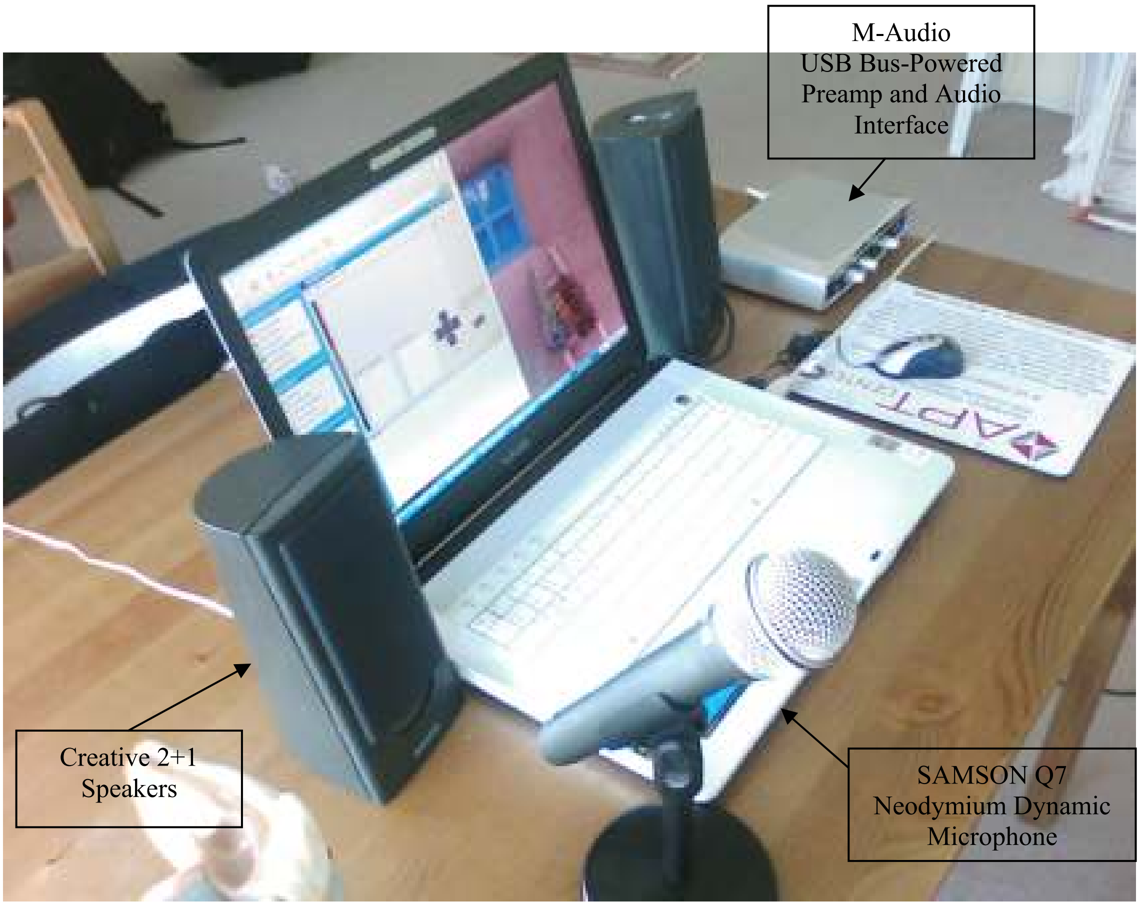


Figure 4.1 Experimental setup.

The recordings took 20 minutes for each subject; some lasted more than 20 minutes because of the irritation of the subjects due to hyoid pressure. At the beginning of the experiment subjects felt difficulty to apply backward pressure to their hyoids.

In the recordings Turkish vowels were used. These vowels are in the order as they are in the alphabet: A, E, I, İ, O, Ö, U, Ü. First, the subjects were asked to pronounce the vowels one by one continuously for three seconds, repeating each vowel twice after a pause for three seconds. The repeated recordings were used for control purposes. After locating the right place for the hyoid, subjects were asked to apply pressure on the hyoid and repeat the recordings. The sound acquisitions were monitored carefully to interrupt and correct false recordings.

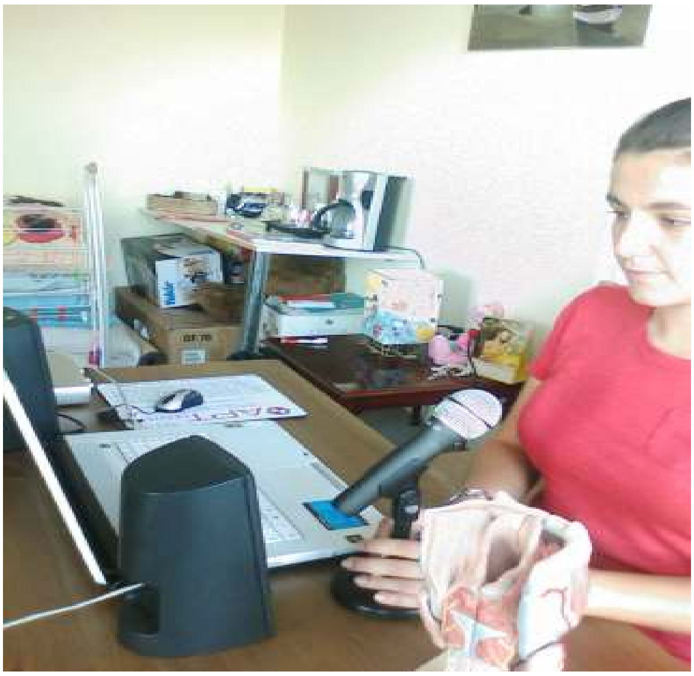


Figure 4.2 Data acquisition without pressure applied to the hyoid bone.

4.2 Subjects

For the study sound data from 20 subjects, 13 males and 7 females were acquired. The average age of the male subjects was 30.8 and those of the female subjects were 32.8. The experiments adhere to NIH ethical guidelines for testing human participants and were approved by the Ethics Committee for Human Subjects of Boğaziçi University.

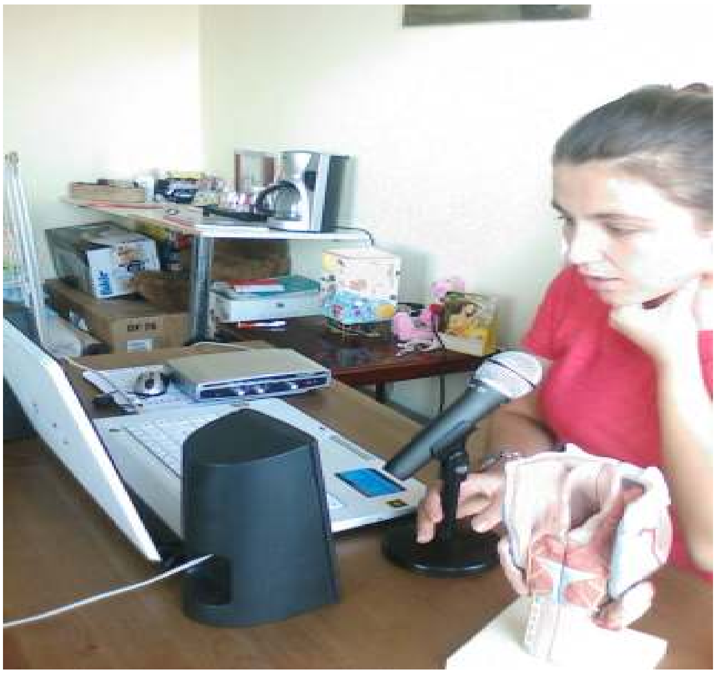


Figure 4.3 Data acquisition with pressure applied to the hyoid bone.

4.3 Software

For processing the acquired speech signals, for spectrogram, linear predictive coding (LPC) and for formant analysis we used *Multi-Speech*, model 3700 from KayPentax. A typical start-up menu of this software is shown in Figure 4.4. It is a low-cost, and easy-to-use speech analysis program with many optional programs and a wide array of functions and features for is suitable for research, teaching, and clinical applications for speech professionals [38]. Multi-Speech requires a host PC computer (Pentium >266MHz, CD-ROM, 16 MB RAM) and sound device. A summary of its basic features are presented in Appendix A. For sound recordings we used SesTek sound recording software with permission from Sestek Ses ve İletişim Bilgisayar Teknolojileri San. ve Tic. A.Ş. [39]. For data analysis and for statistical evaluations we used SPSS 15.0 for Windows [40].

4.4 Formant Analysis

A formant is a peak in an acoustic frequency spectrum which results from the resonant frequencies of any acoustical system. Formants are the distinguishing frequency components of human speech. Formants are the characteristic partials that identify vowels to the listener.

The formant with the lowest frequency is called $f1$, the second $f2$, and the third $f3$. Most often the two first formants, $f1$ and $f2$, are enough to disambiguate the vowel. These two formants are primarily determined by the position of the tongue.

After spectrogram and formant analysis LPC was used for speech analysis to get the direct formant frequency rates. LPC is one of the most useful methods for encoding good quality speech at a low bit rate and provides extremely accurate estimates of speech parameters.

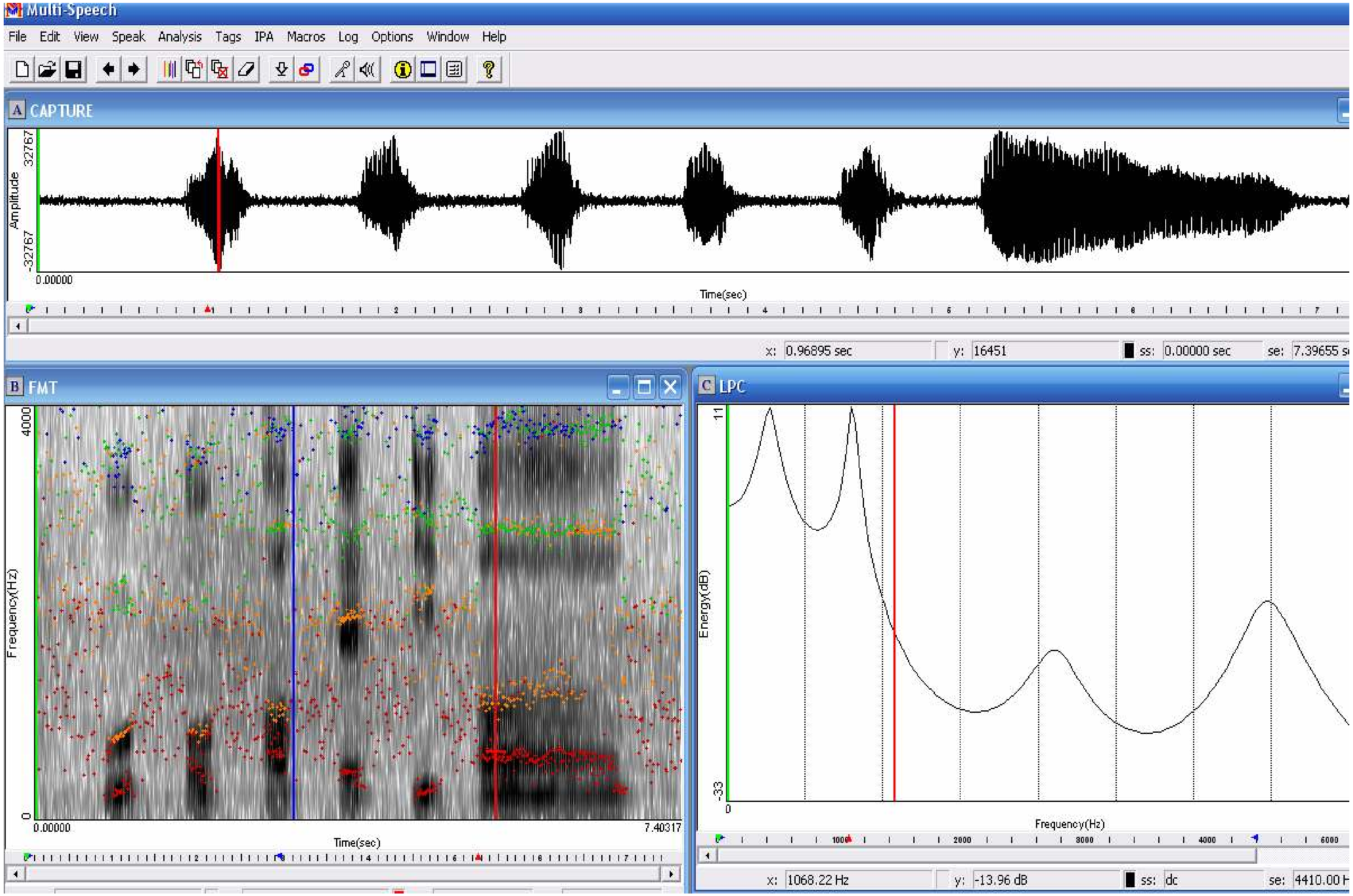


Figure 4.4 KAY Multispeech software for spectrogram, formant analysis and LPC.

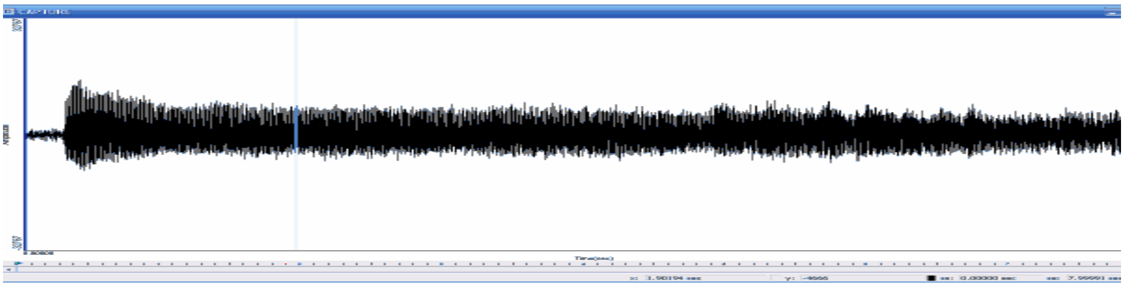


Figure 4.5 Recorded vowel seen as a signal

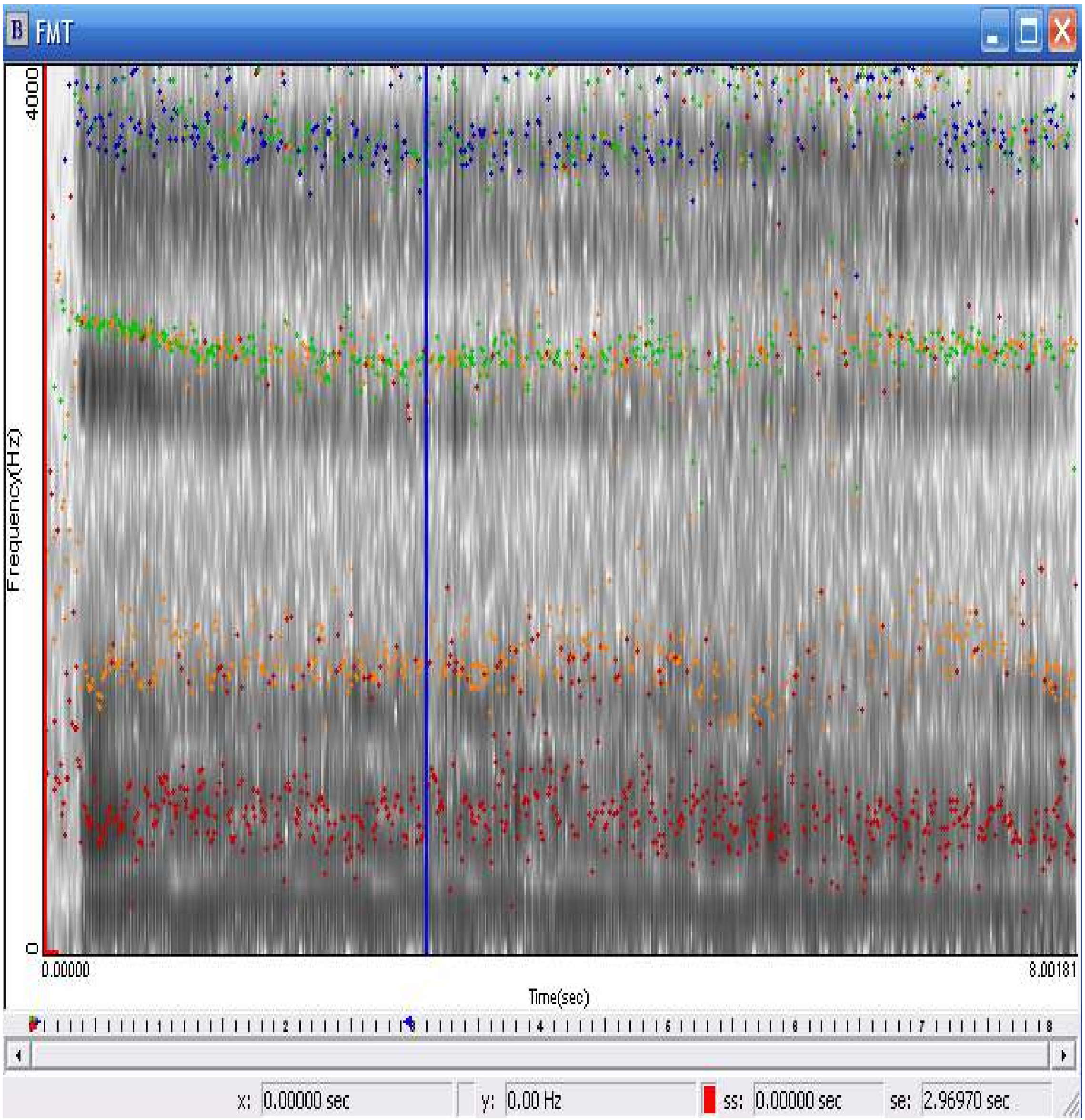


Figure 4.6 Formant analysis.

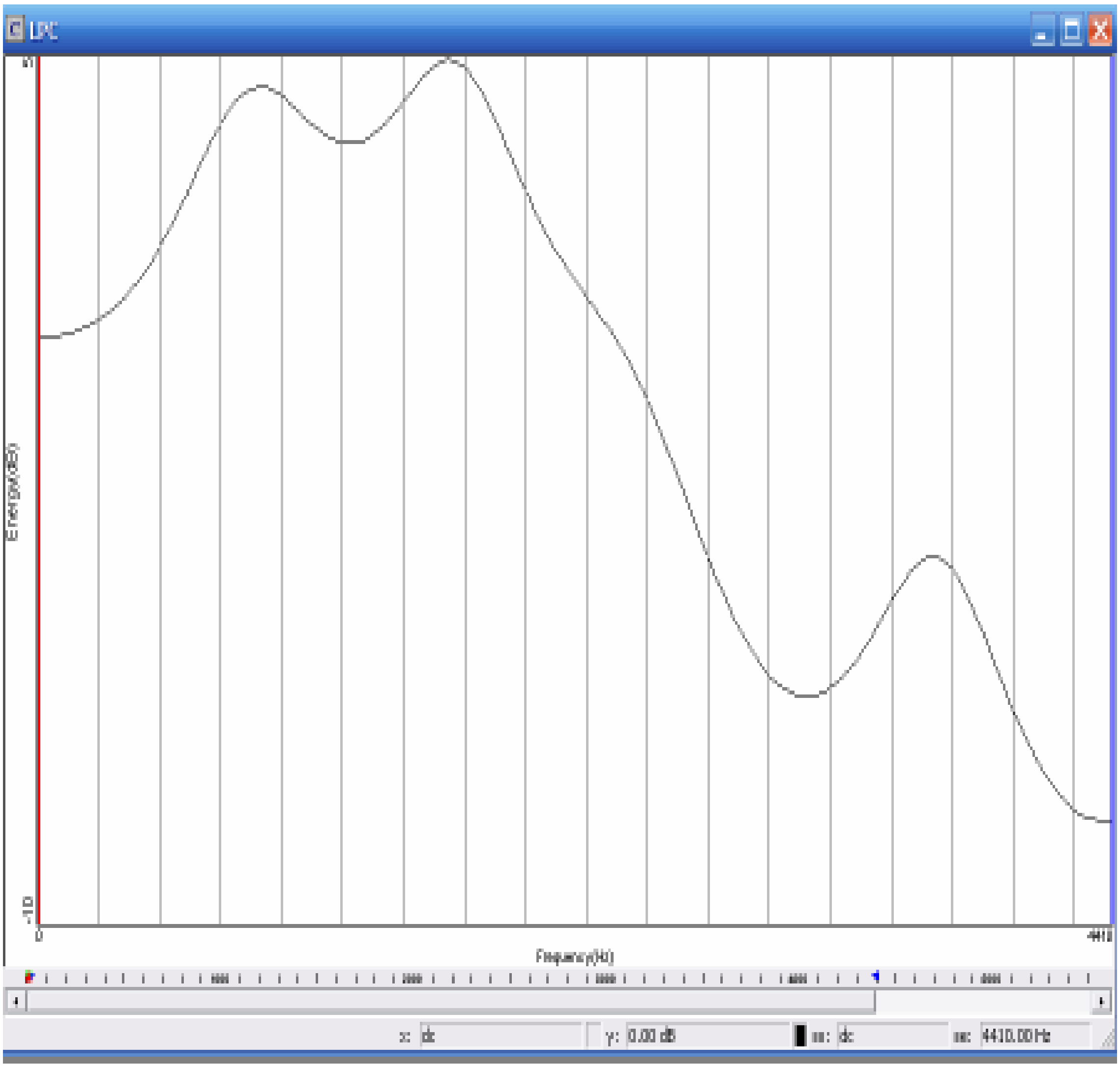


Figure 4.7 LPC (Linear Predictive Coding).

The frequencies obtained from the analysis of each subject's records are noted to use in SPSS later. SPSS is a computer program used for statistical analysis. All the formant frequencies obtained from males and females for each vowel before and after pressure to the hyoids were written into SPSS separately and together. To get a better comparison, 7 of 13 males chosen to equal the number of gender.

Table 4.1
Vowel formants.

Vowel	Main formant region
u	200 to 400
o	400 to 600
a	800 to 1200
e	400 to 600 and 2200 to 2600
i	200 to 400 and 3000 to 3500

Table 4.2
F1 of all 14 subjects for Paired Samples T-Test analysis.

F1 [DataSet1] - SPSS Data Editor

File Edit View Data Transform Analyze Graphs Utilities Window Help

10 : once.o 4436

	cins	once.a	sonra.a	once.e	sonra.e	once.l	sonra.l	once.ii	sonra.ii	once.o	sonra.o	once.oo	sonra.oo	once.u	sonra.u	once.uu
1	1,00	3240,00	3212,00	3853,00	5170,00	5281,00	6359,00	3410,00	4347,00	4640,00	5624,00	4700,00	5988,00	4108,00	4935,00	4710,0
2	1,00	5306,00	5724,00	5174,00	5367,00	3317,00	3278,00	4693,00	4259,00	2955,00	2951,00	3156,00	5064,00	2978,00	3196,00	5159,0
3	1,00	5398,00	4661,00	5373,00	4701,00	3538,00	5576,00	5845,00	4196,00	5941,00	4621,00	4872,00	4874,00	4034,00	3436,00	4532,0
4	1,00	4567,00	5024,00	5003,00	3085,00	5075,00	6155,00	5157,00	2894,00	5311,00	4317,00	1598,00	3226,00	5268,00	4737,00	5342,0
5	1,00	5906,00	4314,00	5660,00	5858,00	6069,00	5956,00	5000,00	4945,00	5343,00	5929,00	5788,00	5900,00	554,00	5988,00	5494,0
6	1,00	2295,00	2826,00	4794,00	4443,00	3072,00	5305,00	4871,00	6824,00	3256,00	6606,00	4987,00	4170,00	3856,00	3220,00	5894,0
7	1,00	3554,00	3611,00	2814,00	3094,00	4046,00	3767,00	2901,00	3553,00	3493,00	3557,00	3173,00	2189,00	4515,00	3161,00	2737,0
8	2,00	4873,00	6073,00	5099,00	4545,00	4775,00	2959,00	5314,00	3627,00	5035,00	5205,00	3142,00	4427,00	5370,00	6073,00	5104,0
9	2,00	4061,00	6065,00	7362,00	5648,00	6329,00	6172,00	4733,00	5381,00	6567,00	6358,00	7144,00	5878,00	5756,00	4337,00	7749,0
10	2,00	1094,00	5000,00	4947,00	5403,00	6150,00	4403,00	5408,00	3357,00	4436,00	4760,00	5009,00	4990,00	4864,00	5576,00	5048,0
11	2,00	5565,00	5773,00	5920,00	5251,00	5348,00	5003,00	5653,00	5372,00	4866,00	4821,00	5189,00	3024,00	4757,00	5529,00	5534,0
12	2,00	3204,00	5243,00	4547,00	4334,00	4386,00	4516,00	5565,00	4317,00	4345,00	5361,00	5338,00	4569,00	5539,00	4169,00	4018,0
13	2,00	5161,00	3748,00	5368,00	5005,00	3309,00	5962,00	4849,00	5183,00	5225,00	4447,00	4814,00	5548,00	5154,00	6387,00	5163,0
14	2,00	4691,00	6017,00	2986,00	4617,00	5148,00	4924,00	3857,00	4564,00	4314,00	6132,00	5072,00	4045,00	4548,00	5048,00	3740,0
15																
16																
17																
18																

Data View | Variable View

SPSS Processor is ready

The Paired Samples T Test compares the means of two variables. It computes the difference between the two variables for each case, and tests to see if the average difference is significantly different from zero. If the significance value is less than 0.05, there is a significant difference. If the significance value is greater than 0.05, there is no significant difference.

5. RESULTS

Formants which appear as peaks in an acoustic frequency spectrum due to resonance are the distinguishing frequency components of human speech and reflect voice quality. We used *Multi-Speech*, model 3700 speech processing software from KayPentax to determine the formants. The results of this analysis are presented in Tables 5.1-5.8. The formants were later studied using SPSS 15 for Windows for paired samples statistics. The results are presented in Tables 5.9-5.20. In these tables, the characters a, e, ı, i, o, ö, u, and ü represent the Turkish vowels. The cap “ˆ” is used to show that the related data is acquired with pressure applied on the hyoid bone.

Table 5.1
F1 for men. The cap “ˆ” shows that the related data is acquired with pressure applied to the hyoid.

a	â	e	ê	ı	ı̂	i	î
3240.00	3212.00	3853.00	5170.00	5281.00	6359.00	3410.00	4347.00
5306.00	5724.00	5174.00	5367.00	3317.00	3278.00	4693.00	4259.00
5398.00	4661.00	5373.00	4701.00	3538.00	5576.00	5845.00	4196.00
4567.00	5024.00	5003.00	3085.00	5075.00	6155.00	5157.00	2894.00
5906.00	4314.00	5660.00	5858.00	6069.00	5956.00	5000.00	4945.00
2295.00	2826.00	4794.00	4443.00	3072.00	5305.00	4871.00	6824.00
3554.00	3611.00	2814.00	3094.00	4046.00	3767.00	2901.00	3553.00
o	ô	ö	ö̂	u	û	ü	ü̂
4640.00	5624.00	4700.00	5988.00	4108.00	4935.00	4710.00	6247.00
2955.00	2951.00	3156.00	5064.00	2978.00	3196.00	5159.00	5220.00
5941.00	4621.00	4872.00	4874.00	4034.00	3436.00	4532.00	1060.00
5311.00	4317.00	1598.00	3226.00	5268.00	4737.00	5342.00	3023.00
5343.00	5929.00	5788.00	5900.00	554.00	5988.00	5494.00	5023.00
3256.00	6606.00	4987.00	4170.00	3856.00	3220.00	5894.00	3159.00
3493.00	3557.00	3173.00	2189.00	4515.00	3161.00	2737.00	2751.00

Table 5.2
F1 for women.

a	â	e	ê	ı	ı̂	i	î
4873.00	6073.00	5099.00	4545.00	4775.00	2959.00	5314.00	3627.00
4061.00	6065.00	7362.00	5648.00	6329.00	6172.00	4733.00	5381.00
1094.00	5000.00	4947.00	5403.00	6150.00	4403.00	5408.00	3357.00
5565.00	5773.00	5920.00	5251.00	5348.00	5003.00	5653.00	5372.00
3204.00	5243.00	4547.00	4334.00	4386.00	4516.00	5565.00	4317.00
5161.00	3748.00	5368.00	5005.00	3309.00	5962.00	4849.00	5183.00
4691.00	6017.00	2986.00	4617.00	5148.00	4924.00	3857.00	4564.00
o	ô	ö	ö̂	u	û	ü	ü̂
5035.00	5205.00	3142.00	4427.00	5370.00	6073.00	5104.00	5161.00
6567.00	6358.00	7144.00	5878.00	5756.00	4337.00	7749.00	5825.00
4436.00	4760.00	5009.00	4990.00	4864.00	5576.00	5048.00	4714.00
4866.00	4821.00	5189.00	3024.00	4757.00	5529.00	5534.00	2503.00
4345.00	5361.00	5338.00	4569.00	5539.00	4169.00	4018.00	5958.00
5225.00	4447.00	4814.00	5548.00	5154.00	6387.00	5163.00	4474.00
4314.00	6132.00	5072.00	4045.00	4548.00	5048.00	3740.00	5189.00

Table 5.3
F2 for men.

a	â	e	ê	ı	ı̂	i	î
6889.00	6981.00	7389.00	7770.00	8115.00	8248.00	7688.00	7940.00
8675.00	8193.00	8968.00	8932.00	6406.00	5491.00	8936.00	8803.00
8844.00	7153.00	8972.00	8709.00	7224.00	8645.00	9147.00	6697.00
8374.00	8616.00	8551.00	8918.00	8721.00	8067.00	8780.00	5019.00
8998.00	8475.00	7987.00	9288.00	8052.00	9235.00	6952.00	6359.00
6912.00	9706.00	8204.00	7804.00	8345.00	8891.00	6410.00	8392.00
8034.00	7729.00	6705.00	7846.00	6702.00	7077.00	7122.00	7731.00
o	ô	ö	ö̂	u	û	ü	ü̂
7932.00	8451.00	7897.00	8021.00	7894.00	8126.00	7900.00	8297.00
6636.00	8853.00	7275.00	8986.00	6075.00	7780.00	8497.00	9021.00
8692.00	7691.00	8341.00	8096.00	7610.00	4690.00	8359.00	4971.00
8836.00	8979.00	7074.00	8821.00	8903.00	8713.00	8911.00	8301.00
9071.00	9275.00	9052.00	8268.00	6179.00	9035.00	6300.00	7454.00
8011.00	9072.00	8427.00	8181.00	8369.00	8106.00	8878.00	7713.00
3850.00	7896.00	7038.00	7212.00	9180.00	7353.00	7096.00	8256.00

Table 5.4
F2 for women.

a	â	e	ê	ı	ı̂	i	î
8673.00	7962.00	8752.00	8932.00	8961.00	8152.00	9136.00	7661.00
7398.00	7029.00	7525.00	6745.00	7874.00	7159.00	7840.00	6746.00
4390.00	8073.00	9072.00	9533.00	8915.00	8223.00	9212.00	7653.00
8735.00	9683.00	9862.00	8146.00	8922.00	8742.00	8041.00	9241.00
7798.00	7741.00	7871.00	8176.00	7169.00	7620.00	7738.00	7634.00
8881.00	8462.00	8524.00	8523.00	7360.00	8260.00	8161.00	8152.00
8163.00	7192.00	6726.00	7640.00	9014.00	8576.00	7504.00	8407.00
o	ô	ö	ö̂	u	û	ü	ü̂
8400.00	8267.00	9763.00	9462.00	8692.00	9687.00	8304.00	8371.00
8186.00	8949.00	7623.00	6980.00	7187.00	6857.00	8021.00	8393.00
8595.00	8713.00	8973.00	9849.00	8493.00	9043.00	8327.00	9132.00
8169.00	9065.00	8509.00	6769.00	8043.00	8578.00	9217.00	5770.00
7925.00	6864.00	7947.00	7791.00	7906.00	7837.00	8347.00	8759.00
9102.00	8177.00	9172.00	8811.00	8629.00	10917.00	8813.00	7113.00
7197.00	6472.00	5541.00	9026.00	8227.00	8853.00	7590.00	8875.00

Table 5.5
F3 for men.

a	â	e	ê	ı	î	i	i̇̂
9341.00	11295.00	10345.00	10884.00	11268.00	10271.00	9225.00	11840.00
11740.00	11995.00	12481.00	12172.00	10590.00	9392.00	12458.00	12021.00
10847.00	10243.00	12233.00	14105.00	9952.00	12117.00	11375.00	9572.00
11213.00	12104.00	12108.00	11445.00	12163.00	10145.00	12052.00	9382.00
10667.00	11986.00	9412.00	10891.00	9474.00	12960.00	9756.00	9412.00
10054.00	12052.00	10469.00	10518.00	9985.00	11365.00	10352.00	12738.00
11046.00	10706.00	9453.00	8932.00	9191.00	9817.00	9827.00	11375.00
o	ô	ö	ö̂	u	û	ü	ü̂
1142.00	11870.00	11120.00	11377.00	11074.00	12064.00	11585.00	11367.00
12658.00	12631.00	10482.00	12256.00	11606.00	10903.00	12037.00	12490.00
12444.00	11866.00	11871.00	11120.00	10637.00	10697.00	11803.00	9542.00
11621.00	10346.00	9408.00	10766.00	12492.00	10397.00	12400.00	10571.00
9845.00	12685.00	11193.00	9122.00	9645.00	12649.00	9381.00	9393.00
10491.00	12944.00	12119.00	10686.00	10021.00	10808.00	12395.00	10553.00
9099.00	11909.00	9489.00	8798.00	12096.00	10900.00	9483.00	9769.00

Table 5.6
F3 for women.

a	â	e	ê	ı	î	i	i̇̂
11557.00	12199.00	11960.00	12841.00	12075.00	12111.00	12805.00	10858.00
10710.00	9972.00	11201.00	10795.00	11254.00	10496.00	11066.00	10440.00
10065.00	11518.00	13184.00	13017.00	12833.00	11679.00	12213.00	11365.00
11901.00	12444.00	11319.00	11741.00	12693.00	11865.00	11677.00	12660.00
10790.00	10738.00	11450.00	11708.00	11314.00	10394.00	12087.00	11547.00
12090.00	11947.00	11052.00	12402.00	11718.00	12075.00	12043.00	11627.00
9940.00	11538.00	11354.00	11087.00	12466.00	11032.00	10955.00	11890.00
o	ô	ö	ö̂	u	û	ü	ü̂
11490.00	12032.00	12609.00	13278.00	12220.00	13488.00	11663.00	12345.00
11376.00	12175.00	11032.00	9920.00	10490.00	9358.00	10936.00	12120.00
11549.00	12249.00	11911.00	10586.00	12194.00	12447.00	11795.00	11477.00
11587.00	12995.00	11896.00	11412.00	12608.00	12185.00	12467.00	10857.00
11253.00	9452.00	11828.00	10326.00	11852.00	11127.00	10904.00	11857.00
12327.00	11541.00	12246.00	11577.00	10810.00	14146.00	11742.00	10891.00
11103.00	11622.00	8961.00	11577.00	10836.00	11880.00	10753.00	11910.00

Table 5.7
F4 for men.

a	â	e	ê	ı	î	i	i̇̂
15237.00	15563.00	14848.00	17309.00	17309.00	17322.00	14941.00	16780.00
15058.00	15464.00	16021.00	15785.00	12516.00	13017.00	15386.00	15802.00
17322.00	13746.00	17418.00	14160.00	14884.00	17202.00	17242.00	13988.00
15035.00	15907.00	14994.00	15056.00	16663.00	15010.00	14769.00	13190.00
17076.00	14357.00	14547.00	16410.00	14964.00	17098.00	15032.00	14925.00
14638.00	14604.00	15960.00	15603.00	15012.00	16820.00	15201.00	16398.00
14994.00	14815.00	13449.00	12737.00	13473.00	14492.00	13823.00	13172.00
o	ô	ö	ö̂	u	û	ü	ü̂
17437.00	17372.00	16327.00	17437.00	17458.00	17732.00	14961.00	17026.00
15479.00	16199.00	14422.00	14830.00	15867.00	14714.00	16008.00	15259.00
18100.00	15309.00	17731.00	14797.00	14324.00	14748.00	17521.00	13167.00
15254.00	16885.00	13193.00	17804.00	15066.00	15370.00	16053.00	16499.00
14707.00	17287.00	17095.00	14432.00	14857.00	16137.00	15254.00	12468.00
15395.00	16209.00	16328.00	16026.00	14587.00	15006.00	16786.00	15158.00
12662.00	16626.00	13030.00	13397.00	17392.00	12339.00	13725.00	12881.00

Table 5.8
F4 for women.

a	â	e	ê	ı	î	i	i̇̂
15269.00	16269.00	14640.00	16362.00	15184.00	16349.00	17157.00	14635.00
13693.00	13314.00	17115.00	16673.00	17082.00	15169.00	17562.00	13588.00
12912.00	14788.00	16196.00	17233.00	17351.00	14788.00	16252.00	15254.00
15428.00	16638.00	14370.00	14939.00	16120.00	15338.00	15155.00	17074.00
15119.00	16756.00	17436.00	17258.00	14944.00	12846.00	12170.00	12873.00
16094.00	14804.00	14621.00	15558.00	15018.00	15684.00	15032.00	15010.00
15042.00	17844.00	13739.00	15245.00	17484.00	15460.00	14898.00	17512.00
o	ô	ö	ö̂	u	û	ü	ü̂
14974.00	15201.00	14581.00	17296.00	15742.00	16758.00	15592.00	15273.00
17283.00	12362.00	17500.00	14076.00	17488.00	13020.00	16943.00	12273.00
15049.00	14711.00	15727.00	16509.00	16189.00	15861.00	15399.00	15128.00
14781.00	17521.00	15082.00	16424.00	16660.00	15428.00	15991.00	13457.00
14922.00	14374.00	17131.00	14910.00	17395.00	15164.00	14842.00	17628.00
16544.00	14387.00	16363.00	13455.00	14084.00	16264.00	15704.00	13573.00
15397.00	14201.00	13231.00	15346.00	17357.00	17650.00	15264.00	17437.00

Table 5.9
Paired samples statistics for F1.

	Mean	N	Std. Deviation	Std. Error Mean
a	4208.21	14	1381.47	369.21
â	4806.50	14	1108.80	296.34
e	4921.43	14	1159.91	310.00
ê	4751.50	14	840.22	224.56
ı	4703.07	14	1113.18	297.51
î	5023.93	14	1111.42	297.04
i	4804.00	14	861.88	230.35
î̇	4487.07	14	1013.59	270.89
o	4694.79	14	1006.25	268.93
ô	5049.21	14	1051.81	281.11
ö	4570.14	14	1378.16	368.33
ö̂	4563.71	14	1152.95	308.14
u	4378.64	14	1332.99	356.26
û	4699.43	14	1138.65	304.32
ü	5016.00	14	1144.89	305.98
ü̂	4307.64	14	1544.86	412.88

Table 5.10
Paired samples correlations for F1.

Pairs	N	Correlation	Sig.
(a, â)	14	.336	.240
(e, ê)	14	.560	.037
(ı, î)	14	.263	.364
(i, î̇)	14	.068	.817
(o, ô)	14	.324	.259
(ö, ö̂)	14	.539	.047
(u, û)	14	.025	.932
(ü, ü̂)	14	.179	.540

Table 5.11
Paired samples test for F1.

Paired Differences								
Pair	Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference		t	df	Sig. (2-tailed)
				Lower	Upper			
(a, â)	-598.29	1451.91	388.04	-1436.59	240.02	-1.542	13	.147
(e, ê)	169.93	979.64	261.82	-395.70	735.56	.649	13	.528
(ı, î)	-320.86	1350.73	361.00	-1100.75	459.03	-.889	13	.390
(i, î)	316.93	1284.92	343.41	-424.96	1058.82	.923	13	.373
(o, ô)	-354.43	1197.18	319.96	-1045.66	336.81	-1.108	13	.288
(ö, ô̈)	6.43	1231.76	329.20	-704.77	717.62	.020	13	.985
(u, û)	-320.79	1731.04	462.64	-1320.26	678.69	-.693	13	.500
(ü, û̈)	708.36	1750.11	467.74	-302.13	1718.84	1.514	13	.154

Table 5.12
Paired samples statistics for F2.

				Std. Error
	Mean	N	Std. Deviation	Mean
a	7911.71	14	1235.54	330.21
â	8071.07	14	878.23	234.72
e	8222.00	14	916.06	244.83
ê	8354.43	14	758.46	202.71
ı	7984.29	14	887.27	237.13
î	8027.57	14	953.21	254.76
i	8047.64	14	894.01	238.93
i̇̂	7602.50	14	1095.80	292.86
o	7900.14	14	1350.64	360.97
ô	8337.43	14	852.49	227.84
ö	8045.14	14	1094.57	292.54
ö̂	8305.21	14	917.49	245.21
u	7956.21	14	934.09	249.65
û	8255.36	14	1439.93	384.84
ü	8182.86	14	777.38	207.76
ü̂	7887.57	14	1219.71	325.98

Table 5.13
Paired samples correlations for F2.

Pairs	N	Correlation	Sig.
(a, â)	14	.336	.240
(e, ê)	14	.560	.037
(ı, î)	14	.263	.364
(i, i̇̂)	14	.068	.817
(o, ô)	14	.324	.259
(ö, ö̂)	14	.539	.047
(u, û)	14	.025	.932
(ü, ü̂)	14	.179	.540

Table 5.14
Paired samples test for F2.

Pair	Paired Differences					t	df	Sig. (2-tailed)
	Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
				Lower	Upper			
(a, â)	-159.36	1446.56	386.61	-994.57	675.86	-.412	13	.687
(e, ê)	-132.43	783.76	209.47	-584.96	320.10	-.632	13	.538
(ı, î)	-43.29	784.27	209.61	-496.11	409.54	-.207	13	.840
(i, î)	445.14	1523.53	407.18	-434.52	1324.80	1.093	13	.294
(o, ô)	-437.29	1386.50	370.56	-1237.83	363.25	-1.180	13	.259
(ö, ȫ)	-260.07	1313.02	350.92	-1018.19	498.05	-.741	13	.472
(u, û)	-299.14	1497.53	400.23	-1163.79	565.50	-.747	13	.468
(ü, ü̂)	295.29	1580.66	422.45	-617.36	1207.93	.699	13	.497

Table 5.15
Paired samples statistics for F3.

				Std. Error
	Mean	N	Std. Deviation	Mean
a	10854.36	14	809.36	216.31
â	11481.21	14	779.30	208.28
e	11287.21	14	1094.62	292.55
ê	11609.86	14	1272.89	340.19
ı	11212.57	14	1205.73	322.24
î	11122.79	14	1055.41	282.07
i	11277.93	14	1118.01	298.80
i̇̂	11194.79	14	1118.52	298.94
o	10570.36	14	2880.78	769.92
ô	11879.79	14	971.30	259.59
ö	11154.64	14	1157.98	309.48
ö̂	10914.36	14	1175.96	314.29
u	11327.21	14	951.76	254.37
û	11646.36	14	1288.60	344.39
ü	11381.71	14	988.38	264.15
ü̂	11081.57	14	1028.62	274.91

Table 5.16
Paired samples correlations for F3.

Pairs	N	Correlation	Sig.
(a, â)	14	0.310	0.281
(e, ê)	14	0.778	0.001
(ı, î)	14	0.067	0.819
(i, i̇̂)	14	0.007	0.982
(o, ô)	14	0.033	0.910
(ö, ö̂)	14	0.298	0.301
(u, û)	14	0.062	0.834
(ü, ü̂)	14	0.303	0.293

Table 5.17
Paired samples test for F3.

Pair	Paired Differences: Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference: Lower	Upper	t	df	Sig. (2-tailed)
(a, â)	-626.86	933.46	249.48	-1165.82	-87.89	-2.513	13	.026
(e, ê)	-322.64	805.81	215.36	-787.90	142.62	-1.498	13	.158
(ı, î)	89.79	1547.90	413.69	-803.94	983.51	.217	13	.832
(i, î)	83.14	1586.64	424.05	-832.96	999.24	.196	13	.848
(o, ô)	-1309.43	3070.46	820.61	-3082.26	463.40	-1.596	13	.135
(ö, ô̈)	240.29	1382.69	369.54	-558.06	1038.63	.650	13	.527
(u, û)	-319.14	1553.94	415.31	-1216.36	578.07	-.768	13	.456
(ü, û̈)	300.14	1191.48	318.44	-387.80	988.08	.943	13	.363

Table 5.18
Paired samples statistics for F4.

				Std. Error
	Mean	N	Std. Deviation	Mean
a	15208.35	14	1137.53	304.01
â	15347.79	14	1247.45	333.40
e	15382.43	14	1315.45	351.57
ê	15737.71	14	1289.73	344.69
ı	15571.71	14	1502.86	401.66
î	15471.07	14	1424.10	380.61
i	15330.00	14	1413.20	377.69
i̊̂	15014.36	14	1537.68	410.96
o	15570.29	14	1381.43	369.20
ô	15617.43	14	1507.83	402.98
ö	15552.93	14	1642.25	438.91
ȫ̂	15481.36	14	1454.90	388.84
u	16033.29	14	1278.84	341.78
û	15442.21	14	1518.21	405.76
ü	15717.36	14	961.35	256.93
ü̂	14801.93	14	1859.39	496.94

Table 5.19
Paired samples correlations for F4.

Pairs	N	Correlation	Sig.
(a, â)	14	-.061	.836
(e, ê)	14	.392	.165
(ı, î)	14	.315	.273
(i, i̊̂)	14	.096	.744
(o, ô)	14	-.330	.249
(ö, ȫ̂)	14	-.246	.396
(u, û)	14	-.106	.719
(ü, ü̂)	14	-.255	.379

Table 5.20
Paired samples test for F4.

Pair	Paired Differences: Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference: Lower	Upper	t	df	Sig. (2-tailed)
(a, â)	-139.43	1738.60	464.66	-1143.27	864.41	-.300	13	.769
(e, ê)	-355.29	1435.97	383.78	-1184.39	473.82	-.926	13	.371
(ı, î)	100.64	1714.53	458.23	-889.30	1090.58	.220	13	.830
(i, î)	315.64	1985.90	530.75	-830.98	1462.27	.595	13	.562
(o, ô)	-47.14	2357.42	630.05	-1408.28	1313.99	-.075	13	.941
(ö, ô̈)	71.57	2447.67	654.17	-1341.67	1484.81	.109	13	.915
(u, û)	591.07	2086.01	557.51	-613.35	1795.50	1.060	13	.308
(ü, û̈)	915.43	2300.57	614.85	-412.88	2243.74	1.489	13	.160

6. CONCLUSIONS

6.1 Conclusions

This study was designed to evaluate the characteristics of phonetic structural changes of Turkish vowels under adjustment at the low pharyngeal (hyoid) region airway and its effects on the voice quality.

Turkish phonetic structure of vowels has not been deeply analyzed. Current knowledge on mathematical modeling studies representing each vowel structure reveals no difference than the standard international parameters. The anatomic correlation of the quality of voice and formantic structure defining each vowel are mostly represented by the constriction sites of lips, tongue and the pharynx. The lower pharynx at the level of hyoid bone has been shown in one study that it has the potential for manipulation in order to change the airway area by surgical intervention [40]. In the same report it has been shown that the change of upper laryngeal output area has a meaningful change, hence this area has a significant role on voice quality.

Since the vowels are supposed to be in the same system, the changes of laryngeal outlet area are expected to change the vowel quality thus the formant levels. In the current study, the subjective evaluation of multi listeners' judgments reveals a considerable change in voice quality. Although this outcome is theoretically expected, formantic analysis does not show any change or shift clarifying the situation. This may lead to a consideration that voice quality change can not be a part of vowel resonance structure.

Our results do not aim to project the absolute reason which is obtained by the Kocak's outcomes in terms of vocalization quality; however this study is focused on the speaking quality to show as formants in vowels as a first line of the research. Principally any voice quality change should not have an effect on the vowel formantic structure if the common anatomic sites are not affected. In our study the change of hyoid region and hypopharyngeal air column area do not show any significant deviation in all of the vowel peak formant frequencies. This also means that the hyoid region has a much more importance in perceived voice quality then the phonetic articulation system.

Hyoidoplasty is a technique experimentally has been designed to change voice quality. Unlike other techniques, hyoidoplasty changes resonance, not laryngeal mechanical parameters and articulation. This is an important finding and emphasizes the fact that hyoidoplasty can be an alternative voice surgery technique and should perhaps be better named as 'acoustic voice surgery'.

6.2 Recommendations for Future Work

This preliminary study on the effects on voice quality of the low pharyngeal (hyoid) region can be complemented by a number of future research projects as listed below:

1. During the acquisition of sound data with pressure applied on the hyoid, we had to rely on the subject's efforts, because a commercial applicator was not available. To improve the reliability of the acquisitions, an applicator can be designed. This applicator should be equipped with controls so that the application pressure can be selected at will, it must be flexible so that it can be used on different subjects, and it should not cause any patient discomfort.

2. It is important to develop new objective parameters and criteria that are easily understandable and usable by clinicians for assessing voice quality. Presently, functional evaluation after a vocal surgery is based on several approaches including videolaryngostroboscopy (VLS), for morphological aspects evaluation, GRBAS scale[1], and Voice Handicap Index (VHI). These approaches should be studied and new objective criteria based on acoustic

[1] The perceptive GRBAS scale comprises five qualitative parameters: grade of dysphony (G), roughness (R), breathiness (B), asthenicity (A), and strainess (S). For each parameter, a value in the range 0–3 is considered, where 0 corresponds to healthy voice, 1 to light disease, 2 to moderate and 3 to severe [6], [7]. Notice that G is related to the degree of abnormality of the voice, B to the airflow size through the glottis. S refers to hyper functional phonation, and A indicates low strength in voice. R reaches intermediate values, and represents the psycho-acoustic impression of the irregular vocal folds vibration, related to jitter and shimmer.

parameters, such as fundamental frequency, noise, formants, etc., should be developed.

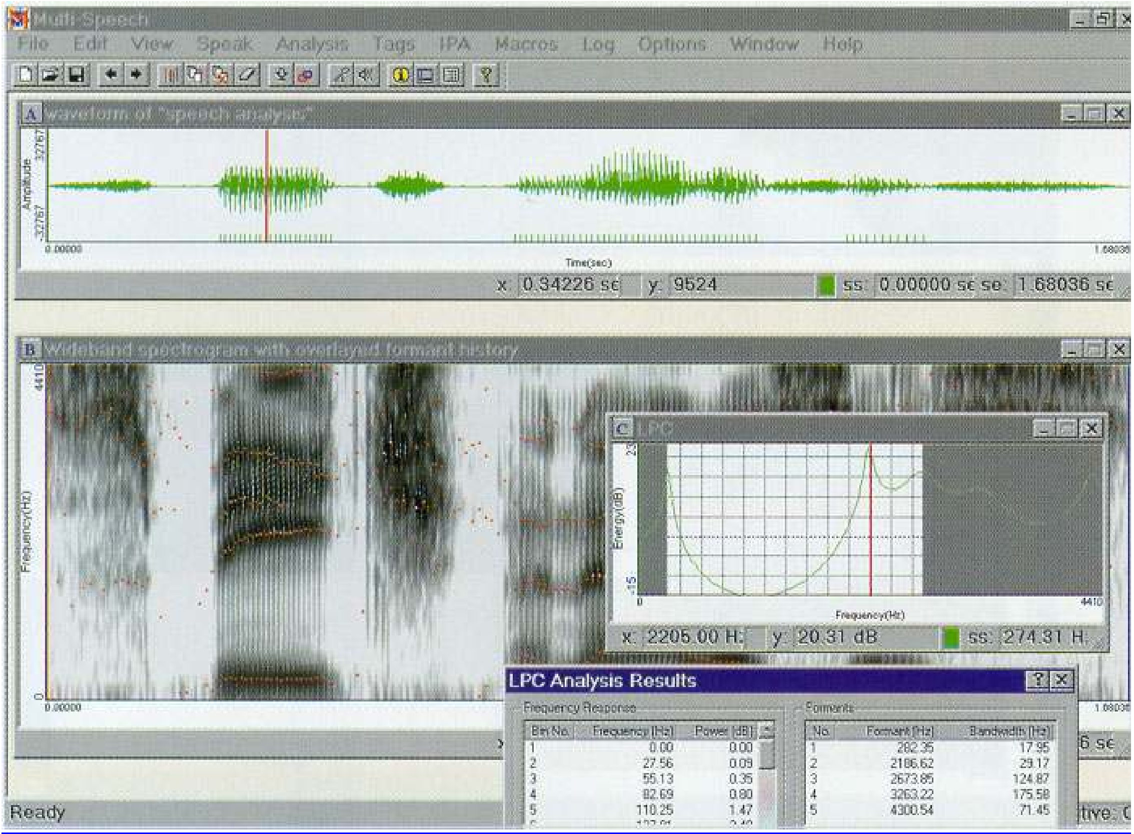
3. A new software tool should be developed that will make it easy to perform presurgical and postsurgical voice quality comparisons.

4. Mathematical models of the phonatory and vocal tract systems should also be re-examined to understand intricate aspects of speech production and the effect of each part of the model on voice quality.

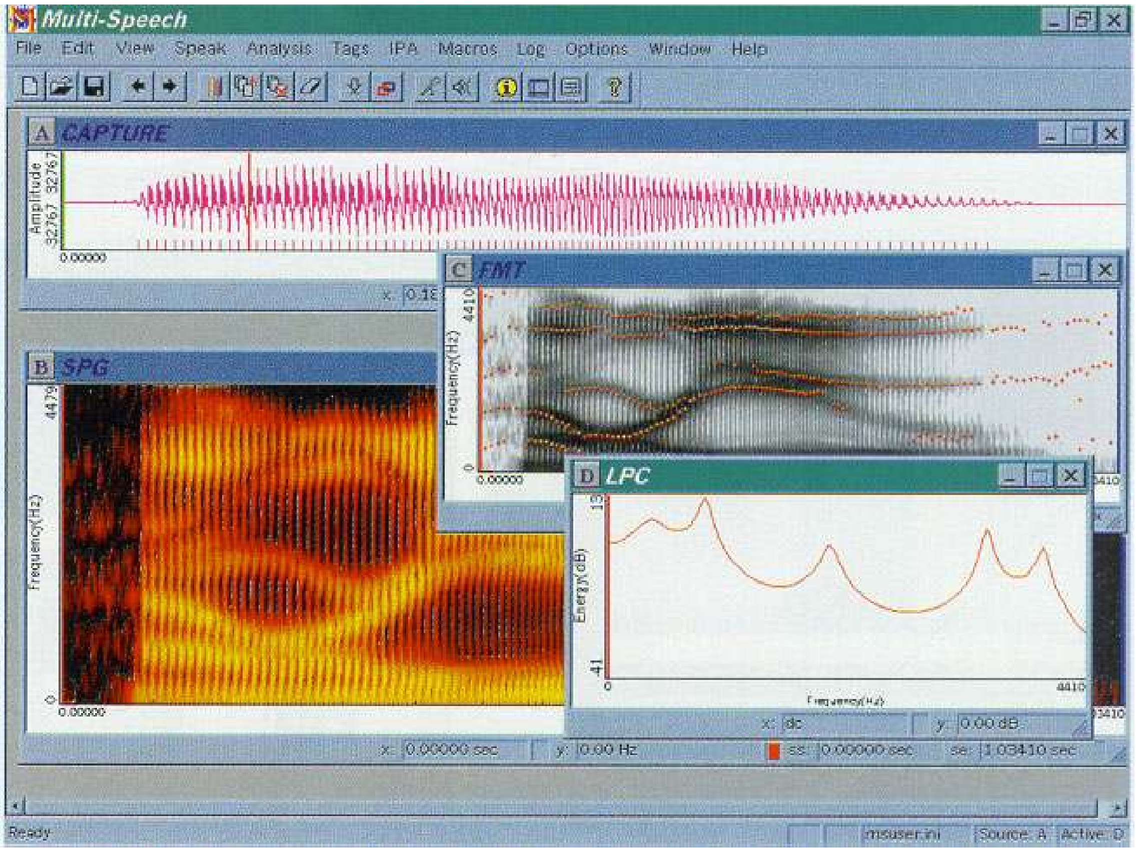
APPENDIX A
MULTI-SPEECH, MODEL 3700 SPEECH ANALYSIS PROGRAM

Multi-Speech, Model 3700, is a, Windows®-based, speech analysis program using standard multimedia hardware to capture, analyze, and play speech samples. It is a comprehensive speech recording, analysis, feedback, and measurement software program. It is the most widely used speech analysis system and also has numerous application specific software options. It is only limited by the specifications, features, and S/N limitations typical of audio devices in the host computer. It can be used for teaching, research, voice measurements, clinical feedback, second language articulation, and forensic work.

The screen below is from a sample macro and signal file delivered with Multi-Speech. Window A shows the IPA symbols time-linked to the speech waveform with glottal impulse markers. Other windows show a spectrogram with formant tracking, a pitch contour, an amplitude contour, and a palatogram, all linked together so that cursor movement shows the different palatometric patterns with the associated acoustic analysis.



Comprehensive graphic and numerical analyses are available for the rich array of easy-to-use Multi-Speech features. The LPC-derived formant frequencies are shown above.



Example of multi-speech screen with waveform (A), colored spectrogram (B), gray scale spectrogram with formant history (C), and LPC formant analysis (D). All graphic views have cursors and numerical analysis. All views are sizable and movable.

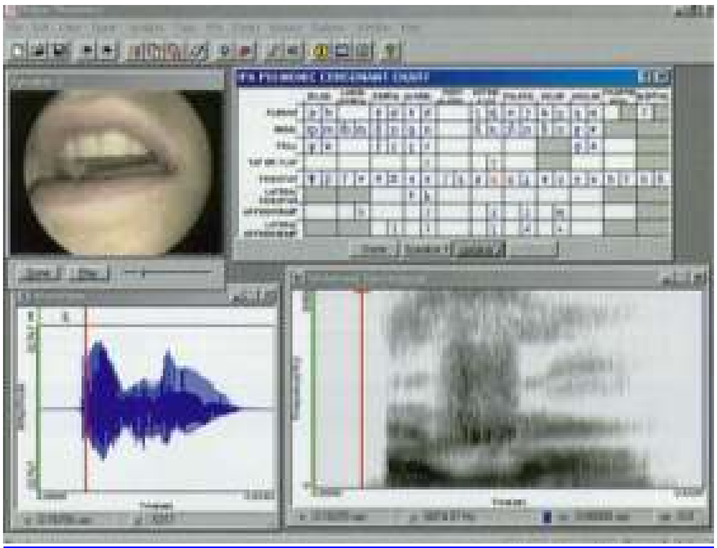
Features

- Capture and play with any sampling rate
- Up to eight channels of data
- Read and write to CSL or .wav file format
- Complete IPA Symbols
- Editing includes cutting, pasting, digital filtering, and appending

- Complete online Help and tutorial

Analysis includes multiple pitch extraction methods, pitch-synchronous LPC, glottal event markers, FFT, formant history, spectrogram, energy, etc.

- Logging of cursor values
- Macros to repeat operational steps
- Learn mode to memorize steps
- Tags, storable with signal
- Complete suite of features for each analysis and editing command
- All 32-bit, MDI-compliant for fast operation



The Video Phonetics Program shown above is one of several optional programs for Multi-Speech for targeted applications. Programs are targeted at clinical, teaching, and research applications.

Summary

Multi-Speech is a powerful, and easy-to-use speech analysis program with many optional programs. These are ideal for research, teaching, and clinical applications. Multi-Speech, because it is based on the CSL (the leading speech analysis system), includes a wide array of functions and features for speech professionals.

Multi-Speech requires a host PC computer (Pentium >266MHz, CD-ROM, 16 MB RAM) and sound device.

REFERENCES

1. Titze, I.R., *Principles of Voice Production*, Prentice-Hall, Englewood Cliffs, 1994.
2. Koçak, İ. , 2002 , "Model of Laryngeal Resonance and its use in improving Voice Quality through Surgery", MS Thesis
3. Snell, R.S., *Clinical Anatomy for Medical Students*,Richard S. Snell, Little Brown, 1995.
4. Gray, H., Anatomy the human body:1918. Avaliable: http://education.yahoo.com/reference/gray/illustrations/figure;_ylt=Al3Wi.Ti..SnlxjwLhEQFuptHokC?id=381
5. INTELLIMED International Corporation:1999.Avaliable: http://www.innerbody.com/image/nerv12.html
6. Propert, M&L, The Propert Encyclopaedia: 1993. Avaliable: http://www.probertencyclopaedia.com/j/Tongue.jpg
7. Language online, ,History of speech recognition: 1987. Avaliable: http://www.languageonline.in.th/images/h3.jpg
8. Scientific library online: 2002.Avaliable: http://www.scielo.cl/fbpe/img/ijmorphol/v24n1/fig04-02.jpg
9. Russell, K., General phonetics: 2003.Avaliable: http://www.umanitoba.ca/linguistics/russell/phonetics/anatomy/soft-palate.html
10. Hart, M., Project Gutenberg: 1971.Avaliable: http://www.gutenberg.org/files/19880/19880-h/images/fig47.jpg
11. Dechant, E., Improving the Teaching of Reading: 1970. Avaliable: http://en.wikipedia.org/wiki/Palatine_uvula
12. Mlynarek, A., Discover the larynx class notes: 2002. Avaliable: http://sprojects.mmi.mcgill.ca/larynx/notes/n_frames.htm
13. Eastroy community school class notes: 2005. Avaliable: http://www.easttroy.k12.wi.us/hs/dept/science/bottum/Adv%20Biology/respiratory/rphysiology/larynx.JPEG
14. BoneHarry M. Tucker, The Larynx, Thieme 1987, p.2, Fig.1.1
15. Netter F, Atlas of Human Anatomy 2nd ed, Novartis 1997, p71
16. Armstrong, E., Voice and speech source: 1995. Avaliable: http://www.yorku.ca/earmstro/journey/images/thyroid.gif
17. Lindberg, D., Medline Plus: 1998. Avaliable: http://www.nlm.nih.gov/medlineplus/ency/images/ency/fullsize/12668.jpg
18. Mcgill medical students, molson medical informatics: 2002. Avaliable: http://sprojects.mmi.mcgill.ca/larynx/notes/hist/nhintro.htm
19. Titze, I.R. "The physics of small-amplitude oscillation of the vocal folds," *J. Acoust. Soc. Am.,* 83 (1988), 1536-1552.
20. Lucero, J.C.,"A theoretical study of the hysteresis phenomenon at vocal fold oscillation onset-offset," *J. Acoust. Soc. Am.,* 105 (1999), 423-431 (1999).
21. Herzel, H. and C. Knudsen, "Bifurcation in a vocal fold model*," Nonlinear Dyn.,* 7 (1995), 53-64.
22. Trevisan, M.A., Eguia, M.C. and G.B. Mindlin, "Nonlinear aspects of analysis and synthesis of speech time series data," *Phys. Rev.* E, 63 (2001), 026216.

23. Jiang, J.J., Zhang, Y. and J. Stern, "Modeling of chaotic vibrations in symmetric vocal folds," *J. Acoust. Soc. Am.,* 110 (2001), 2120-2128.

24. Lucero, J.C., "Dynamics of a two-mass model of the vocal folds for men, women, and children," *Proceedings of the 4th International Conference on Voice Physiology and Biomechanics* (2004), 191-195.

25. Lucero J.C. and L.L. Koenig, "Simulations of VhV sequences in children," *Proceedings of the 15th International Conference on Phonetic Science* (2003), 2905-2908.

26. Flanagan, J. and L. Landgraf, "Self-oscillating source for vocal-tract synthesizers," *IEEE Trans. On Audio and Eletroacoustics,* 16 (1968), 57-64.

27. Berry, D.A., Herzel, H., Titze, I.R. and B.H. Story, "Bifurcations in excised larynx experiments," *J. Voice,* 10 (1996), 129-138.

28. Pelorson, X., Hirschberg, A., van Hassel, R.R., Wijnands, A. P., J. and Y. Auregan, "Theoretical and experimental study of quasisteady- flow separation within the glottis during phonation. Application to a modified two-mass model," *J. Acoust. Soc. Am.,* 96 (1994), 3416-3431.

29. Lucero, J.C., "Dynamics of the vocal fold oscillation", Department of Mathematics, UnB, 70910-900 Brasilia, DF, Brazil.

30. I.R. Titze, The physics of small-amplitude oscillation of the vocal folds, J. Acoust. Soc. Am., 83 (1988), 1536-1552.

31. M.A. Trevisan, M.C. Eguia and G.B. Mindlin, Nonlinear aspects of analysis and synthesis of speech time series data, Phys. Rev. E, 63 (2001), 026216.

32. Perko, L., *Differential Equations and Dynamical Systems*, Springer-Verlag, New York, 1991.

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

34. Thompson, J.M.T. and H.B. Stewart, *Nonlinear Dynamics and Chaos*, Wiley, New York, 1996.

35. Hirose, H. and S. Niimi, "The relationship between glottal opening and the transglottal pressure differences during consonant production," in *Laryngeal Function in Phonation and Respiration* (T. Baer, C. Sasaki, e K. Harris, eds.), pp. 381-390, College-Hill, Boston.

36. Koenig, L.L., "Laryngeal factors in voiceless consonant production in men, women, and 5-year-olds," *J. Speech Lang. Hear. Res*., 43 (2000), 1211-1228.

37. Acoustic Properties of the Vocal Tract Tampere University of Technology SGN-4106 - Speech Recognition

38. PENTAX Medical Company Available:http://www.kayelemetrics.com/Productinfo/3700/3700.htm

39. Turk, O., 2005. Available:http://www.sestek.com.tr/Company_Profile.asp

40. SPSS Inc., 1968 http://www.spss.com/spss/index.htm.source