



**T.C.
İSTANBUL UNIVERSITY
INSTITUTE OF GRADUATE STUDIES
IN SCIENCE AND ENGINEERING**



Ph.D. THESIS

**ANALYSIS OF THE FUNDAMENTAL FREQUENCY AND NASAL
SOUND IN PATIENTS WITH NASAL POLYPS**

Emrah ORUN

Department of Biomedical Engineering

Biomedical Engineering Programme

**SUPERVISOR
Prof. Dr. Aydın AKAN**

January, 2017

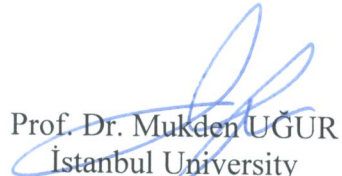
İSTANBUL

This study was accepted on 16/1/2017 as a Ph. D. thesis in Department of Biomedical Engineering, Biomedical Engineering Programme by the following Committee.

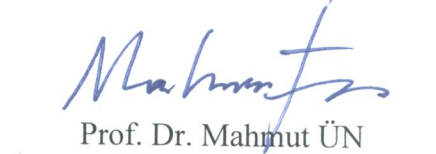
Examining Committee Members



Prof. Dr. Aydın AKAN(Supervisor)
İstanbul University
Faculty of Engineering



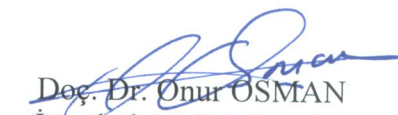
Prof. Dr. Mukden UĞUR
İstanbul University
Faculty of Engineering



Prof. Dr. Mahmut ÜN
İstanbul Yeni Yüzyıl University
Faculty of Engineering and
Architecture



Doç. Dr. Fırat KAÇAR
İstanbul University
Faculty of Engineering



Doç. Dr. Onur ÖSMAN
İstanbul Arel University
Faculty of Engineering and
Architecture



As required by the 9/2 and 22/2 articles of the Graduate Education Regulation which was published in the Official Gazette on 20.04.2016, this graduate thesis is reported as in accordance with criteria determined by the Institute of Graduate Studies in Science and Engineering by using the plagiarism software to which İstanbul University is a subscriber.

This study was supported by TUBITAK with the project number of 112E167.

The author of this thesis was supported under the TUBITAK BİDEB 2211-C scholarship program.

FOREWORD

I had had my BS when I started working as a lecturer in 2008. However, with the responsibility of working in an academic institution, it was necessary to get MS and PhD to improve and learn new information.

I was glad that I got accepted to a PhD program in engineering. But it was annoying to be included in the orientation course with undergraduate students. I would like to express my thanks to Prof. Dr. Sadık KARA, who was always right beside me from the first day I started PhD in 2011. In his own words “do not disturb the comfort of the head”, and I continued my way by doing so.

I thank Asst. Prof. Dr. Şükrü OKKESİM, Asst. Prof. Dr. Haşim Özgür TABAKOĞLU, Asst. Prof. Dr. Saime AKDEMİR AKAR, who, during one and a half years due to the intensive course program, relieved my tiring journeys between Ankara and Istanbul with a smiling face and understanding attitude.

I thank Dr. Ahmet AKKOZ, Asst. Prof. Dr. Alper YÜKSEL, Assoc. Dr. Nebil ARK, who at the hospital provided full support to get records and communicate with patients.

I thank my colleagues Asst. Prof. Dr. Murat YEŞİLYURT, lecturer Nuh Naci ATAY, who were always ready to assist in writing my thesis and technical matters.

I express my gratitude to Prof. Dr. Aydın AKAN, who honored me by accepting to be my supervisor after being transferred to Istanbul University and supported me continually until I completed my thesis; and Asst. Prof. Dr. Aysel ERSOY YILMAZ, who while receiving my registration information made me connect with the university with her positive energy.

And of course, I would like to thank my wife and other family members who have always supported me in this difficult process of five and a half years.

January 2017

Emrah ORUN

TABLE OF CONTENTS

	Page
FOREWORD.....	i
TABLE OF CONTENTS.....	ii
LIST OF FIGURES	v
LIST OF TABLES	vii
LIST OF SYMBOLS AND ABBREVIATIONS	viii
ÖZET.....	x
SUMMARY	xii
1. INTRODUCTION.....	1
1.1. LITERATURE REVIEW	1
1.2. PURPOSE OF THE THESIS	2
1.3. HYPOTHESIS.....	3
1.4. NASAL POLYPS	4
1.4.1. Epidemiology	4
1.4.2. Etiology	5
1.4.3. Diagnosis.....	5
1.4.4. Treatment	7
1.4.4.1. Medical Treatment.....	7
1.4.4.2. Surgical Treatment	8
1.5. THE NATURE AND FORMATION OF SOUND/SPEECH.....	9
1.5.1. Characteristics of Sound	10
1.5.1.1. Pitch.....	10
1.5.1.2. Intensity/Loudness	11
1.5.1.3. Tone or Quality.....	11
1.5.2. Anatomic Structures That Form Sound/Speech.....	11
1.5.2.1. Lungs.....	11
1.5.2.2. Trachea	12
1.5.2.3. Larynx (Voice Box).....	12
1.5.2.4. Vocal Cords (Vocal Folds)	13

1.5.2.5.	<i>Pharynx</i>	14
1.5.3.	Organs Forming Oral Cavity.....	14
1.5.3.1.	<i>Lips</i>	15
1.5.3.2.	<i>Teeth</i>	15
1.5.3.3.	<i>Alveolar Ridge</i>	15
1.5.3.4.	<i>Palate</i>	15
1.5.3.5.	<i>Tongue</i>	15
1.5.4.	Formation of Speech and Modelling.....	16
1.5.4.1.	<i>Formation Stages of Speech</i>	16
1.5.4.2.	<i>Source-Filter Model of Speech Production</i>	20
2.	MATERIALS AND METHODS	21
2.1.	CLINIC AND DATA ACQUISITION	21
2.2.	INSTRUMENTATION AND MEASUREMENT	23
2.2.1.	Turkish Alphabet Acoustics.....	23
2.2.2.	Electroglottograph.....	23
2.2.2.1.	<i>Function</i>	24
2.2.2.2.	<i>EKG waveform</i>	24
2.2.2.3.	<i>EKG Device</i>	26
2.2.3.	Instrumentation and Measurement.....	28
2.3.	SIGNAL PROCESSING AND CLASSIFICATION	29
2.3.1.	Classification Models.....	29
2.3.1.1.	<i>Artificial Neural Network (ANN) / Multi-Layer Perceptron (MLP) Model</i>	29
2.3.1.2.	<i>Spegasos Model</i>	31
2.3.1.3.	<i>Naive Bayes Model</i>	31
2.3.1.4.	<i>Logistic Model</i>	32
2.3.1.5.	<i>RBF Network Model</i>	32
2.3.1.6.	<i>Simple Logistic</i>	33
2.3.1.7.	<i>SMO Model</i>	33
2.3.1.8.	<i>J48 Model</i>	33
2.3.2.	Training Methods	34
2.3.3.	WEKA Application.....	34
3.	RESULTS	36
3.1.	EXPERIMENTAL RESULTS	36
3.2.	FUNDAMENTAL FREQUENCY AND LOUDNESS ANALYSIS	39

3.3. NASAL REGION ANALYSIS (SPE AND NX/SP RATIO GRAPH).....	47
4. DISCUSSION.....	59
5. CONCLUSION AND RECOMMENDATIONS	64
REFERENCES.....	65
APPENDICES	71
APPENDIX A.....	71
APPENDIX B.....	74
CURRICULUM VITAE.....	77



LIST OF FIGURES

	Page
Figure 1.1: Anterior rhinoscopy.....	6
Figure 1.2: Use of nasal endoscope.	6
Figure 1.3: Sound transmission and perception.	9
Figure 1.4: Airflow through the lungs.	12
Figure 1.5: Anatomy of the Larynx.	12
Figure 1.6: Parts of Larynx.	13
Figure 1.7: The Position of Vocal Cords during Respiration.	14
Figure 1.8: The Position of Vocal Cords during Normal Tone of Sound.....	14
Figure 1.9: Parts of the Tongue.....	15
Figure 1.10: Respiration.....	17
Figure 1.11: Phonation.....	17
Figure 1.12: Resonance.....	18
Figure 1.13: Articulation.....	19
Figure 1.14: Source-filter model of speech production.	20
Figure 2.1: Audiology Hearing Balance Lab. at Turgut Ozal University.....	22
Figure 2.2: The process steps of proposed work.....	22
Figure 2.3: Raw waveform of EGG.	25
Figure 2.4: Phases of the EGG signal.	26
Figure 2.5: During of speech position vocal cord.	26
Figure 2.6: EGG coupling/application.	27
Figure 2.7: The early EGG devices (EGG-A100 model).....	27
Figure 2.8: A measuring with microProcessor Laryngograph EGG-D400 device.	28

Figure 2.9: A simple three layer feed forward neural network.	30
Figure 3.1: Selection of the microphone calibration.	36
Figure 3.2: Selection of the device parameters.	37
Figure 3.3: The spectrum of recorded speech.	37
Figure 3.4: The QA Settings and Quantitative Analysis of recorded speech.	38
Figure 3.5: Selecting the area to be calculated with MDVP.	39
Figure 3.6: MDVP Results (Sample).	40
Figure 3.7: Loudness graph.	41
Figure 3.8: ROC Curve, RBF Network and Simple Logistic Algorithms, Leave- One-Out Cross-Validation (LOOCV) (Class value= with NP).	46
Figure 3.9: ROC Curve, RBF Network and Simple Logistic Algorithms, Leave- One-Out Cross-Validation (LOOCV) (Class value= without NP).	47
Figure 3.10: The graphs obtained after QA.	48
Figure 3.11: Speech Pattern Elements, Individual with Nasal Polyps.	49
Figure 3.12: Speech Pattern Elements, Individual without Nasal Polyps.	49
Figure 3.13: Nx/Sp Ratio, Individual with Nasal Polyps.	50
Figure 3.14: Nx/Sp Ratio, Individual without Nasal Polyps.	50
Figure 3.15: ROC Curve, Naive Bayes and N-fold cross validation	57
Figure 3.16: ROC Curve, Naive Bayes and N-fold cross validation	58
Figure 4.1: The formation of speech as a complete system.	61

LIST OF TABLES

	Page
Table 1.1: Frequency values of human voice.....	10
Table 2.1: Turkish Alphabet Acoustics.....	23
Table 2.2: A confusion matrix for classification.....	34
Table 3.1: Min-max values for measured fundamental frequency (F0).....	41
Table 3.2: Min-max values for Ax graph mean value.....	41
Table 3.3: The success of classification model for N-fold cross validation (a) and Leave-One-Out Cross-Validation (LOOCV) (b).	43
Table 3.4: The classification model results in detail.....	45
Table 3.5: Statistical values of the Individuals with NP.	51
Table 3.6: Statistical values of the Individuals without NP.	52
Table 3.7: The success of classification model for N-fold cross validation (a) and Leave-One-Out Cross-Validation (LOOCV) (b).	54
Table 3.8: The classification model results in detail.....	56

LIST OF SYMBOLS AND ABBREVIATIONS

Symbol	Explanation
--------	-------------

β	: Beta
---------	--------

Abbreviation	Explanation
--------------	-------------

AB	: Ada Boost
-----------	-------------

ANN	: Artificial Neural Network
------------	-----------------------------

ARFF	: Attribute-Relation File Format
-------------	----------------------------------

BC	: Before Christ
-----------	-----------------

CT	: Computed Tomography
-----------	-----------------------

DM	: Data Mining
-----------	---------------

dB	: Decibel
-----------	-----------

ECG	: Electrocardiography
------------	-----------------------

EEG	: Electroencephalography
------------	--------------------------

EGG	: Electroglottography
------------	-----------------------

EMG	: Electromyography
------------	--------------------

ENT	: Ear, Nose, and Throat
------------	-------------------------

ESS	: Endoscopic Sinus Surgery
------------	----------------------------

FN	: False Negative
-----------	------------------

FP	: False Positive
-----------	------------------

Hz	: Hertz
-----------	---------

LL	: Lib Linear
-----------	--------------

LM	: Levenberg-Marquardt
-----------	-----------------------

LOOCV	: Leave-One-Out Cross-Validation
--------------	----------------------------------

MDVP	: Multidimensional Voice Profile
-------------	----------------------------------

MLP	: Multilayer Perceptron
------------	-------------------------

NB	: Naive Bayes
-----------	---------------

NP	: Nasal Polyp
-----------	---------------

Pegasos	: Primal Estimated sub-GrAdient SOLver for SVM
----------------	--

QA	: Quantitative Analysis
-----------	-------------------------

RBF	: Radial Basis Function
RF	: Random Forest
ROC	: Receiver Operating Characteristic
SMO	: Sequential Minimal Optimization
SPE	: Speech Pattern Element
SPL	: Sound Pressure Level
SVM	: Support Vector Machine
TN	: True Negative
TP	: True Positive
VFCA	: Vocal Fold Contact Area
WEKA	: Waikato Environment for Knowledge Analysis

ÖZET

DOKTORA TEZİ

NAZAL POLİPLİ HASTALARDA TEMEL FREKANS ve NAZAL SES ANALİZLERİ

Emrah ORUN

İstanbul Üniversitesi

Fen Bilimleri Enstitüsü

Biyomedikal Mühendisliği Anabilim Dalı

Danışman : Prof. Dr. Aydın AKAN

Bu çalışmanın temel amacı, Nazal Polip (NP) hastalarının ses kayıtlarını elde ederek bu kayıtlar üzerinden sesin oluşumundaki değişimleri analiz etmektir. Ses akciğerlerden gelen havanın, ses tellerinin de bulunduğu larinks bölgesinden (kaynak bölge) geçerek ağız, burun boşlukları ve paranasal sinüslerden oluşan rezonans bölgesinde filtre edilmesi ile oluşur. Bu çalışmada kaynak-filtre modeli olarak ifade edilen bu modelde filtre bölgesinde değişime sebep olan nazal poliplerin hem rezonans bölgesindeki hem de kaynak bölgesindeki etkileri ortaya konmak istenmiştir. Bu amaçla elde edilen ses kayıtlarından öznitelikler çıkarılmış (feature extraction) ve öznitelik seçimi yapılmıştır. Seçilen öznitelikler çeşitli algoritmalar ile (MLP, Spegasos, NB, Logistic, RBF Network, Simple Logistic, SMO, J48) , Leave-One-Out Cross-Validation (LOOCV) ve N Fold Cross-Validation eğitim metodları kullanılarak sınıflandırılmıştır.

Bu çalışmada gönüllü olanların mikrofon ile normal sesleri kaydedilmiştir. Buna ilave olarak, larinks bölgesinde ses tellerinin yer aldığı glottis kısmından elektroglottografi yardımı ile EGG sinyalleri ve burun üzerine bağlanan titreşim transdüseri ile konuşma anındaki nazal titreşimler kaydedilmiştir. Çalışmada gönüllülere nazal sesler (m, n) içeren kelimeler söylenmiştir. Her kelime için temel frekanslar (F0) hesaplanmış,

kaydın tamamı için sesin grlė (loudness) ve burun yzeyine baėlanan titreřim transdseri ile kayıt boyunca nazal titreřimler yzde olarak tespit edilmiřtir.

Sonuç olarak elde edilen verilerle 2 farklı analiz yapılmıřtır. İlk olarak temel frekans ve grlk deėerleri kullanılarak nazal poliplerin ses oluřumunda kaynak blgeyi ve dolayısıyla temel frekansı etkileyip etkilemediėi ortaya konmuřtur. İkinci olarak ta elde edilen nazal titreřim verilerinden konuřmanın tamamına oranlama yapılarak, nazal poliplerin konuřma zerine etkileri ortaya konmuřtur.

Ocak 2017, 91 sayfa.

Anahtar kelimeler: Nazal Polip, Temel Frekans, Sınıflandırma, Çapraz Doğrulama



SUMMARY

Ph.D. THESIS

ANALYSIS of THE FUNDAMENTAL FREQUENCY and NASAL SOUND in PATIENTS WITH NASAL POLYPS

Emrah ORUN

İstanbul University

Institute of Graduate Studies in Science and Engineering

Department of Biomedical Engineering

Supervisor : Prof. Dr. Aydın AKAN

The main purpose of this study is to obtain the voice recordings of patients with the Nasal Polyp (NP) and to analyze the changes in voice formation through these recordings. Sound occurs when the air from the lungs passes through the larynx region (source region) where the vocal cords are located and is filtered in the resonance region of the mouth, nasal cavities, and paranasal sinuses. In this study, in this model, which is also called as the source-filter model, it is aimed to reveal the effects of the nasal polyps which causes the change in the filter region both on resonance region and on source region. For this purpose, features were extracted from the obtained sound recordings and feature selection was made. Selected features were classified by using various algorithms (MLP, Spegasos, NB, Logistic, RBF Network, Simple Logistic, SMO, J48) and by using Leave-One-Out Cross-Validation (LOOCV) and N Fold Cross-Validation training methods.

In this study, normal sounds of the volunteers were recorded with a microphone. In addition, EGG signals from the glottis part of the larynx region, where the vocal tracts are located, were recorded with an Electroglottograph. The nasal vibrations at the time of speech were recorded with a vibration transducer attached on the nose. In the study,

volunteers were told to say words containing nasal sounds (m, n). Fundamental frequency (F0) for each word were calculated. Loudness for the whole recording and nasal vibrations with a vibration transducer attached on the nose were determined as percentage.

As a result, two different analyses were made using the obtained data. First, it was revealed if nasal polyps affect source region and thus fundamental frequency in sound formation by using fundamental frequency and loudness. Second, the effects of nasal polyps on speech were revealed assessing the whole speech from the obtained nasal vibration data.

January 2017, 91 pages.

Keywords: Nasal Polyp, Fundamental Frequency, Classification, Cross Validation



1. INTRODUCTION

1.1. LITERATURE REVIEW

Physically, sound is mechanical vibrations that occur in gaseous, liquid or solid environments. In order for these vibrations to be perceived as sound, they must have a propagation environment, ear sensitivity (sense ability) and energy. A movement in any part of the material environment causes the other parts to start moving because of the flexibility of the material. This movement allows the sound to be heard. As the air cannot propagate in an airless environment, the sound can not be heard [1].

Edels states that there are mainly four events in the formation of human voice. They are initiation, vibration resonance and articulation respectively. The formation of raw sound starts in the lungs. Vibration takes place in the larynx in vocal folds and raw sound occurs. Pharynx, mouth, nasal cavities and paranasal sinuses constitute the resonance regions and they play a role in amplifying the sound. Tongue, teeth, lips and palate make raw sound whose articulation organs have been amplified an understandable phoneme [2]. Today, nasal polyps are known to have affected speech due to the changes which they caused in the resonance region [3, 4]. The medical effects of nasal polyps have been the focus of many studies [5]. In a study conducted on children aged 4-12, the effect of barriers in the nasal region on the sound quality has been examined and nasal polyps as one of these barriers have been discussed. As a result, it has been revealed that the barriers in the nasal region affect the sound quality [6]. In another study, nasometer measures and audio spectrum of patients with nasal polyps have been studied [7]. In addition, there are studies examining the effects of changes in nasal resonance on sound [8, 9]. To sum up, the general view we encountered in the literature review is that there are effects of nasal polyps on human life or the effects of changes in nasal resonance on sound.

On the other hand, many studies have been done relating to fundamental frequency measurements which we examine in our study. Today, voice processing is used in many fields. In particular, voice recognition, the use for security purposes using voice

characteristics of people, making various classifications with the fundamental frequency are most used methods [10-13].

As the fundamental frequency includes the features of the person's speech, it's been used in many academic studies. In one study, fundamental frequency of men and woman has been compared and analyzed. In one study, the voice sensitivity of 18-45 males and females have been examined by the basic frequency data [14]. In another study, fundamental frequency has been used to analyze and the classify mutational dysphonia disease [15]. Also for the gender perception and male-female voice comparison, fundamental frequency has been used in some studies [16, 17].

As a result of the literature review, it has not been encountered any method that contributes to the diagnosis of nasal polyp disease using a similar approach to the one used in this thesis.

1.2. PURPOSE OF THE THESIS

When we look at the field of medicine, many traditional methods have been replaced with methods called modern medicine. In modern medicine, the contribution of many sub-fields of science in Biomedical Engineering is obvious [18, 19]. Despite the use of biomedical devices in the diagnosis and treatment of diseases, the ratio of total benefit/total loss a patient will get is taken into account and even if the used device causes damage, it is preferred because of the total benefit ratio. For example, X-ray machines used extensively in many branches of medicine expose patient directly to ionized radiation. But being exposed to radiation from the x-ray machines is considered reasonable [18, 19]. They are also preferred by doctors because of the ease of diagnosis.

On the other hand, today the study of physiological parameters obtained from human by transforming into different forms and the presentation of them to the medical world has gained an important place in Biomedical Engineering. Thanks to especially biomedical signal processing, successful results have been obtained in diagnosis of many diseases [20].

When we look at the current methods used in the diagnosis of nasal polyps, we see medical history, physical examination, nasal endoscopy and CT [3]. However, it would

be appropriate to use these methods when they are really needed since they are demanding for patients and they expose patients to radiation. Considering the effects of nasal polyps on the speech of individuals, the diagnosis of the disease before advancing further by analyzing the sound recordings of individuals would support the existing methods in this regard and doctors, before adopting other methods, will have employed a preliminary diagnosis method. In this respect, the thesis topic will help protect patients from the unnecessary effects of endoscopy and radiation.

The main purpose of this study is to obtain the voice recordings of patients with the Nasal Polyp (NP) and to analyze the changes in voice formation through these records. Sound occurs when the air from the lungs passes through the larynx region (source region) where the vocal cords are located and filtered in the resonant region of the mouth, nasal cavities, and paranasal sinuses. In this study which is also referred to as the source-filter model, it is aimed to reveal the effects of the nasal polyps which causes the change in the filter region. Sound recordings obtained for this purpose have been classified by using mathematical properties with classification algorithms and the results have been interpreted.

In addition, it is aimed that the results obtained within the scope of this study will support the preliminary diagnosis of nasal polyps.

1.3. HYPOTHESIS

Nasal polyps generally can be diagnosed with examination and nasal endoscopy. A detailed computed tomography is needed to determine the extent of polyps in sinuses and plan a surgical treatment if it will be used [3]. Nasal polyps cause changes in sound because they affect nasal cavities and paranasal sinuses which play a major role in sound resonance [3, 4].

In this study, normal sounds of the volunteers will be recorded with a microphone. In addition, EGG signals from the glottis part of the larynx region, where the vocal tracts are located, will be recorded with an Electroglottograph. The nasal vibrations at the time of speech will be recorded with a vibration transducer attached on the nose. In the study, volunteers will be told to say words containing nasal sounds (m, n). Fundamental

frequency (F0) for each word will be calculated. Loudness for the whole recording and nasal vibrations with a vibration transducer attached on the nose will be determined as percentage.

As a result, two different analyses will be made using the obtained data. First, it will be revealed if nasal polyps affect source region and thus fundamental frequency in sound formation by using fundamental frequency and loudness. Second, the effects of nasal polyps on speech will be revealed assessing the whole speech from the obtained nasal vibration data.

The basic hypothesis about thesis is that the findings from sound recordings will be meaningful and achieve higher success in classification.

1.4. NASAL POLYPS

Nasal polyps, formed by edematous swelling of the nasal mucosa, are hanging masses that grow inside the nasal passages and sinuses. Nasal polyps, which are the most common benign masses in otolaryngology practice, are derived from ancient Greek words “poly” and “pous” meaning “much” and “foot” respectively [3].

Nasal polyps is a chronic inflammatory disease of paranasal and nasal sinus mucosa whose etiology is not fully clear. Infections, allergies, vasomotor disorders, autonomic dysfunction and endocrine disorders are considered in the etiology. They are characterized by symptoms such as nasal congestion, deterioration in smell and taste, headache, runny nose, postnasal drip which are seen more in diseases like rhino sinusitis and allergies and disrupts the quality of life of patients [21].

Diagnostic tests include medical history, physical examination (anterior rhinoscopy), endoscopy and radiological CT paranasal [21].

1.4.1. Epidemiology

Different studies have been conducted about the prevalence of nasal polyps in society. According to a study carried out in 2000, the prevalence of nasal polyps is around 1-4% [22]. According to another study conducted in 2003, the prevalence is between 0.2 and

4.3% [23]. The majority of studies has shown that nasal polyps are more common in men [23]. According to the same study, the prevalence increases with age.

1.4.2. Etiology

Although there have been many studies on nasal polyps, the etiology has not been fully explained [24].

Many factors have been proposed regarding the etiology of nasal polyps. Some of them are internal and some are external factors. Probably none of these factors (at least in some cases) alone does not lead to the formation of polyps. Various combinations of these factors most likely work together and coalescence of internal and external factors leads to this complexity [24]. Today, etiologic factors considered important in the development of nasal polyps are as follows [24]:

- 1) Allergy
- 2) Super antigens
- 3) Infections
 - a) Viral infections
 - b) Bacterial infections
 - c) Fungal infections
- 4) Related diseases
 - a) Samter's triad
 - b) Churg-Strauss syndrome
 - c) Cystic fibrosis
 - d) Young Syndrome
 - e) Primary ciliary dyskinesia
 - f) Non-allergic Rhinitis with Eosinophilia Syndrome (NARES)

1.4.3. Diagnosis

In patients, it is not difficult to diagnose polyps with severe nasal congestion. These can be noticed by anterior rhinoscopy. (Figure 1.1). To see small polyps confined to the middle meatus, it may require the use of nasal endoscope. (Figure 1.2). However it is necessary to get biopsy to exclude inverted papilloma and malignancy in unilateral cases [24].



Figure 1.1: Anterior rhinoscopy.

<http://krauthammerlab.med.yale.edu/imagefinder/> (18.02.2016)

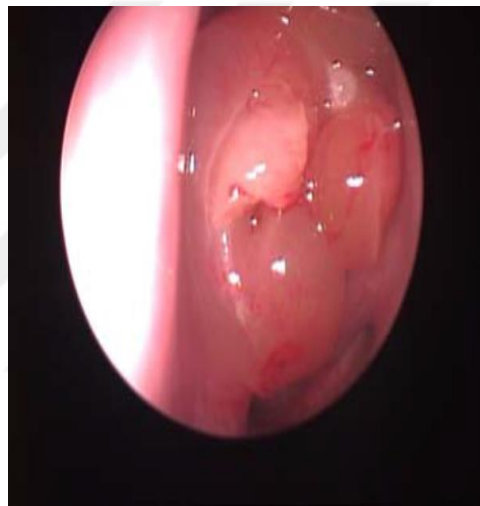


Figure 1.2: Use of nasal endoscope.

<http://krauthammerlab.med.yale.edu/imagefinder/> (18.02.2016)

Paranasal sinus CT scan is the currently preferred method in the evaluation of paranasal sinuses and the adjacent structures. The optimal imaging of bone tissues, soft tissues and optimal air enable correct evaluation of anatomy, paranasal sinuses and the prevalence of the disease in peripheral tissues [24].

As the MRI of paranasal sinus is expensive, it doesn't have primary place in assessment. MRI has soft tissue resolution which is much better than CT. Especially it outclasses CT in bacterial and fungal sinusitis distinction and in inflammation and tumor distinction [24].

We can classify nasal polyps into 4 groups according to their size [25]:

1. Polyps that can be seen with endoscopic medical examination in ostiometal complex
2. Polyps that fill out ostiometal complex
3. Polyps that fill out ostiometal complex, nasal cavity, maksils and frontal sinus
4. Polyps that fill out nasal cavity and paranasal sinus

In recent years, the measurements of diagnostic studies of nasal polyps in the nasal patency (nose opening measurement) and nasal airflow measurement also play an important role [3].

Methods used in the measurement of the Nasal Patency [3]:

1. CT
2. MRI
3. Rinostereometr
4. Acoustic Rhinometry
5. Volumetry Acoustic Rhinometry
6. Nasal Peak Flow (peakflow)
7. Other Airflow methods

1.4.4. Treatment

Treating Nasal Polyps is difficult and today it's still problematic. In 1950, medical treatment became important by using steroids in treatment [26].

1.4.4.1. Medical Treatment

The main purpose of medical treatment in nasal polyp is reduction of polyps or elimination, fixing rhinitis symptoms and avoiding recurrences.

Steroids are powerful anti-inflammatory drug which have been proven scientifically in treatment of nasal polyposis [26].

Systematic steroids are significantly effective in the treatment of nasal polyposis. In the studies nasal polyposis was significantly shrank after systematic steroids applications [26-29]. Sense of smell disorders also give positive reaction to systematic steroid treatment. However, there are a lot of side effects and long-term usage is the disadvantage of systemic steroids. In a study done after 3 weeks of oral steroid therapy, in the early period 72 % improvement was observed in paranasal CT findings of

patients with nasal polyposis but in the long run it was determined that patients returned back to their initial state [28].

The local effective steroids which are not as effective as systematic steroids are the first drugs to be used in the treatment of nasal polyposis because of their few side effects and long term use. Intranasal steroids reduce the size of polyps, relieve nasal flow and reduces polyp recurrence after surgery [29]. After an average of 15 daily oral steroid therapy, in order to achieve further improvement, long-term treatment of topical steroid is a method that can be used as maintenance [30]. If allergic rhinitis accompanies the patients with nasal polyps, antihistamine can be used.

In recent years, Lokotrien antagonist which is used for the treatment of asthma has been suggested to contribute to the treatment of nasal polyps. In a study conducted, it has been reported that it can be used as an alternative to steroids in reducing chronic sinonasal symptom in the of control nasal polyps [31].

In the medical treatment of nasal polyps intranasal decongestants, intranasal furosemide, intranasal capsaicin, antibiotics and antifungal agents may be used more rarely [26].

1.4.4.2. Surgical Treatment

When drug treatment fails, surgical methods should be used. However any surgical treatment could not be curative and patients with nasal polyposis is exposed to medical and surgical treatments several times along their life. Successful results have been obtained with the emergence of Endoscopic Sinus Surgery [32]. The main purpose of Nasal Polyp surgery is to annihilate sinonasal pathology, to cure nasal obstruction, to drain sinuses, to cure disorder of sense of smell, to decrease attack frequency of accompanying illnesses like asthma and to reduce the need of medical treatment [33].

In the surgical treatment of nasal polyps, a variety of surgical techniques are used. Surgical initiatives that are defined at the time of Hippocrates began to be applied in the 1800s, Caldwell-Luc operation and fronto-ethmoidectomy surgery that are started to be implemented in 1900s and ESS surgery that is practiced today are some of them [21].

1.5. THE NATURE AND FORMATION OF SOUND/SPEECH

Sound is the periodic change of pressure in the atmosphere perceived by our ear. In physical dimension, sound is a simple mechanical disturbance formed in air, liquid or gaseous medium. It is the result of vibration of molecules in a substance. Everything that vibrates is a source of sound. Sounds propagate in vibrations with energy from the source. Vibrating objects are flexible and create sound. Flexible objects can create sound waves and propagate sound waves. As sound is a mechanical wave, it needs a medium to propagate. Sound waves are longitudinal waves travelling through media in the form of compression and expansion. Sound waves have pressure and create beat. Sound has frequency, wavelength, period, and speed. The frequency of a wave is the number of cycles in a second, and is measured in Hertz (Hz) [34].

Sound is a physical phenomenon that directs the stimulating effect to the brain through ear. For sound to exist there must be a source to create this impact, a medium through which stimulating effect comes to ear and additionally an ear and a brain to detect this impact. In the absence of one of these there is no sound. Sound only exists when these three elements are present [34]:

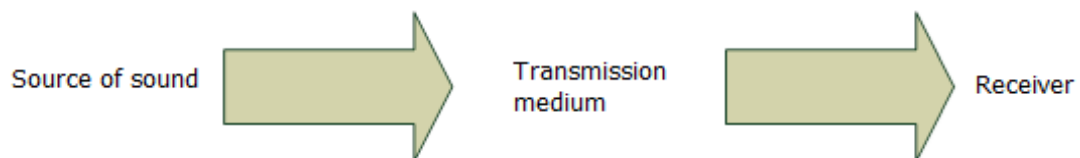


Figure 1.3: Sound transmission and perception.

Sound can be studied as pure and complex by its frequency; periodic and aperiodic by its repeating pattern. Pure tone has only one frequency. It has a sinusoidal waveform. It is periodic by its repeating pattern. A complex sound consists of a combination of a large number of pure tone. Most of the sounds in nature, particularly speech, are in the form of complex sound. Frequency components of complex tone are called "partials". Among the partials, frequency with the smallest frequency value is called F₀ (fundamental frequency). An integer multiple of fundamental frequency is called "harmonics". Fundamental frequency is the first harmonic. For example, if the fundamental frequency is 120 Hz, harmonics are 240, 360, 480, 600 Hz [35].

Sound can be classified into four categories by its frequency value [36]:

- 1) Infrasound : frequency range between 0-20 Hz,
- 2) Audible Sound : frequency range between 20-20.000 Hz,
- 3) Ultrasound : frequency range between 20.000 Hz and 1GHz,
- 4) Hypersound : above 1 GHz.

Humans normally hear sound frequencies between approximately 20 Hz and 20,000 Hz [37]. Sound with high frequency is called treble and sound with low frequency is called bass. Frequencies are used to determine the highness of sound. To calculate this, the frequency of sound in a second is found. As frequencies are higher, sound becomes more high-pitched and as they are lower, sound becomes more low-pitched. With a few exceptions, frequency of female voice is higher than that of male voice. In this case, it appears that female voice is more high-pitched [38].

Table 1.1: Frequency values of human voice.

Voice Type	Fundamental Frequency (Hz)	
	Minimum	Maximum
Bass	87	340
Baritone	90	400
Tenor	125	500
Alto	130	700
Soprano	225	1100
Female Voice	150	
Male Voice	100	

1.5.1. Characteristics of Sound

A sound can be characterized by the following three quantities [34].

1.5.1.1. Pitch

Pitch, the frequency of sound as perceived by human ear, is a distinguishing feature. It is one of the most basic characteristics of sound [34].

The pitch or fundamental frequency of sound is related to the thickness and length of vocal cords. Fundamental frequency is directly related to the fact that how many times vocal cords open and close in a second. Vibration rate is related to vibrator's thickness, length and flexibility. Short, thick and flexible vocal cord (produces high frequency) vibrates at a slower rate than a long, thin vocal cord (produces low frequency) [34].

1.5.1.2. *Intensity/Loudness*

Intensity or amplitude refers to the degree that the balance in the air is distorted by sound source and the compression intensity of molecules. The distance between the pre-vibration balance state of an object which starts to vibrate and the highest peak that vibration waves reach shows intensity. At this point, intensity should not be confused with loudness of sound. Loudness of sound is specific to the ear, so the ear perceives intensity as loudness of sound. For example, when the intensity of sound is doubled, the ear does not perceive this increase in equal degree. Doubling the number of people in a hall does not necessarily mean doubling the loudness of sound. The loudness of sound is primarily related to the sound pressure level (SPL). Generally, sounds with greater SPL are perceived more loudly while sounds with lesser SPL are perceived more softly [34, 39, 40].

Sound intensity is the energy per second per unit of area (watt/ cm²) arriving at that point to the propagation direction and its unit is decibel (dB). Intensity of sound is proportional to the amplitude of vocal vibration. It is dependent on cavity resonators, muscle, ligament, mucosa, lung capacity, strength and tension of thoracoabdominal muscles and the pressure of expiration air [41].

1.5.1.3. *Tone or Quality*

It is the psychoacoustic perception of subjective sound frequency [41]. Tone is also defined as sound quality [34]. Normal tone range for people is (1-1,5) octave. Factors like gender, hearing loss and age can affect tone [41].

1.5.2. Anatomic Structures That Form Sound/Speech

1.5.2.1. *Lungs*

Lungs provide the energy required for the formation of sound. While speaking, vocal cords vibrate through the air from the lungs. Therefore, presenters and singers who receive voice training focus on the control of mechanism of breathing [34, 41].

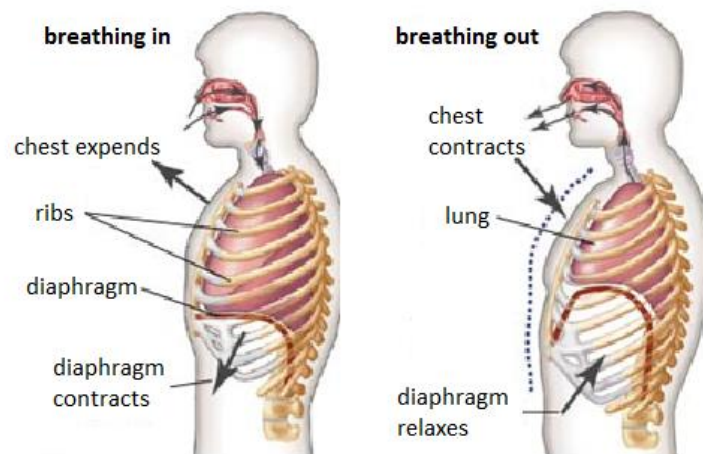


Figure 1.4: Airflow through the lungs.

<http://www.voicedoctorla.com/vocal-anatomy> (03.03.2016)

1.5.2.2. *Trachea*

The trachea is a tube that connects the pharynx and larynx to the lungs, allowing the passage of air.

1.5.2.3. *Larynx (Voice Box)*

Various components are involved in the production of sound in the larynx. These components are tension of vocal cords with air pressure, the shape of rima glottidis, the condition and width of respiratory tract [34].

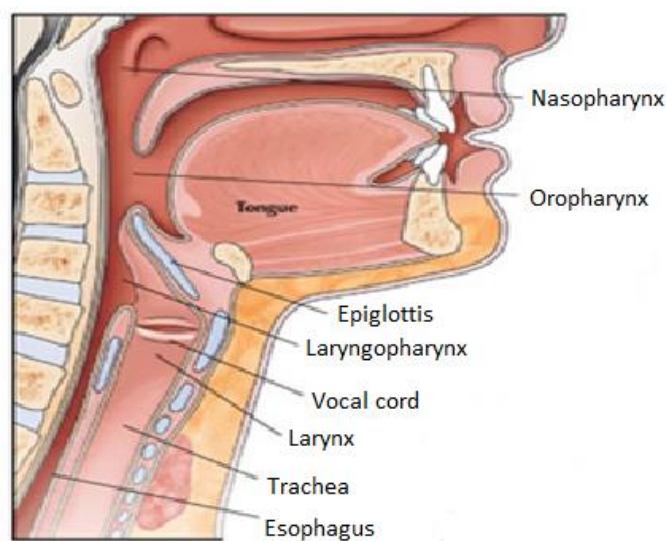


Figure 1.5: Anatomy of the Larynx.

<http://dnbhelp.wordpress.com/pharynx-larynx-2-2> (03.03.2016)

Larynx can be studied in three main sections [41]. These are:

Supraglottic area: Upper part of glottic area, between Hyoid bone and vocal cords. In this area, there are epiglottis, ventricular bands and ventricles.

Glottic area: This is where airstream from subglottic area changes into vibration with vocal cords. Larynx in this area has three important functions as respiratory, sphincteral and phonatory.

Subglottic area: The lower part of the larynx, the area from just below the vocal cords down to the top of the trachea

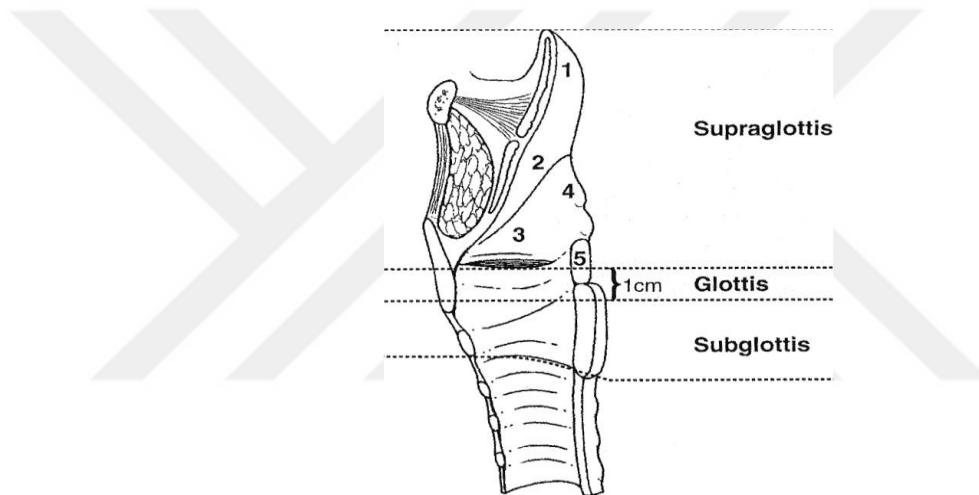


Figure 1.6: Parts of Larynx.

(Picture: Merati, A.L., Bielamowicz, S.A., (2007). Textbook of Laryngology, 31-349)

1.5.2.4. Vocal Cords (Vocal Folds)

Vocal cords are composed of twin infolding of mucous membrane stretched horizontally, from back to front, across the larynx. Larynx area between vocal cords is called glottis. In the upper part, there are ventricles and false vocal cords. These structures both affect the resonance quality of sound and protect the respiratory tract during swallowing.

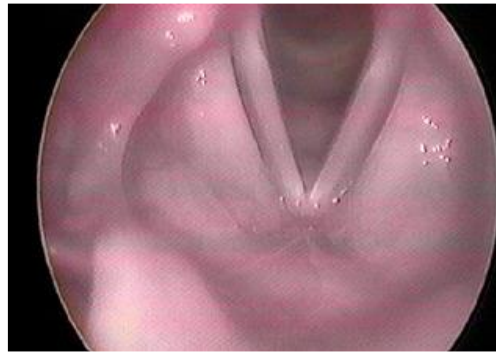


Figure 1.7: The Position of Vocal Cords during Respiration.

<http://professionalvoice.org/default.aspx> (15.03.2016)



Figure 1.8: The Position of Vocal Cords during Normal Tone of Sound.

<http://professionalvoice.org/default.aspx> (15.03.2016)

1.5.2.5. Pharynx

Pharynx is the part behind the oral and nasal cavity of the digestive system. It allows food taken to pass through oesophagus. It also plays an active role in the respiratory system. Pharynx is a 50 cm-long, muscular and mucosal tube. The tube begins at the base of the skull and ends inferior to the cricoid cartilage. It makes up the part of the throat situated immediately posterior to the nasal cavity, posterior to the mouth and superior to the oesophagus and larynx. The oesophageal opening is the narrowest part of the pharynx [42].

1.5.3. Organs Forming Oral Cavity

It is the most changing element during speech so it is very important [43]. The shape of oral cavity can change with tongue movement or change of mouth shape [43].

1.5.3.1. *Lips*

Lips are very important for the articulation of consonant and vowel phonemes. Most external articulators are the lips [43].

1.5.3.2. *Teeth*

Teeth are parts that form the lateral and anterior side of the oral cavity and that are involved in sound production with anterior part of the tongue [43].

1.5.3.3. *Alveolar Ridge*

The alveolar ridge is a small protuberance just behind the upper front teeth. The part between the teeth and hard palate is alveolar ridge [43].

1.5.3.4. *Palate*

Palate is the roof of the mouth. It extends from alveolar ridge to velum [43].

1.5.3.5. *Tongue*

Tongue is a mucous membrane-covered muscle mass. The posterior part is, at its root, connected with the hyoid bone. Larynx is also connected with the hyoid bone. Therefore, the tongue position affects the height of the hyoid bone. In this way, it also affects the position of larynx in the neck. When the tongue is extended forward, larynx goes up a little bit; when the tongue is drawn back, larynx goes down. Tongue is very important in producing healthy and accurate sound. Especially the position of the tongue in the mouth and tongue root are among the factors that affect the sound resonance [34]. Below are the parts of the tongue [43].

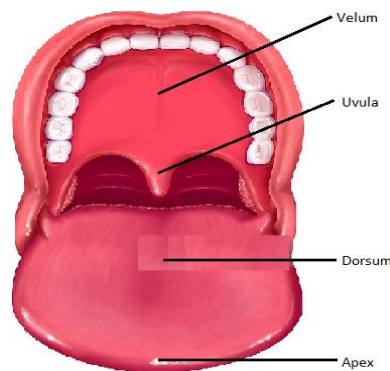


Figure 1.9: Parts of the Tongue.

1.5.4. Formation of Speech and Modelling

That humans can transform sound to human speech and communicate is the most important feature that makes them different from all other creatures. Sound, the fundamental element of speech, has been the subject of curiosity and research throughout history. First recorded studies on sound date back to 5th century BC. Hippocrates noted the importance of lung, trachea, lips and tongue for speech. Aristotle did scientific research on sound and identified the relationship between sound and the emotions. Claudius Galen, who lived between the years 131-201, in addition to vital functions such as breathing and swallowing, defined the larynx which plays an important role in phonation which is a part of man's social life [37]. In 1854, Manuel Garcia invented the indirect laryngoscopy technique and tried to observe the vocal cords. In 1940, sound spectrograph was developed by Potter, Kopp and Green. Over time with the development of technology, it has been found that different anatomical structures and physiological systems are involved in the formation of sound [38].

1.5.4.1. *Formation Stages of Speech*

Speech sound can be studied in four main sections. These are respiratory, phonation, resonance and articulation.

Respiration

The primary task of respiratory system is to provide the body's oxygen need. Its secondary function is to provide the energy for sound and speech [41]. Respiration enables vocal cords to open [38].

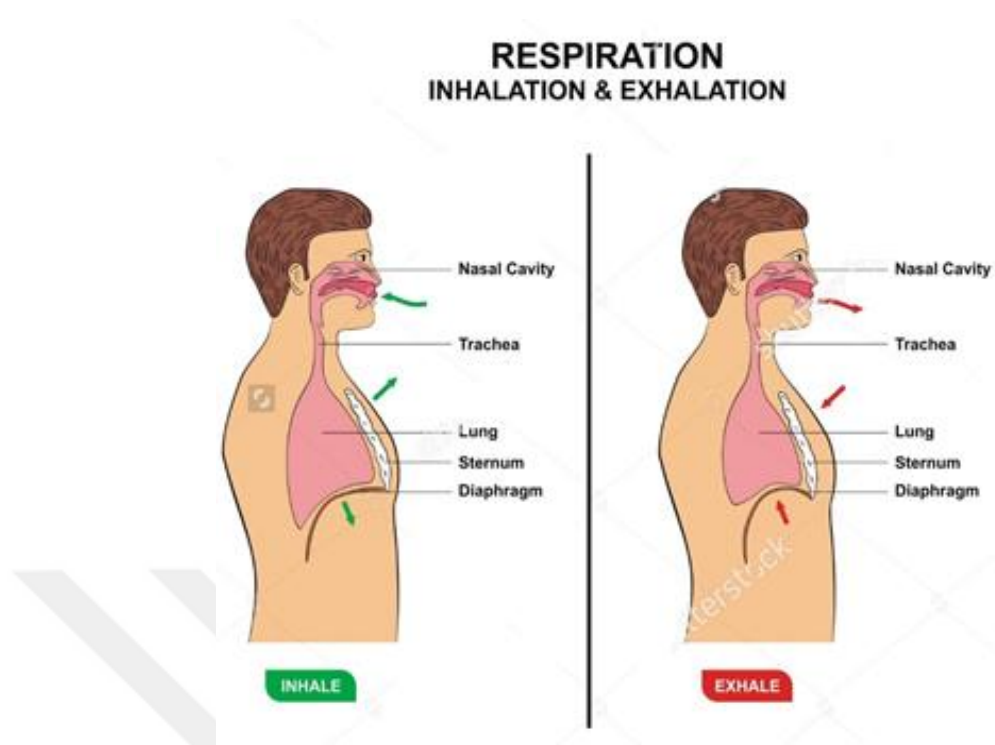


Figure 1.10: Respiration.

<http://www.shutterstock.com/pic-90513595/stock-vector-vector-respiration-inhalation-exhalation.html> (13.05.2016)

Phonation

It is the vibration of vocal cords [41]. The five conditions must be satisfied for proper phonation. These are adequate breath support, approximation of vocal folds, favourable vibratory properties, favourable vocal fold shape, control of length and tension [44].

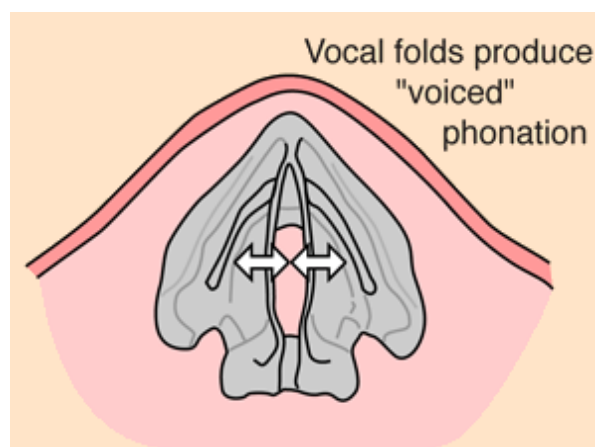


Figure 1.11: Phonation.

<http://hyperphysics.phy-astr.gsu.edu/hbase/music/voice2.html> (13.05.2016)

Resonance

Resonance, depending on the current status of the vocal tract, is the selective amplification of some harmonics. The resonance of the vocal tract is called “formant”. Formants are simple harmonic sounds that larynx produces in multiples of F_0 . Two formants below F_0 are needed to recognize sound. The bass one is F_1 (the first formant) and the treble one is F_2 (the second formant). One of these (F_1) belongs to pharynx and (F_2) belongs to oral cavity. The resulting sound waves are converted to sound and speech through articulatory organs such as teeth, lips and tongue [41].

Formant frequencies are the most important criteria for recognizing vowel sounds. For every person, these F_1 and F_2 values and the distance between the two formants are almost fixed and typical for vowel letters used in the same content. The first three formants, mainly F_1 and F_2 , give the sound its tone and colour [45].

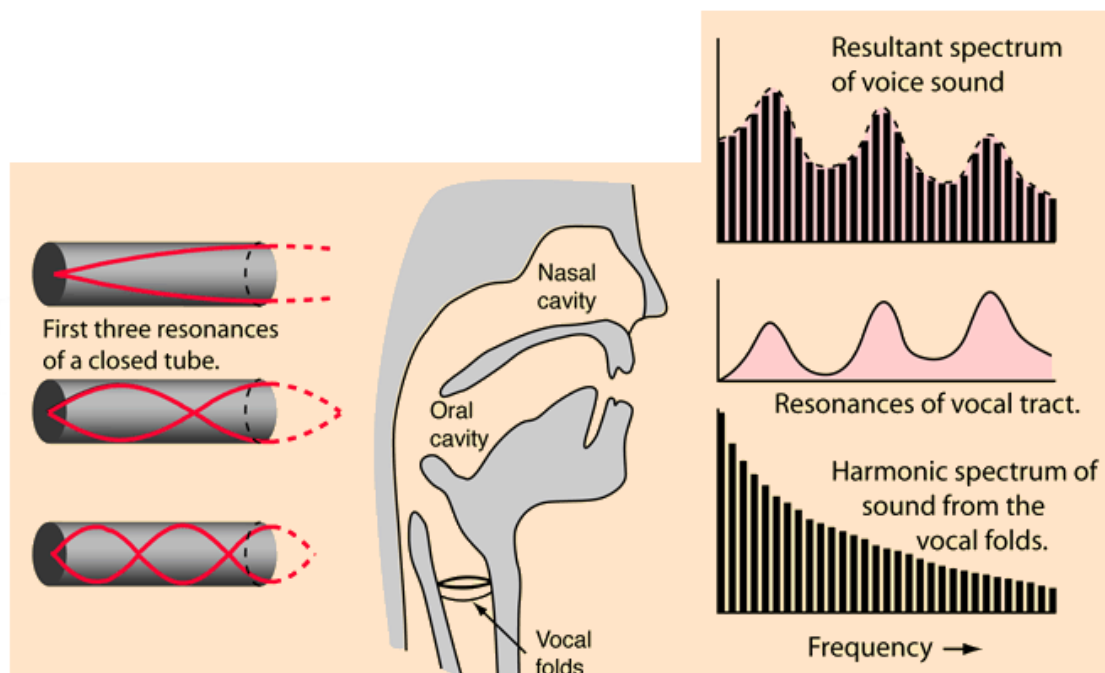


Figure 1.12: Resonance.

<http://hyperphysics.phy-astr.gsu.edu/hbase/music/vocres.html#c1> (13.05.2016)

Articulation

It is the formation of speech sound by processing the sound produced with phonation with the help of organs like tongue, lips and soft palate [41].

According to the source-filter theory, larynx is the source for fixed sound formation. Tongue, lip, palate and pharynx are filters that convert this sound to words. Accordingly, lingual movements, mainly vowels and consonants, are the result of dynamic movements of vocal tract. However recent studies have shown that the position and shape of glottis change during the production of different letters. Voiceless consonants (such as p, t, s) are produced from half-closed glottic opening. Voiced letters (such as b, d, l) are produced in conditions that vocal cords participate in vibration with all their structures (with the effect of aerodynamic muscle activity and elastic powers). Therefore, the larynx is the acoustic producer as well as an important articulator organ [46].

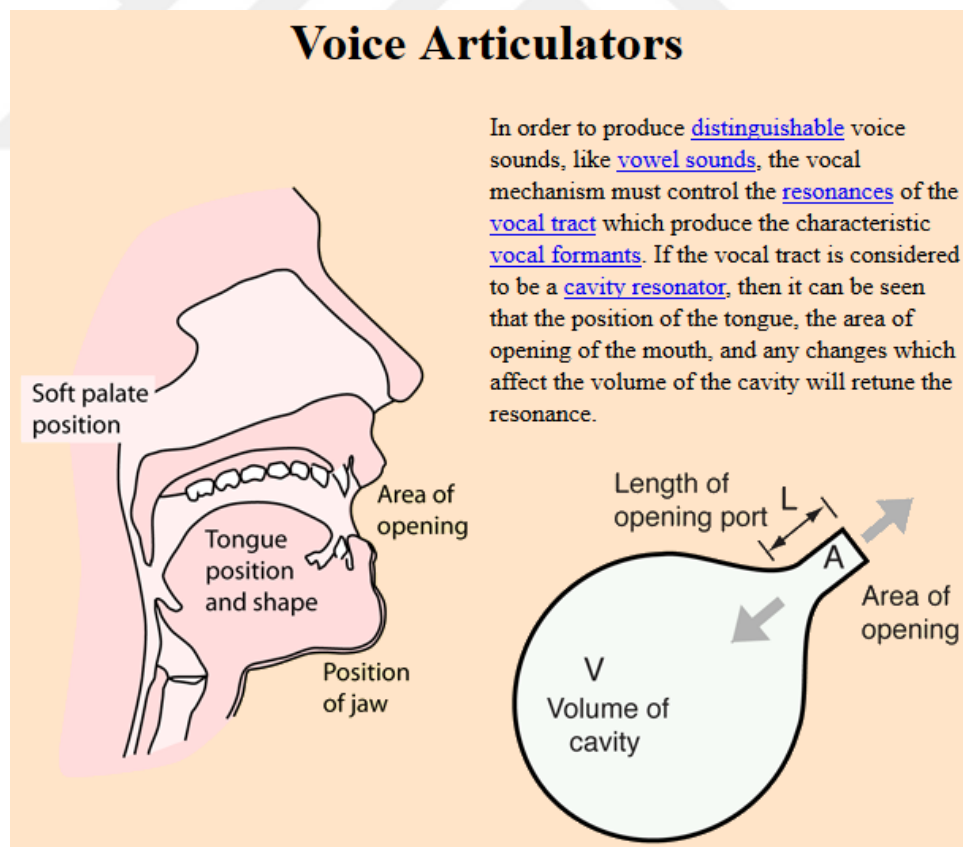


Figure 1.13: Articulation.

<http://hyperphysics.phy-astr.gsu.edu/hbase/music/vocres.html#c2> (13.05.2016)

1.5.4.2. Source-Filter Model of Speech Production

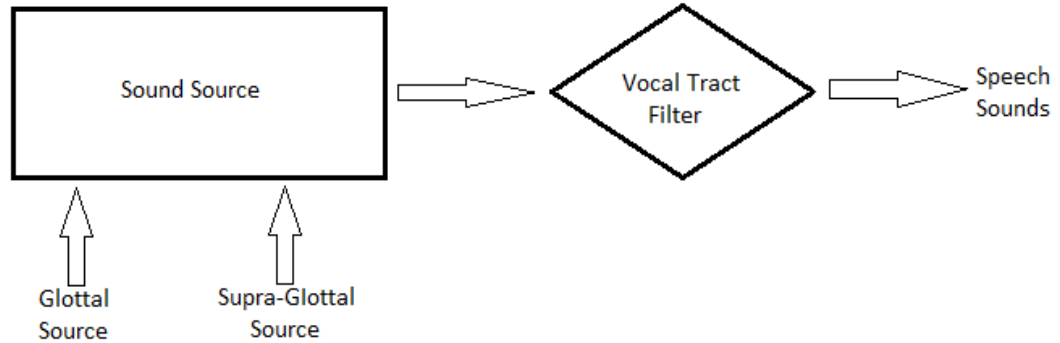


Figure 1.14: Source-filter model of speech production.

The source-filter theory describes speech production as a two stage process involving the generation of a sound source, with its own spectral shape and spectral fine structure, which is then shaped or filtered by the resonant properties of the vocal tract. Most of the filtering of a source spectrum is carried out by that part of the vocal tract anterior to the sound source. In the case of a glottal source, the filter is the entire supra-glottal vocal tract. The vocal tract filter always includes some part of the oral cavity and can also, optionally, include the nasal cavity (depending upon whether the velum is open or closed). Sound sources can be either periodic or aperiodic. Glottal sound sources can be periodic, aperiodic or mixed. Supra-glottal sound sources that are used contrastively in speech are aperiodic although some trill sounds can resemble periodic sources to some extent. Periodic and aperiodic sources can be generated simultaneously to produce mixed voiced and aperiodic speech typical of sounds such as voiced fricatives. In voiced speech the fundamental frequency is a characteristic of the glottal source acoustics whilst features such as vowel formants are characteristics of the vocal tract filter (resonances) [47].

In addition, the source-filter model assumes that the source is linearly separable from the filter and that there is no interaction between them. This is not strictly true since the vibration of the vocal cords is affected by the sound pressure inside the vocal tract. There is coupling between the vocal tract and the lungs during the period when the glottis is open, thereby modifying the filter characteristics every cycle of the excitation [48].

2. MATERIALS AND METHODS

The aim of this chapter is to introduce the properties of the materials that were used and the methods that were performed throughout this research.

2.1. CLINIC AND DATA ACQUISITION

The study design was a prospective cohort study and only voluntary patients were included into the study consecutively after the application of exclusion criterias.

All patients enrolled in the study underwent complete clinical history, otorhinolaryngological physical examination and flexible nasopharyngolaryngoscopy. Patients were eliminated from the research who had basic problems on breathing, voice, and structural abnormalities in the larynx, mouth, or throat. All subjects met the following criteria: no history of smoking, heavy drinking, hearing impairment, vocal cord surgery, respiratory problems.

This study was conducted in 34 patients with NP and 33 healthy individuals in the Research Hospital of the Medicine Faculty at Turgut Ozal University between November 2013 and December 2015. During this period, the total number of individuals applied to Ear, Nose, and Throat (ENT) clinic of the hospital with a NP complaint was 42, and 34 of them volunteered to participate in the study.



Figure 2.1: Audiology Hearing Balance Lab. at Turgut Ozal University.

The study group included 34 patients with nasal polyp (25 males, 9 females) with age range 27-62 years. The control group included 33 healthy volunteers (25 males, 8 females) with no voice disorder and age range of 24-56 years. The average age in NP was 44.1 years and in healthy was 37.2 years.

The procedures applied in this study were approved by the Ethical Committee.

The proposed work was carried out the following steps:

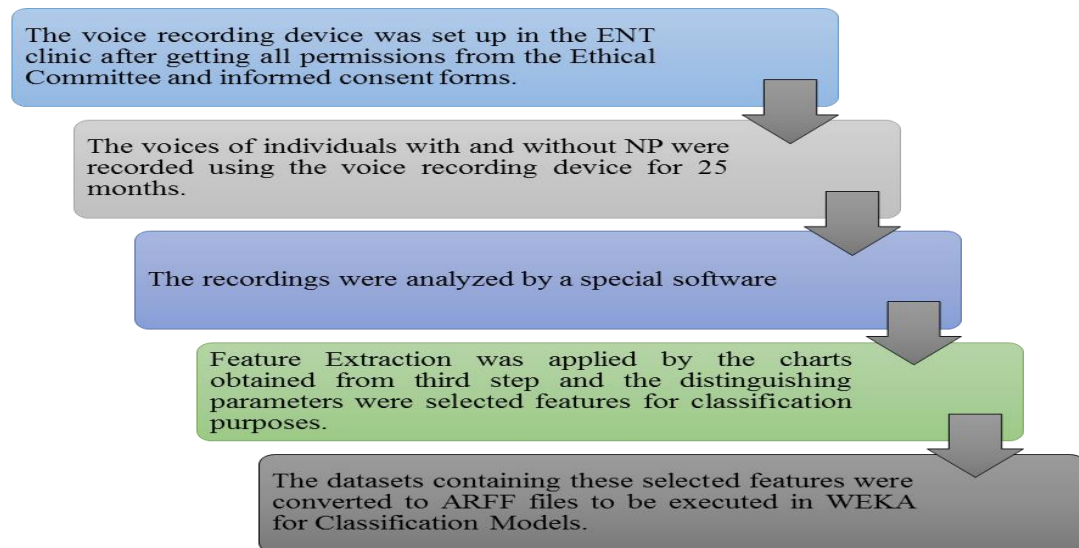


Figure 2.2: The process steps of proposed work.

2.2. INSTRUMENTATION AND MEASUREMENT

2.2.1. Turkish Alphabet Acoustics

When examining the Turkish Alphabet Acoustics as shown in Table 2.1, m and n are expressed as nasal voices and these voices originate from nasal regions especially from nasopharynx [1]. Nasal polyp causes nasal obstruction [49]. Accordingly, this obstruction causes distortion in the formation of the nasal voice. In this work, “mmmm” and “nnnnnn” sounds of individuals were recorded. In addition, the patients were requested to say the words “engin”, mangal”, “cengiz”, “kangal” in order to obtain a sound with "ng" emphasis originating from nasal regions [50].

Table 2.1: Turkish Alphabet Acoustics.

Vowels	Wide	Narrow	Wide	Narrow
Velarized	a	I	o	U
Front	e	I	ö	Ü
Consonants	Fricative	Explosion	Nasal Passage	Slippery
Voiced	c, j, v, z	b, d, g	m, n	ğ, l, r, y
Voiceless	ç, f, h, s, ş	t, k, p	-	-

2.2.2. Electroglottograph

The electroglottograph, or EGG, (sometimes referred to as a laryngograph) is a device for the noninvasive measurement of the time variation of the degree of contact between the vibrating vocal folds during voice production. Though it is difficult to verify the assumption precisely, the aspect of contact being measured by a typical EGG unit is considered to be the vocal fold contact area (VFCA) [51, 52].

The technique was first applied to voice work by Fabre (1957) and he used a constant current, impedance based, method of observation. He also introduced the term “électroglottographie” considering that the main information derived from the device was in respect of the glottal opening. In fact, however, the main information concerns vocal fold contact, and it was for this reason, early on, that the term “Laryngograph” was used [53].

Initially, essentially only the Lx signal was presented in the form of a waveform going positive (up) for increasing vocal fold contact area (this was also Fabre’s aim, but his

actual waveforms were positive for increasing glottal opening). As the result of the work described above, and many other contributions, the near universal practice is now to use both EGG as well as Lx information with positive contact polarity [53].

Baken states that in normal voice the main acoustic excitation of the vocal tract coincides with vocal fold closure and since this is a salient aspect, the Lx signal is widely relied on to provide larynx frequency information. In recent years this has led to the combined use of Lx waveform information together with laryngeal stroboscopy. A first step towards routine clinical use was made by Anastaplo and Karnell in 1988 so that both synchronisation and waveshape came from an Lx signal source [53].

2.2.2.1. Function

EGG or Electroglottograph is a system which gives an information on the closure of vocal folds by measuring the electrical resistance between two electrodes placed around the neck. Even if the signal provides only an approximate value of the glottis surface, it gives a very good information about period of vocal folds vibration, especially because we are at the source of the phonation (no influence of vocal tract) and because there is no aerodynamic noise [54].

To measure VFCA, an EGG records variations in the transverse electrical impedance of the larynx and nearby tissues by means of a small A/C electric current in the megaHertz region applied by electrodes on the surface of the neck. This electrical impedance will vary slightly with the area of contact between the moist vocal folds during that part of the glottal vibratory cycle in which the folds are in contact [55].

2.2.2.2. EGG waveform

Early commercial available EGG units were compared quite thoroughly by Baken. However, using modern low noise electronics, EGG noise levels can be brought down enough so that the noise is approximately 40dB (a factor of 100) less than a typical EGG signal from an adult voice [55].

Electroglottographic wavegrams are a new technique for displaying and analyzing EGG signals. The technique provides an intuitive means for quickly assessing vocal fold contact phenomena and their variation over time [56].

Electroglottograph signals are used in stroboscope synchronization, voice fundamental frequency tracking, tracking vocal fold abductory movements and the study of the singing voice [57].

Egg signal is composed of:

a high-frequency component which is relative to vocal folds vibration (voice)

a low-frequency component which is relative to low movement of larynx (ex : swallowing).

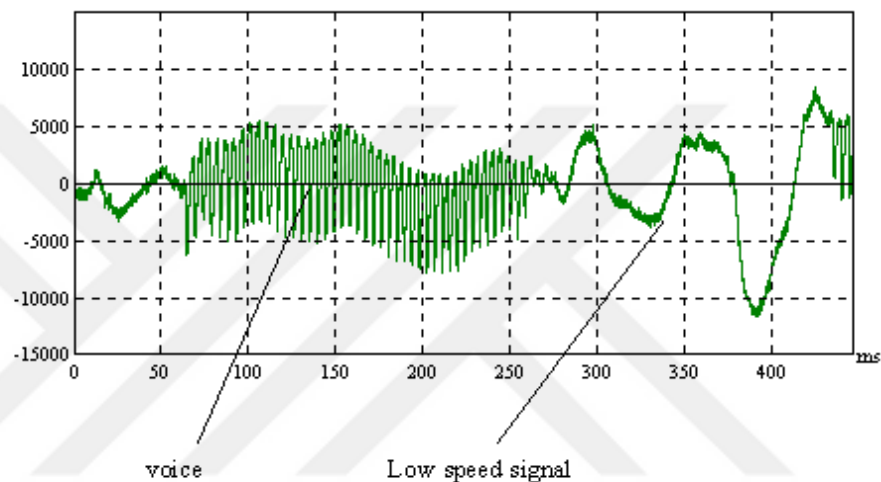


Figure 2.3: Raw waveform of EGG.

<http://www.lpl-aix.fr/~ghio/pedago-EggUK.htm> (14.05.2016)

For a good analysis, the system filters as following:

1) filtering for voice

- high-pass filter (to cut low frequencies)
- anti-aliasing low-pass filter

2) filtering for swallowing

- low-pass filter (to keep low frequencies)

Egg waveform display

The waveform of the glottal cycle can be displayed in different ways. Upward movement of the curve can indicate more (1) or less (2) vocal fold contact:

(1) upward $\Leftrightarrow$ closure

(2) upward $\Leftrightarrow$ opening



Figure 2.4: Phases of the EGG signal.

<http://www.lpl-aix.fr/~ghio/pedago-EggUK.htm> (14.05.2016)

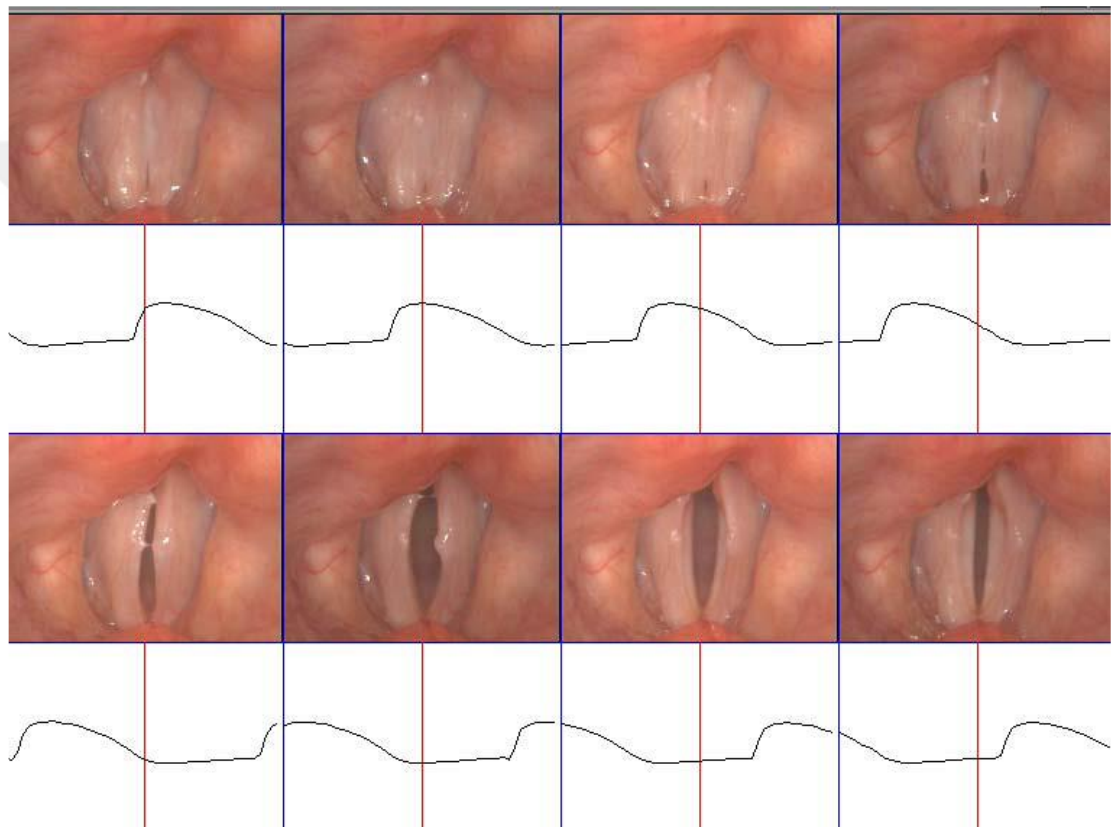


Figure 2.5: During of speech position vocal cord.

<http://www.laryngograph.com/pdffdocs/paper3.pdf> (14.05.2016)

2.2.2.3. EGG Device

EGG allows for laryngographic assessment of phonatory function during connected speech, soft and loud singing. It gives qualitative and quantitative information about the contact, separation and open phase of vocal fold vibration and is therefore of great use in the interpretation of pathological conditions and follow-ups after surgical treatment or conservative voice therapy. It has proved particularly useful in assessment and documentation of organic voice disorders [58].

Before each new laryngographic examination of vocal fold vibration two electrodes are placed on the subject's neck at the level of the thyroid cartilage. The EGG detects variation in the electrical resistance of the current passed between the electrodes. The resistance is produced by changes in the nature and area of vocal fold contact during phonation. The electrolaryngographic output from the speaker's voice activity is displayed as a laryngograph waveform showing horizontal opening and closing movements of the vocal folds. Many of commercially available devices make it possible to record an additional signal from a microphone, which is synchronized with the laryngographic output, so that both waveforms can be directly compared when displayed one on top of another [58].

EGG

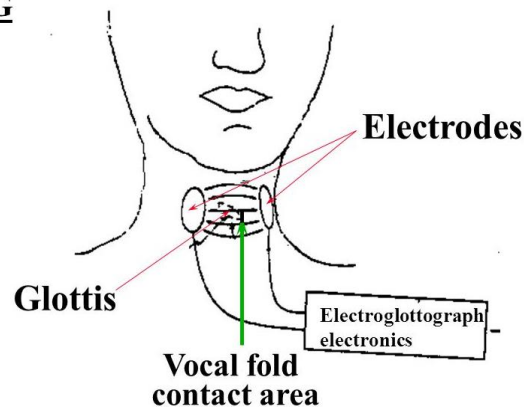


Figure 2.6: EGG coupling/application.

<http://slideplayer.com/slide/5289154/> (14.05.2016)



Figure 2.7: The early EGG devices (EGG-A100 model).

The early EGG devices were produced by Laryngograph Ltd and the first model EGG-A100 model as shown above. We have made data collection by using the Speech Studio Software and Laryngograph microProcessor EGG-D400 device which is latest model.

The model we used Laryngograph microProcessor EGG-D400 device was produced according to the standards of IEC 60601 and IEC 60950. Furthermore, the software used by the device was certified by European Communities Council Directive 93/42/EEC Concerning Medical Devices. The device was used in many published articles in journals indexed in esteemed international databases [14, 15, 59].

2.2.3. Instrumentation and Measurement

The speeches of individuals were recorded by the device of Laryngograph Micro Processor EGG-D400 and analyzed with its own software called Speech Studio Software. Attached to the recording device as shown in Figure 2.8, there were two surface electrodes, one of which was used to measure the fundamental frequency produced by vocal cords and connected to a part of the larynx, and the other one, vibration transducer, was used for measuring the value of nasal sound and connected to the outer surface of the polyp region. A microphone was used to record the sound of speech.

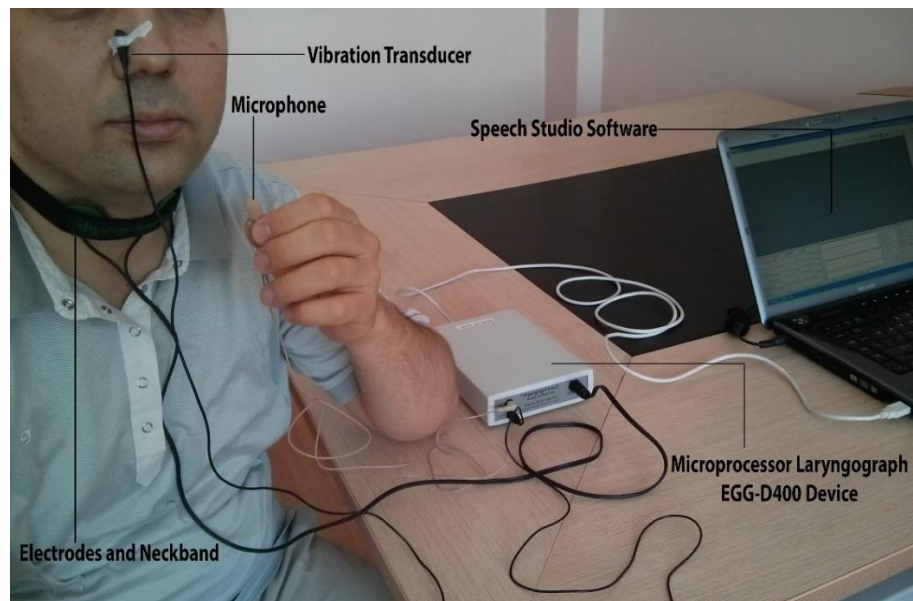


Figure 2.8: A measuring with microProcessor Laryngograph EGG-D400 device.

(Picture: Emrah Orun)

2.3. SIGNAL PROCESSING AND CLASSIFICATION

2.3.1. Classification Models

In this study, we used different classification methods in Waikato Environment for Knowledge Analysis (WEKA). The selected data sets were converted to ARFF form, which is a standard file for WEKA to be used as training sets for the formation of the classification model. WEKA, standing for the Waikato Environment for Knowledge Analysis, is a java-based, open-source data mining (DM) platform developed at the University of Waikato, New Zealand [60, 61]. It is actually a collection of machine learning algorithms used for many real applications. In the literature, there are many publications using WEKA in biomedical applications, such as prostate cancer [62], diagnosis of colon cancer [63], and rheumatoid arthritis [64]. Due to the machine learning property, WEKA is a powerful DM tool for tasks of data clustering, classification, regression, and preprocessing. There are different classification models in the machine learning library provided by WEKA: Ada Boost (AB) [65], Multi-Layer Perceptron (MLP) [66], Naive Bayes (NB) [67], Random Forest (RF) [68], LibLinear (LL) [69], J48 [70] and Spegasos. We, in our study, used 8 different algorithms in WEKA.

2.3.1.1. *Artificial Neural Network (ANN) / Multi-Layer Perceptron (MLP) Model*

The ANN is a mathematical model that is used to solve many problems in different fields. It imitates the decision making process of human beings although it is not accepted as a good simulation of the human brain. An ANN is trained with the available data set to extract the relation between inputs and outputs to reach desired outputs when new data is input. In an ANN, there are generally three layers; namely, the input layer, the output layer and hidden layer as seen in Figure 2.9. The hidden layer can consist of one or more hidden layers depending upon the network topology. In the input layer, raw input data is given to the network. In the hidden layer, the weights of each input are adjusted to produce the desired output. The behavior of the output layer merely depends upon the hidden layer and connection weight between these two layers. The ANN is broadly used in many different fields, such as pattern recognition, intelligent control, optimization, regression analysis, data filtering and data processing, diagnostic processes in medical applications and game programming [71].

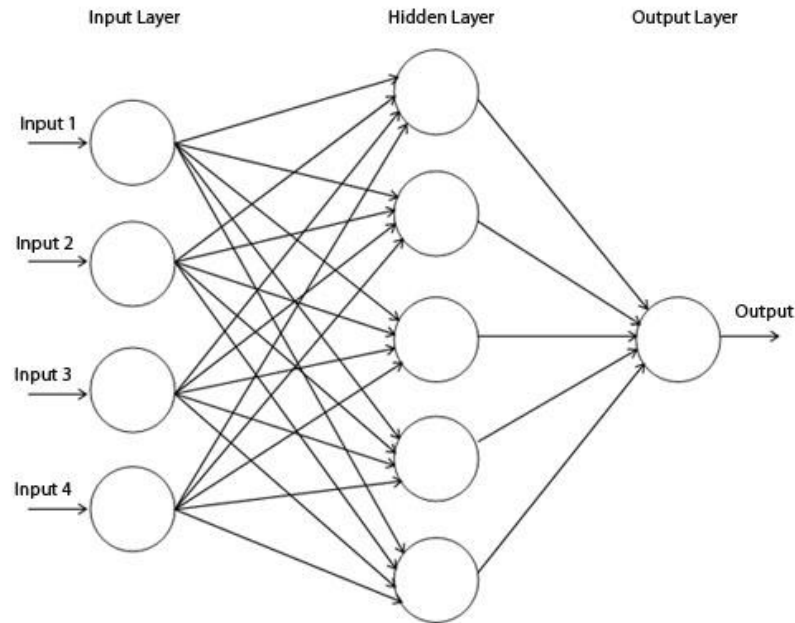


Figure 2.9: A simple three layer feed forward neural network.

(Picture: Witten, I.H. & Frank, E. (2011). Data Mining: Practical Machine Learning Tools and Techniques, New York, USA)

The ANN is employed in many applications in the biomedical engineering field. Since the diagnosis processes can be handled as a classification problem, the usage of an ANN or other expert system based classifiers become valuable. In classification of ECG signals, the ANN is used for making decisions. Moreover, the ANN is used to classify features obtained from EEG, EMG, Doppler audio recordings of blood flow, images obtained from X-rays, Tomography, MRI, etc. [71].

MLP (Multi-Layer Perceptron)

In biomedical applications, the most commonly applied network structure is MLP. There are several training algorithms for MLP networks but the LM error back-propagation training algorithm is generally used in biomedical applications. Similar to other ANN architectures, a MLP network can have one input layer, one output layer and one or more hidden layers. The number of neurons at the input layer is equal to the number of features of the input set. For classification problems, the number of neurons at the output layer depends upon the number of classes. If, for example, there are two classes to be categorized, then one neuron is enough. If there are more than two classes, then the number of neurons must be equal to the number of classes. Generally there is

one hidden layer in the literature but it can be two or more depending upon properties like the size of problem, solution method, and number of features. Hidden layers increase the ability of classification performance, especially in nonlinear classification problems. However, when the number of hidden layers increases, the training time increases considerably. In addition to deciding the number of hidden layers, setting the appropriate number of neurons at each layer is also a critical task. The most widely used strategy here is the conventional trial and-error method. One important issue related to network design is deciding the number of neurons at the hidden layers. It highly affects the classification performance of the ANN. If the number of neurons is set too low, then the classification performance can fail, especially for complex patterns. If the number is set too high, then training time will increase unnecessarily. Furthermore, the network can lose its generalization capability because of memorization of individual samples during training. After selecting the network architecture and deciding the number of hidden layers and the number of neurons at these layers, one must decide the training algorithm. In biomedical applications, typically the Levenberg-Marquardt Back-propagation (LM) algorithm is utilized with MLP architecture. This algorithm was proposed by Levenberg in 1944 and by Marquard in 1963. It is an approximation of the Gauss-Newton method based on least squares estimation for nonlinear parameters. It processes much faster than conventional back propagation algorithms while requiring more memory [71].

2.3.1.2. *Spegasos Model*

Implements the stochastic variant of the Pegasos (Primal Estimated sub-GrAdient Solver for SVM) method of Shalev-Shwartz et al. in 2007. This implementation globally replaces all missing values and transforms nominal attributes into binary ones. It also normalizes all attributes, so the coefficients in the output are based on the normalized data [72].

2.3.1.3. *Naive Bayes Model*

The Naive Bayes classifier is an anticipated classifier used in machine learning methods. Naive Bayes classifier is the version of Bayes theorem simplified with proposition of independence. Bayes theorem is expressed by the following equation [68];

$$P(A/B) = \frac{P(B/A)P(A)}{P(B)}$$

$P(A/B)$: When event B occurs, is the probability of event A

$P(B/A)$: When event A occurs, is the probability of event B

$P(A)$ and $P(B)$: Is the priori probabilities of event A and B

Naive Bayes classifier is generally used in data mining, in biomedical engineering, identifying the medical diseases or abnormalities [73], in classification chart of electrocardiogram (ECG) [74], in separation of graph of electroencephalography (EEG) [75] and used in some other areas.

2.3.1.4. Logistic Model

Logistic regression is a part of a category of statistical models called generalized linear models. Logistic regression allows one to predict a discrete outcome, such as group membership, from a set of variables that may be continuous, discrete, dichotomous, or a mix of any of these. Generally, the dependent or response variable is dichotomous, such as presence/absence. The independent or predictor variables in logistic regression can take any form. That is, logistic regression makes no assumption about the distribution of the independent variables. They do not have to be normally distributed, linearly related or of equal variance within the group. There are two main uses of logistic regression. The first is the prediction of group membership. Since logistic regression calculates the probability of success over the probability of failure, the results of analysis are in the form of an odds ratio [76].

2.3.1.5. RBF Network Model

Radial basis function neural network can be used for a wide range of application primarily because it can approximate any regular function and its training is faster than that of feed-forward network. The RBF used in this work consists of three layers: input layer, which corresponds to the input vector space, output layer, which corresponds to the pattern classes and hidden layer based on radial neurons. Training of RBF means determining the number of basis functions (hidden layer neurons), centers and widths of each basis function and output layer weights. In this case the output layer is a single

neuron with logistic activation function. Therefore architecture of neural network is fixed by determining the hidden layer: number of neurons and width of radial functions [76].

2.3.1.6. *Simple Logistic*

Simple logistic regression is analogous to linear regression, except that the dependent variable is nominal, not a measurement. One goal is to see whether the probability of getting a particular value of the nominal variable is associated with the measurement variable; the other goal is to predict the probability of getting a particular value of the nominal variable, given the measurement variable [77].

2.3.1.7. *SMO Model*

Sequential minimal optimization (SMO) is an algorithm for solving the quadratic programming (QP) problem that arises during the training of support vector machines. It was invented by John Platt in 1998 at Microsoft Research. SMO is widely used for training support vector machines and is implemented by the popular LIBSVM tool. The publication of the SMO algorithm in 1998 has generated a lot of excitement in the SVM community, as previously available methods for SVM training were much more complex and required expensive third-party QP solvers [78].

2.3.1.8. *J48 Model*

An important knowledge structure that can result from data mining activities is the decision tree (DT) that is used for the classification of future events. Decision trees are typically built recursively, following a top-down approach. The acronym TDIDT, which stands for top-down induction on decision trees, refers to this kind of algorithm. One branch is a chain of nodes from root to a leaf; and each node involves one attribute. The occurrence of an attribute in a tree provides the information about the importance of the associated attribute. J48 algorithm (a WEKA implementation of C4.5 algorithm) is a widely used one to construct decision trees. Decision tree algorithm (J48) has two phases: building and pruning. The building phase is also called as the ‘growing phase’. J48 uses entropy based information gain as the selection criterias [79].

2.3.2. Training Methods

In this study, we used eight classification algorithms on 40 datasets. We applied algorithms two different classification methods. These methods are Leave-One-Out Cross-Validation (LOOCV) method and N-Fold Cross-Validation Method.

The Leave-One-Out Cross-Validation (LOOCV) and N-Fold Cross-Validation Method [80] was used for classification model formation in WEKA. Each of 40 datasets was formed by including all of the samples except for one. The remaining sample was the test set. Therefore, for n samples, n different training sets and n different tests sets were obtained. The test dataset was then executed over the trained model, and the accuracy ratio for the test fold was calculated. This procedure was repeated for each data fold. In our study, the total number of executions was 40.

The test dataset (were divided randomly in 10 folds, $N=10$) is then executed over the trained model and accuracy ratio for this test fold is calculated. The accuracy of the classification model is the average of the N -folds accuracy ratios.

2.3.3. WEKA Application

This section presents the analysis results of proposed method using the dataset, which was obtained from the WEKA. Measures of the quality of classification are built from a confusion matrix which records correctly and incorrectly recognized examples for each class. Table 2.2 presents a confusion matrix for classification, where TP – True Positive, FP – False Positive, FN – False Negative, and TN – True Negative counts.

Table 2.2: A confusion matrix for classification.

<i>Classification</i>	<i>Positive</i>	<i>Negative</i>
<i>Positive</i>	TP	FN
<i>Negative</i>	FP	TN

$$\mathbf{Precision} = \frac{TP}{TP + FP} \quad (2)$$

and

$$\mathbf{Recall} = \frac{TP}{TP + FN} \quad (3)$$

All three measures distinguish the correct classification of data within different classes. They focus on positive example classes. Precision is a function of TP examples misclassified as FP examples Eq. (2). Similarly Recall (or sensitivity) is correctly classified TP examples and its misclassified FN examples. Recall gives information about a fraction of positives selected from the test set, whereas the precision defines the correctness of positive instances prediction Eq. (3) [81]. Low values of the latter parameter indicate a high rate of false positives.

$$\mathbf{F - Measure} = 2 * (\beta) \frac{\mathbf{Precision} * \mathbf{Recall}}{\mathbf{Precision} + \mathbf{Recall}} \quad (4)$$

F-Measure is the weighted harmonic mean of precision and sensitivity Eq. (4). The $\beta=1$ is a metric case of the more general measure. A detailed assessment of classification models performance can be obtained by the Receiver Operating Characteristic (ROC):

$$\mathbf{ROC} = \frac{(n|positive\ class)}{(n|negative\ class)} \quad (5)$$

In this formula $(n|class)$ indicates the conditional probability.

The ranges of F-Measure, Recall and Precision, ROC are from 0 to 1.

3. RESULTS

3.1. EXPERIMENTAL RESULTS

In this chapter, the analysis of audio recordings obtained from nasal polyposis patients and healthy individuals was presented.

The data collected by Laryngograph microprocessor device EGG-D400, have been analyzed through Speech Studio Software contained in the EGG device.

After making the connections of the device as seen in the figure 2.8, we run Speech Studio Software. We selected the microphone calibration and parameters that we will measure.

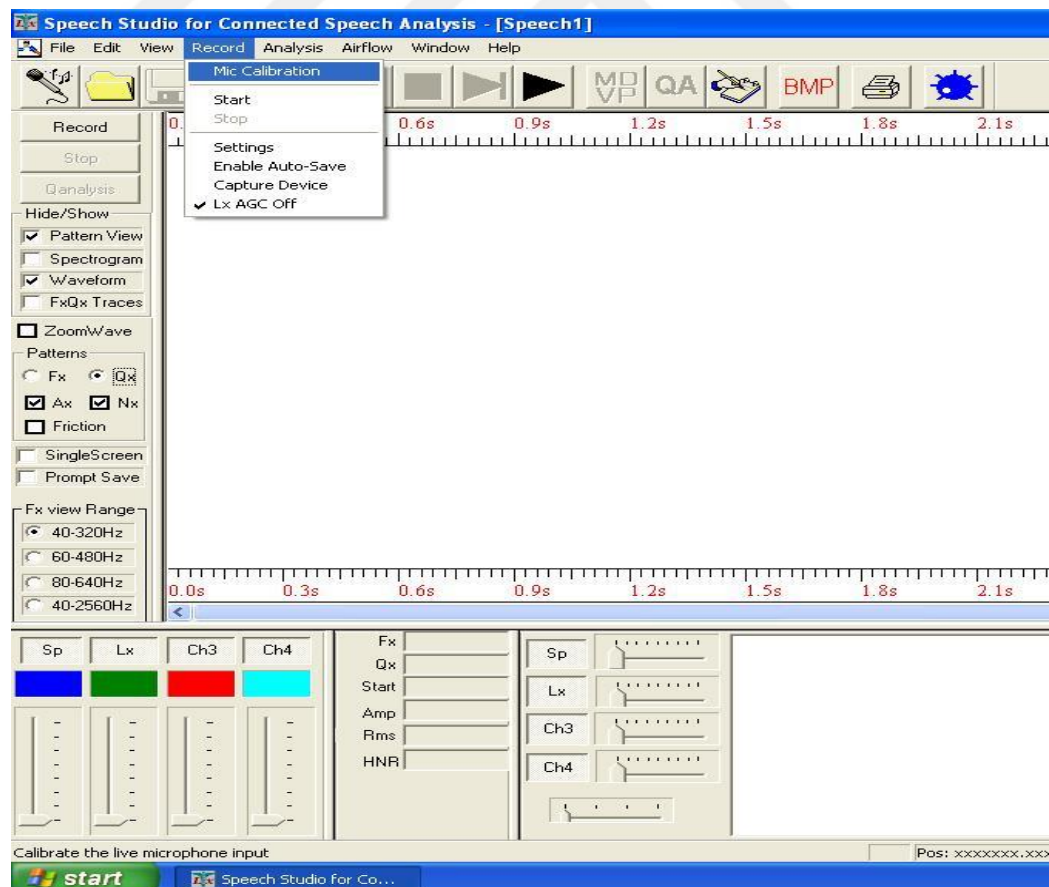


Figure 3.1: Selection of the microphone calibration.

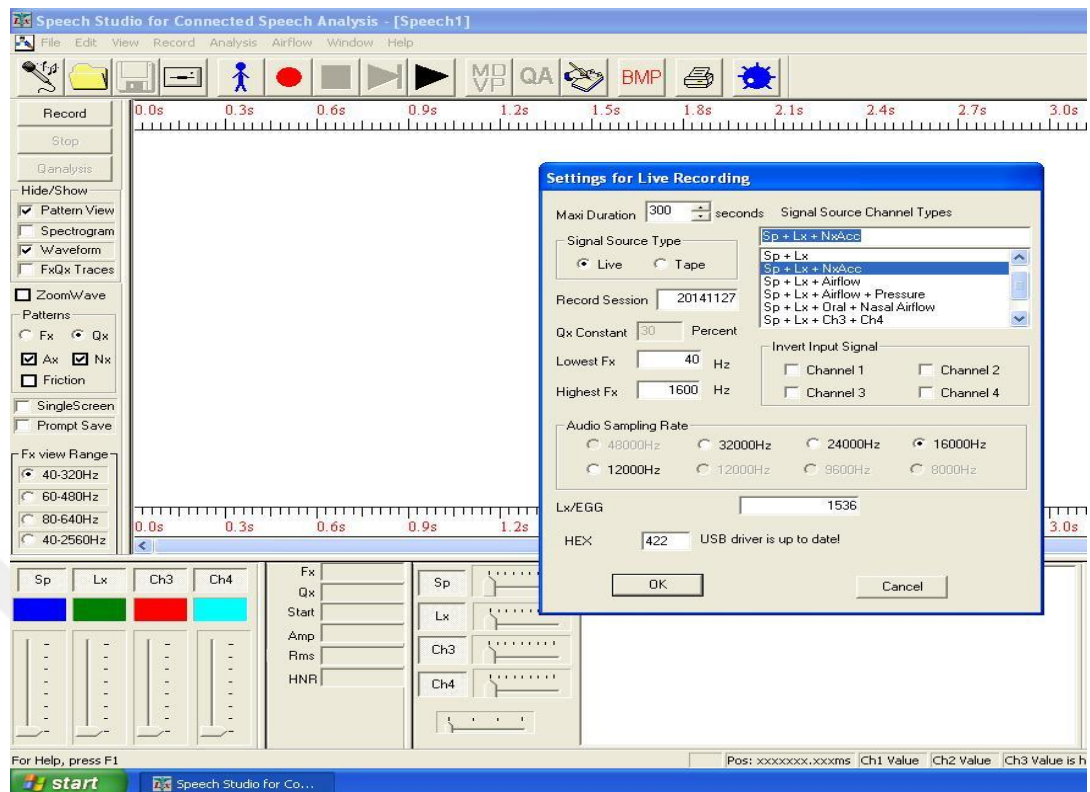


Figure 3.2: Selection of the device parameters.

After completing recording process we obtained the spectrum of sound.

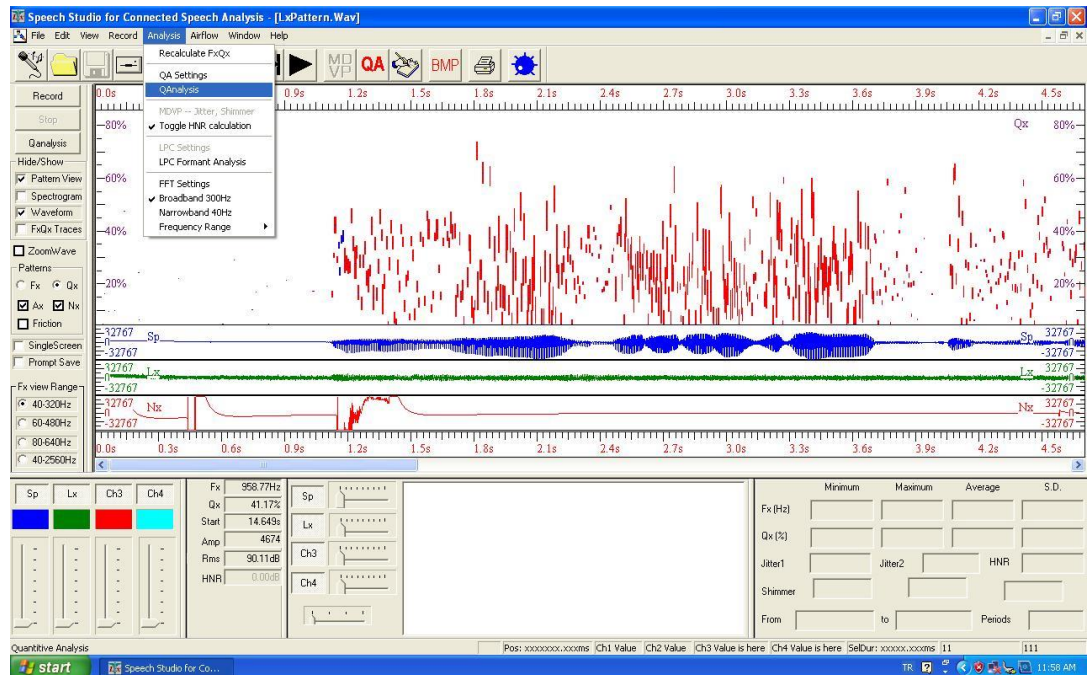


Figure 3.3: The spectrum of recorded speech.

The records we obtained in the above spectrum;

- With Sp (blue line) Acoustic signals,
- With Lx (green line) Laryngograph,
- With Nx (red line) Nasality,

measured these values.

We have released special graphics and methods obtained from values Quantitative Analysis(QA).

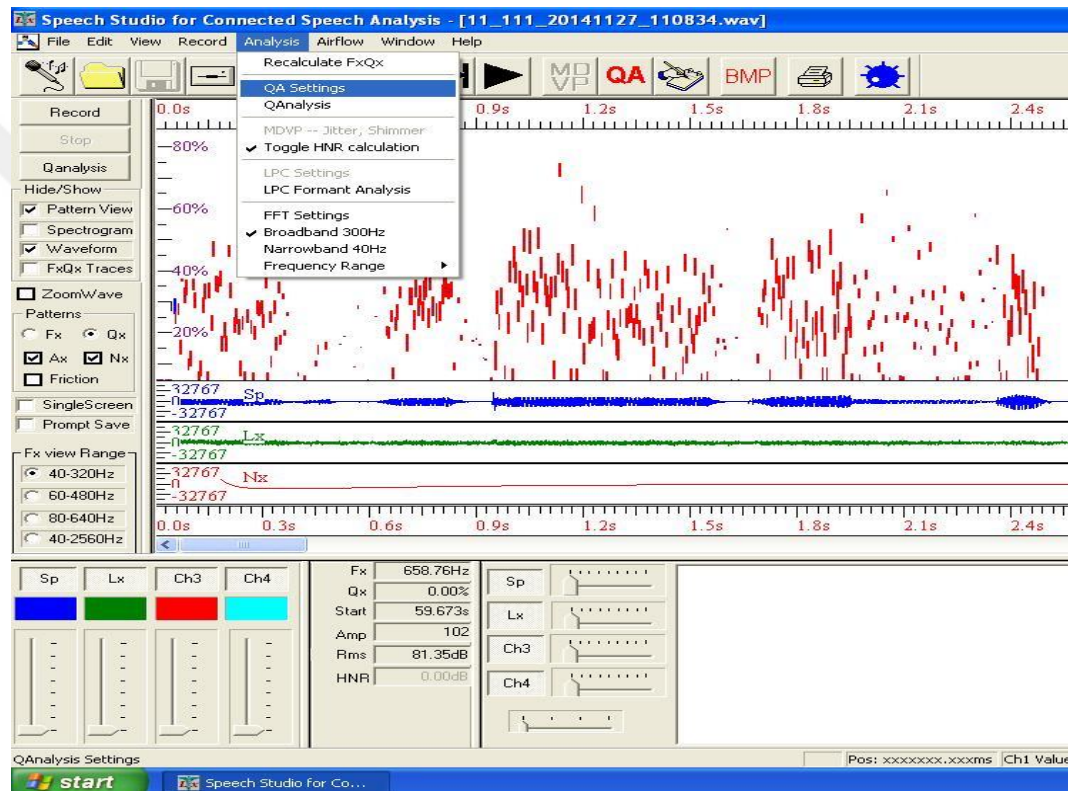


Figure 3.4: The QA Settings and Quantitative Analysis of recorded speech.

QA provides a range of connected speech and voice analyses that is available from Laryngograph. It has proved to be important both clinically and in voice research [82].

Speech Studio Software offers many analysis options. In our study, we use fundamental frequency, loudness, Speech Pattern Element (SPE) graphs and Nx/Sp Ratio graphs.

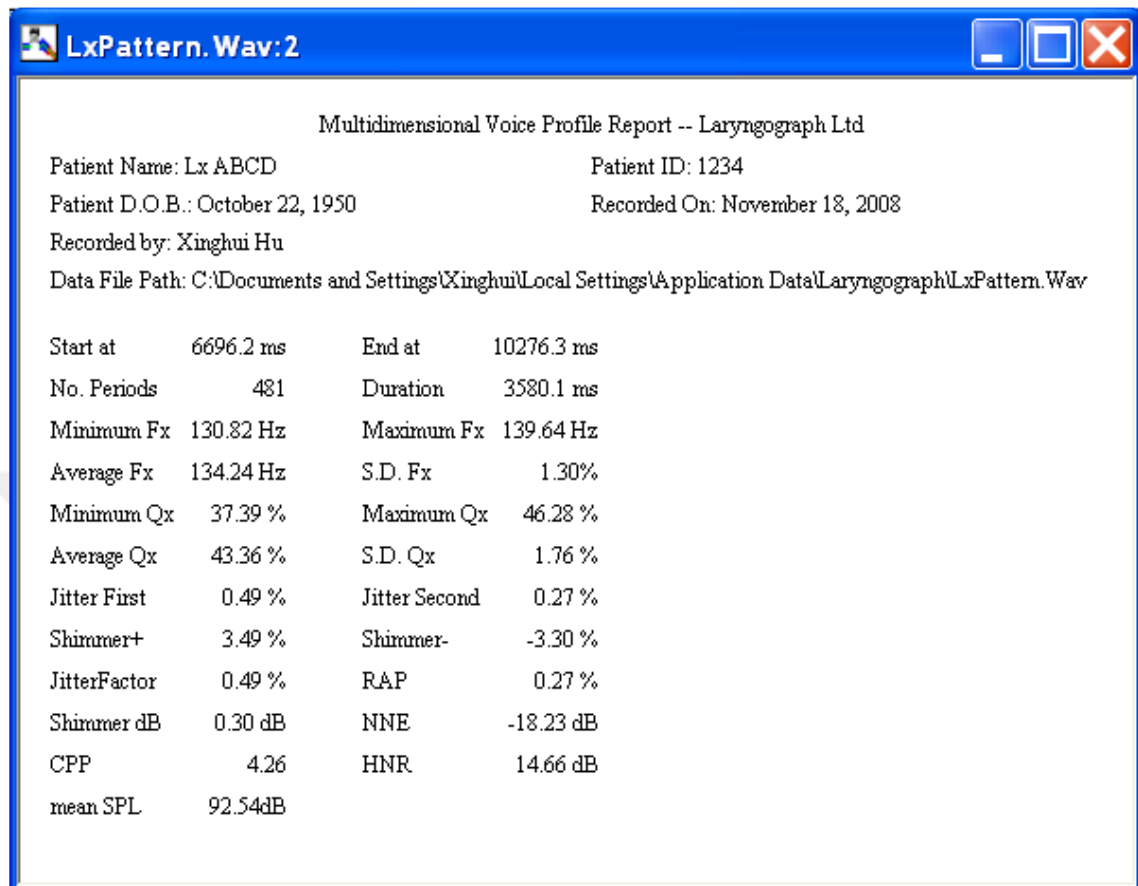


Figure 3.6: MDVP Results (Sample).

The next important voice parameter is loudness and the analyses of its correlate, amplitude (Ax), logically follow on in the analysis groups. Ax is the probability distribution for the peak amplitude of the acoustic (microphone) trace during each vocal fold cycle. The peak of the plot shows the most commonly occurring value (modal value). The plot is ordinarily symmetrical about the modal value with a suitable range.

Later, loudness graph (figure 3.7) of the sound recordings for each individual was obtained with the help of Speech Studio Software.

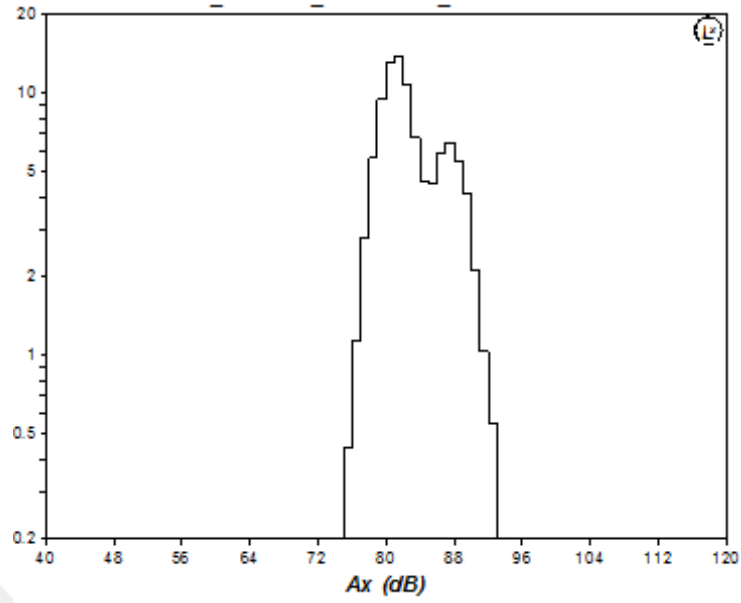


Figure 3.7: Loudness graph.

Minimum fundamental frequency values obtained for each word and the mean value of loudness graph were used as feature and classification was done. Table 3.1 shows the value range of the fundamental frequency measured for each word.

Table 3.1: Min-max values for measured fundamental frequency (F0).

	min F0 (Hz) for without NP		min F0 (Hz) for with NP	
	min	max	min	max
mmm	102.05	134.03	72.64	111.45
nnnn	103.35	134.91	85.90	123.65
engin	101.80	132.22	76.69	121.32
cengiz	91.10	132.38	85.69	125.90
kangal	91.33	132.57	59.06	130.56
mangal	97.65	138.94	70.20	121.56

Table 3.2: Min-max values for Ax graph mean value.

	Min	Max
Loudness (dB) for without NP	55.82	87.21
Loudness (dB) for with NP	62.59	91.79

By using the mean and standard deviation values that we obtained, we've prepared .arff file to make classification in WEKA.

Weka Arff Codes

@relation 'F0&Ax1'

@attribute 'mmmm' numeric

@attribute 'nnnn' numeric

@attribute 'engin' numeric

@attribute 'cengiz' numeric

@attribute 'kangal' numeric

@attribute 'mangal' numeric

@attribute 'Ax1_Mean' numeric

@attribute class {'polipli','saglikli'}

@data

72.64,91.74,84.46,85.69,77.83,70.20,91.65,'polipli'

104.69,100.20,99.07,94.26,81.87,84.42,66.60,'polipli'

80.95,91.38,80.07,95.07,92.74,79.57,62.59,'polipli'

122.85,117.42,111.04,92.40,110.98,103.20,82.16,'saglikli'

121.88,116.54,108.71,101.70,115.40,106.21,66.03,'saglikli'

128.71,118.72,112.27,91.10,95.97,97.65,58.52,'saglikli'

111.45,101.76,121.32,125.90,130.56,109.25,85.83,'polipli'

115.17,124.95,109.45,97.43,123.75,108.30,69.85,'saglikli'

105.20,121.70,113.50,96.80,91.33,99.70,71.45,'saglikli'

86.80,112.77,82.17,103.69,102.07,104.37,87.46,'polipli'

110.71,106.25,114.65,104.86,116.37,106.17,91.41,'polipli'

97.29,123.65,96.62,113.90,107.07,95.96,67.72,'polipli'

128.70,127.10,115.40,102.10,97.61,126.45,71.54,'saglikli'

95.87,99.43,103.43,95.90,103.98,119.99,87.48,'polipli'

124.95,116.04,107.53,109.74,111.48,105.57,87.21,'saglikli'

.

.

.

.

.

.

Classification models in WEKA were established using the speech recordings of 20 patients with NP and 20 without NP individuals. The results of classification model are tabulated in from Tables 3.3 and 3.4 as shown below:

Table 3.3: The success of classification model for N-fold cross validation (a) and Leave-One-Out Cross-Validation (LOOCV) (b).

Using Algorithm	Test mode: N-fold cross-validation, N=10						
	TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
Multilayer Perceptron	0.85	0.15	0.85	0.85	0.85	0.913	With NP
	0.85	0.15	0.85	0.85	0.85	0.913	Without NP
	0.85	0.15	0.85	0.85	0.85	0.913	Average
Spegasos	0.8	0.15	0.842	0.8	0.821	0.825	With NP
	0.85	0.2	0.81	0.85	0.829	0.825	Without NP
	0.825	0.175	0.826	0.825	0.825	0.825	Average
Naive Bayes	0.9	0.1	0.9	0.9	0.9	0.953	With NP
	0.9	0.1	0.9	0.9	0.9	0.953	Without NP
	0.9	0.1	0.9	0.9	0.9	0.953	Average
Logistic	0.85	0.15	0.85	0.85	0.85	0.894	With NP
	0.85	0.15	0.85	0.85	0.85	0.893	Without NP
	0.85	0.15	0.85	0.85	0.85	0.893	Average
RBF Network	0.9	0.15	0.857	0.9	0.878	0.92	With NP
	0.85	0.1	0.895	0.85	0.872	0.918	Without NP
	0.875	0.125	0.876	0.875	0.875	0.919	Average
Simple Logistic	0.9	0.1	0.9	0.9	0.9	0.955	With NP
	0.9	0.1	0.9	0.9	0.9	0.955	Without NP
	0.9	0.1	0.9	0.9	0.9	0.955	Average
SMO	0.9	0.15	0.857	0.9	0.878	0.875	With NP
	0.85	0.1	0.895	0.85	0.872	0.875	Without NP
	0.875	0.125	0.876	0.875	0.875	0.875	Average
J48	0.9	0.1	0.9	0.9	0.9	0.898	With NP
	0.9	0.1	0.9	0.9	0.9	0.898	Without NP
	0.9	0.1	0.9	0.9	0.9	0.898	Average

(a)

Using Algorithm	Test mode: Leave-One-Out Cross-Validation (LOOCV)						
	TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
Multilayer Perceptron	0.8	0.1	0.889	0.8	0.842	0.905	With NP
	0.9	0.2	0.818	0.9	0.857	0.905	Without NP
	0.85	0.15	0.854	0.85	0.85	0.905	Average
Spegasos	0.85	0.15	0.85	0.85	0.85	0.85	With NP
	0.85	0.15	0.85	0.85	0.85	0.85	Without NP
	0.85	0.15	0.85	0.85	0.85	0.85	Average
Naive Bayes	0.85	0.05	0.944	0.85	0.895	0.958	With NP
	0.95	0.15	0.864	0.95	0.905	0.958	Without NP
	0.9	0.1	0.904	0.9	0.9	0.957	Average
Logistic	0.85	0.15	0.85	0.85	0.85	0.903	With NP
	0.85	0.15	0.85	0.85	0.85	0.905	Without NP
	0.85	0.15	0.85	0.85	0.85	0.904	Average
RBF Network	0.95	0.1	0.905	0.95	0.927	0.913	With NP
	0.9	0.05	0.947	0.9	0.923	0.913	Without NP
	0.925	0.075	0.926	0.925	0.925	0.913	Average
Simple Logistic	0.95	0.1	0.905	0.95	0.927	0.975	With NP
	0.9	0.05	0.947	0.9	0.923	0.975	Without NP
	0.925	0.075	0.926	0.925	0.925	0.975	Average
SMO	0.9	0.1	0.9	0.9	0.9	0.9	With NP
	0.9	0.1	0.9	0.9	0.9	0.9	Without NP
	0.9	0.1	0.9	0.9	0.9	0.9	Average
J48	0.9	0.1	0.9	0.9	0.9	0.903	With NP
	0.9	0.1	0.9	0.9	0.9	0.903	Without NP
	0.9	0.1	0.9	0.9	0.9	0.903	Average

(b)

Table 3.3 gives performance values (TP rate, FP rate, Precision, Recall, F-Measure and ROC values) for both training methods (N-Fold Cross-Validation and Leave-One-Out Cross-Validation-LOOCV) used with the classification models (MLP, Spegasos, NB, Logistic, RBF Network, Simple Logistic, SMO, J48).

Table 3.4: The classification model results in detail.

Using Algorithm	N-fold cross-validation (N=10)	
	Correctly Classified Instances	Incorrectly Classified Instances
Multilayer Perceptron	85%	15%
Spegasos	82.5%	17.5%
Naive Bayes	90%	10%
Logistic	85%	15%
RBF Network	87.5%	12.5%
Simple Logistic	90%	10%
SMO	87.5%	12.5%
J48	90%	10%

(a)

Using Algorithm	Leave-One-Out Cross-Validation (LOOCV)	
	Correctly Classified Instances	Incorrectly Classified Instances
Multilayer Perceptron	85%	15%
Spegasos	85%	15%
Naive Bayes	90%	10%
Logistic	85%	15%
RBF Network	92.5%	7.5%
Simple Logistic	92.5%	7.5%
SMO	90%	10%
J48	90%	10%

(b)

Table 3.4 shows the success rate of the classification models for correct classification. When the results were examined, it was seen that the most succesful method was the model in which RBF Network and Simple Logistic classification algorithm and Leave-One-Out Cross-Validation (LOOCV) training method were applied to dataset together. In these models, 92.5% classification success was achieved and 37 out of 40 data were successfully detected.

At the end of this classification process, ROC (Receiver Operating Characteristic) curves were drawn to show the classification performance obtained.

The ROC curve has a False Positive Rate on the x-axis and a True Positive Rate on the y-axis. The larger the area under the curve in the ROC curves, the better the classification done. Figures 3.8 and 3.9 show the RBF Network and Simple Logistic classification algorithm and the ROC curve for the Leave-One-Out Cross-Validation (LOOCV) training method with the highest percentage success rate.

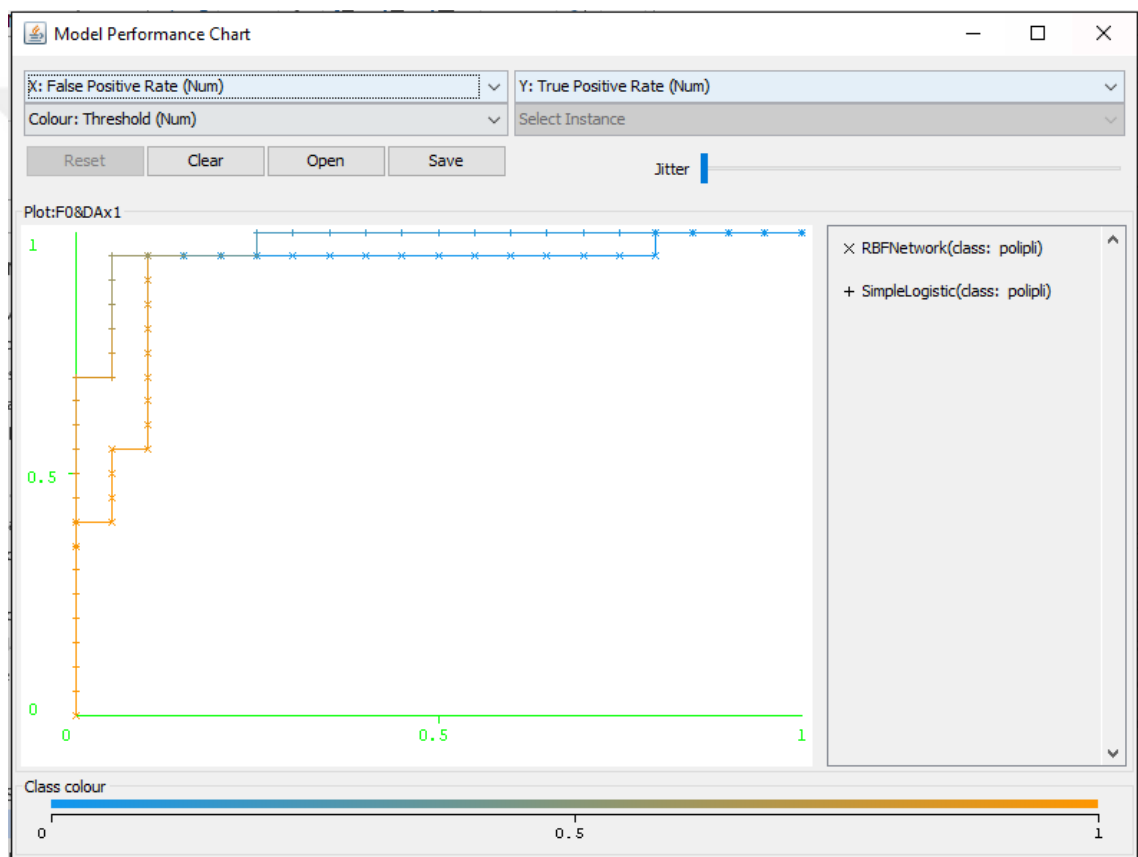


Figure 3.8: ROC Curve, RBF Network and Simple Logistic Algorithms, Leave-One-Out Cross-Validation (LOOCV) (Class value= with NP).

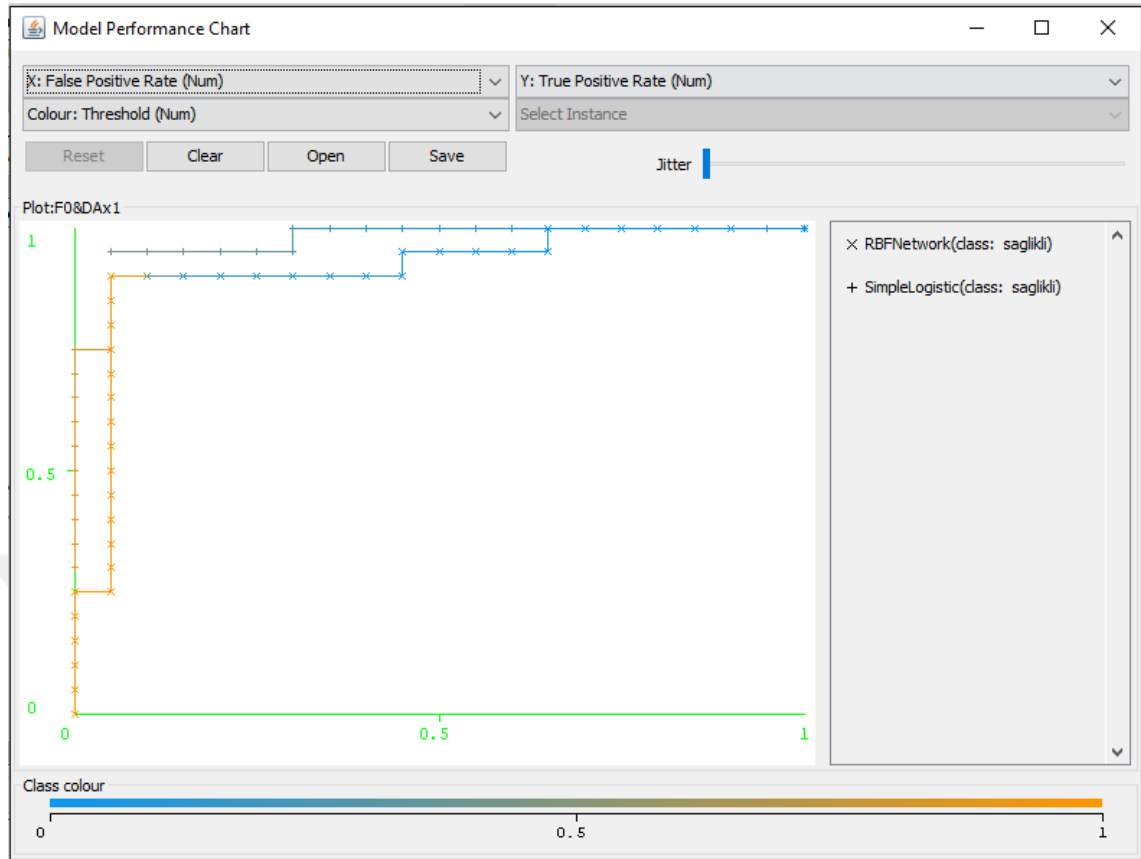


Figure 3.9: ROC Curve, RBF Network and Simple Logistic Algorithms, Leave-One-Out Cross-Validation (LOOCV) (Class value= without NP).

3.3. NASAL REGION ANALYSIS (SPE AND NX/SP RATIO GRAPH)

For the data set given the information in the section 3.1, the nasal region analysis was made. We separate SPE (Speech Pattern Elements) and Nx/Sp Ratio graph from 6 graphics that we obtained above which we will use in this study and we have achieved the mathematical values of these graphs through Speech Studio Software.

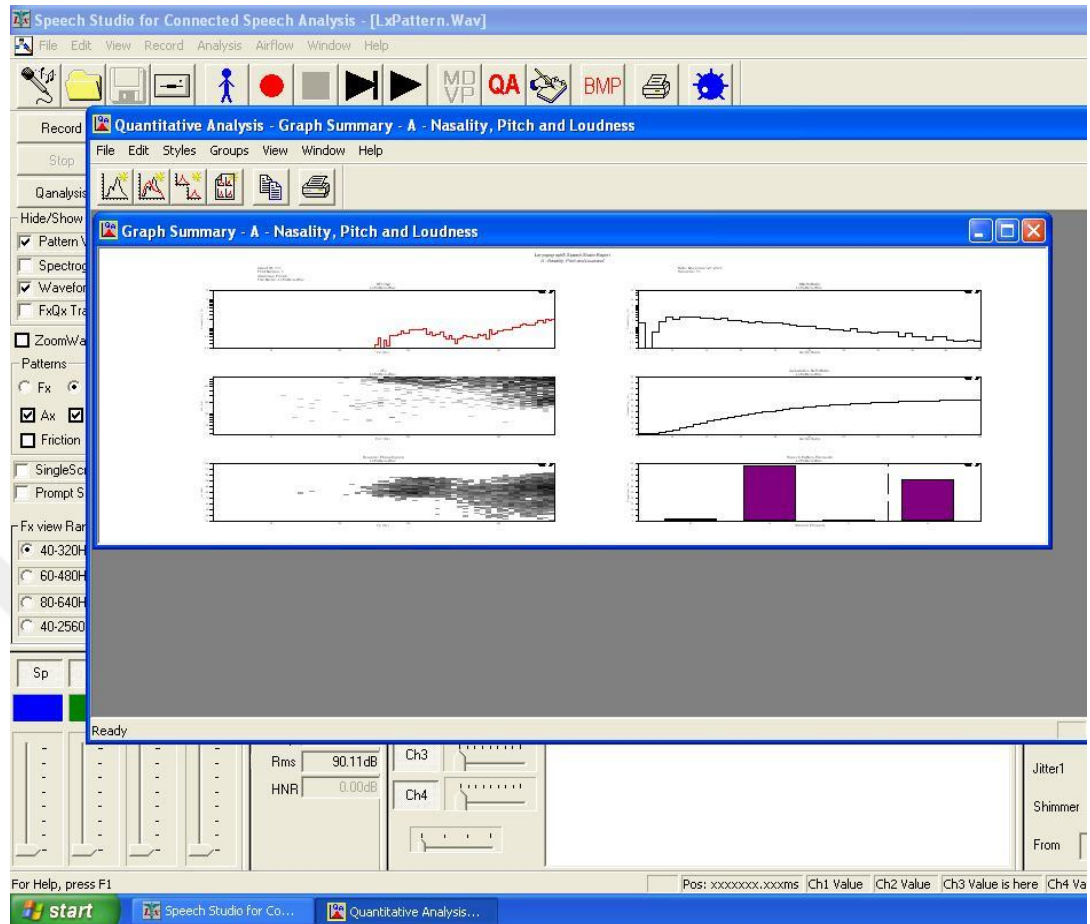


Figure 3.10: The graphs obtained after QA.

We've given two sample of SPE and Nx/Sp Ratio chart which we will use for classification below. These sample graphics are belong to those with nasal polyps and healthy individuals.

The Speech Pattern Element (figure 3.11 and 3.12) display is an indicator of fluency. It shows, as a bar graph, the proportion of time in the recording spent in silence or non-voice (Sx), voice (Vx), frication (Fr) and nasalance (Nx) if enabled [82].

As Nx values in SPE chart contains a description of the formation of the nasal sound, we use the classification as a Feature Extraction.

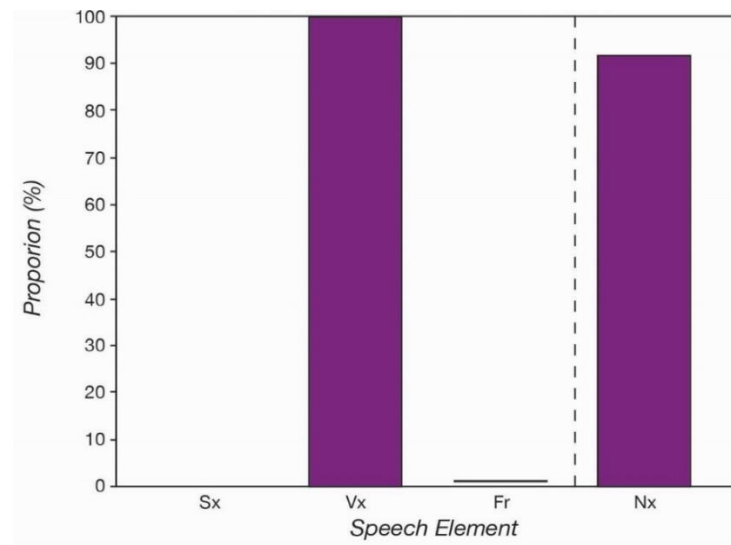


Figure 3.11: Speech Pattern Elements, Individual with Nasal Polyps.

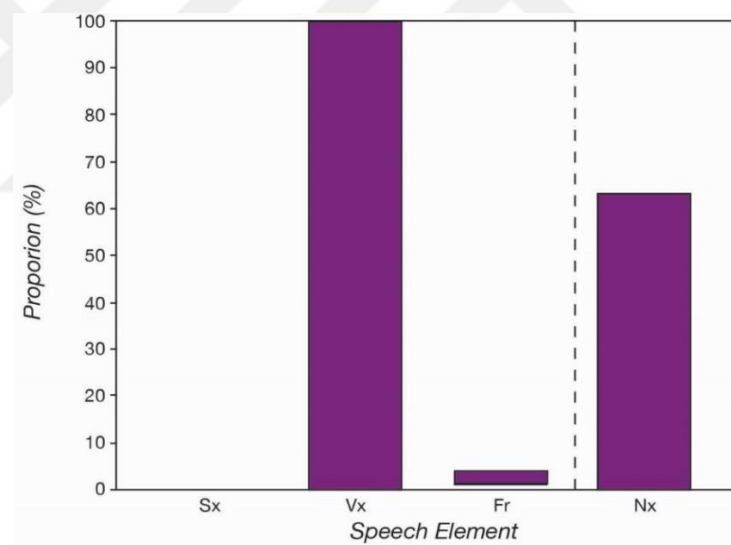


Figure 3.12: Speech Pattern Elements, Individual without Nasal Polyps.

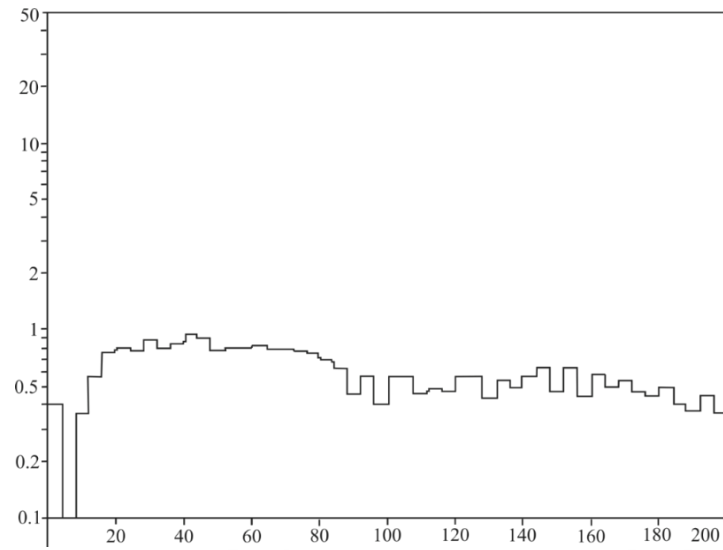


Figure 3.13: Nx/Sp Ratio, Individual with Nasal Polyps.

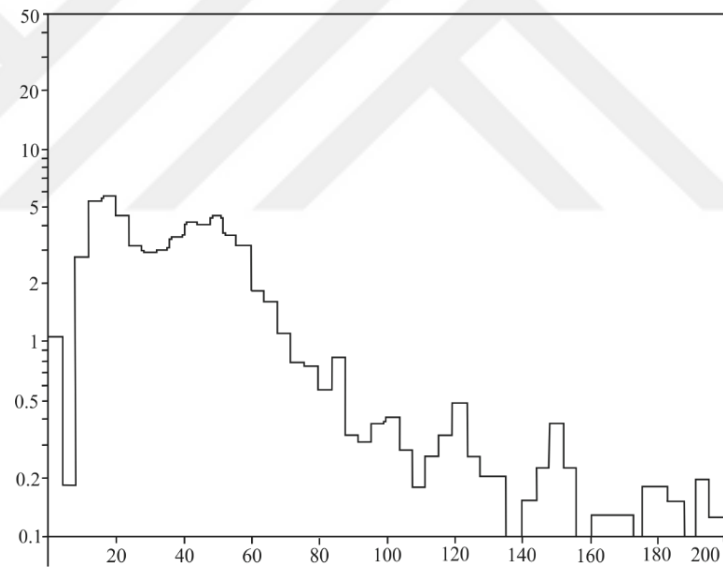


Figure 3.14: Nx/Sp Ratio, Individual without Nasal Polyps.

The device had a calibrated Nx/Sp ratio for charting. For calibration, a reference value of a true nasal sound was used and this value was formulated as $NxSpRef = Nx_{ref}/Sp_{ref}$. For every laryngeal cycle, a measured NxSp ($=Nx/Sp$) was also obtained. The NxSpRatio of the graphs (x-axes) was calculated by the formula $Nx/SpRatio (\%) = 100 * NxSp / NxSpRef$. The x-axes of the graphs were limited to 200 since sometimes the nasal vibration was stronger than that of the calibration sound [78].

After Quantitative Analyses we obtained the graphs as Nx/Sp ratio and SPE. From Nx/Sp ratio graph we obtained mod, median, mean, standard deviation, and from SPE graph we obtained Nx/Vx percentage values (mathematical values) which we calculated by using Speech Studio Software (Feature Extraction).

It was observed that the basic distinguishing parameters for individual with and without NP were mean, standard deviation, and Nx/Vx percentage (Feature Selection).

Table 3.5: Statistical values of the Individuals with NP.

	<i>Mean</i>	<i>Std. Dev.</i>	<i>SPE-Nx Value</i>	<i>Age</i>	<i>Sex</i>
<i>Patient 1</i>	67,6	51,15	88	49	M
<i>Patient 2</i>	76,61	51,38	88	44	M
<i>Patient 3</i>	57,49	44,88	73	32	M
<i>Patient 4</i>	69,94	47,11	88	42	M
<i>Patient 5</i>	92,67	61,33	96	29	M
<i>Patient 6</i>	88,52	47,33	97	33	M
<i>Patient 7</i>	83,6	46,36	88	37	M
<i>Patient 8</i>	60,27	42,06	79	25	M
<i>Patient 9</i>	111,88	55,3	97	26	M
<i>Patient 10</i>	50,77	38,13	75	35	M
<i>Patient 11</i>	65,43	51,31	77	45	M
<i>Patient 12</i>	67,7	51,88	78	43	M
<i>Patient 13</i>	58,13	45,6	83	54	M
<i>Patient 14</i>	79,67	53,11	69	40	M
<i>Patient 15</i>	74,27	49,5	90	49	M
<i>Patient 16</i>	76,6	39,1	93	53	M
<i>Patient 17</i>	93,45	60,08	71	51	M
<i>Patient 18</i>	97,53	58,43	65	44	M
<i>Patient 19</i>	68,25	50,35	89	43	M
<i>Patient 20</i>	59,9	44,63	84	47	M

Table 3.6: Statistical values of the Individuals without NP.

	Mean	Std. Dev.	SPE-Nx Value	Age	Sex
<i>Person 1</i>	47,46	33,45	68	49	M
<i>Person 2</i>	52,92	40,68	73	44	M
<i>Person 3</i>	57,73	47,19	69	32	M
<i>Person 4</i>	65,53	46,11	69	42	M
<i>Person 5</i>	45,23	34,03	68	29	M
<i>Person 6</i>	46,36	38,14	70	33	M
<i>Person 7</i>	35,71	34,07	49	37	M
<i>Person 8</i>	52,11	46,5	78	25	M
<i>Person 9</i>	92,95	51,35	78	26	M
<i>Person 10</i>	42,43	35,22	63	35	M
<i>Person 11</i>	38,11	38,03	51	45	M
<i>Person 12</i>	43,12	39,12	63	43	M
<i>Person 13</i>	50,75	43,71	55	54	M
<i>Person 14</i>	45,88	41,29	67	40	M
<i>Person 15</i>	38,93	37,01	59	49	M
<i>Person 16</i>	60,11	49,35	65	53	M
<i>Person 17</i>	58,56	44,3	71	51	M
<i>Person 18</i>	41,4	37,74	51	44	M
<i>Person 19</i>	47,18	35,3	54	43	M
<i>Person 20</i>	70,01	48,66	65	47	M

By using the mean and standard deviation values that we obtained, we've prepared .arff file to make classification in WEKA.

Weka Arff Codes

```
@relation 'NxSp_Ratio_and_Nx'
@attribute 'mean' numeric
@attribute 'std_dev' numeric
@attribute 'SPE-Nx' numeric
@attribute 'class' {'polipli','saglikli'}
@data
67.60,51.15,88,'polipli'
76.61,51.38,88,'polipli'
57.49,44.88,73,'polipli'
47.46,33.45,68,'saglikli'
52.92,40.68,73,'saglikli'
57.73,47.19,69,'saglikli'
69.94,47.11,88,'polipli'
65.53,46.11,69,'saglikli'
45.23,34.03,68,'saglikli'
92.67,61.33,96,'polipli'
88.52,47.33,97,'polipli'
83.60,46.36,88,'polipli'.
.
.
.
.
.
.
.
.
.
.
```

Classification models in WEKA were established using the speech recordings of 20 patients with NP and 20 without NP individuals. The results of classification model are tabulated in from Tables 3.7 and 3.8 as shown below:

Table 3.7: The success of classification model for N-fold cross validation (a) and Leave-One-Out Cross-Validation (LOOCV) (b).

Using Algorithm	Test mode: N-fold cross-validation, N=10						
	TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
Multilayer Perceptron	0.9	0.1	0.9	0.9	0.9	0.923	With NP
	0.9	0.1	0.9	0.9	0.9	0.923	Without NP
	0.9	0.1	0.9	0.9	0.9	0.923	Average
Spegasos	0.8	0.1	0.889	0.8	0.842	0.85	With NP
	0.9	0.2	0.818	0.9	0.857	0.85	Without NP
	0.85	0.15	0.854	0.85	0.85	0.85	Average
Naive Bayes	0.9	0.05	0.947	0.9	0.923	0.925	With NP
	0.95	0.1	0.905	0.95	0.927	0.925	Without NP
	0.925	0.075	0.926	0.925	0.925	0.925	Average
Logistic	0.8	0.1	0.889	0.8	0.842	0.881	With NP
	0.9	0.2	0.818	0.9	0.857	0.883	Without NP
	0.85	0.15	0.854	0.85	0.85	0.882	Average
RBF Network	0.85	0.2	0.81	0.85	0.829	0.783	With NP
	0.8	0.15	0.842	0.8	0.821	0.783	Without NP
	0.825	0.175	0.826	0.825	0.825	0.783	Average
Simple Logistic	0.85	0.1	0.895	0.85	0.872	0.92	With NP
	0.9	0.15	0.857	0.9	0.878	0.92	Without NP
	0.875	0.125	0.876	0.875	0.875	0.92	Average
SMO	0.9	0.1	0.9	0.9	0.9	0.9	With NP
	0.9	0.1	0.9	0.9	0.9	0.9	Without NP
	0.9	0.1	0.9	0.9	0.9	0.9	Average
J48	0.85	0.05	0.944	0.85	0.895	0.946	With NP
	0.95	0.15	0.864	0.95	0.905	0.946	Without NP
	0.9	0.1	0.904	0.9	0.9	0.946	Average

(a)

Using Algorithm	Test mode: Leave-One-Out Cross-Validation (LOOCV)						
	TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
Multilayer Perceptron	0.9	0.15	0.857	0.9	0.878	0.898	With NP
	0.85	0.1	0.895	0.85	0.872	0.898	Without NP
	0.875	0.125	0.876	0.875	0.875	0.898	Average
Spegasos	0.75	0.1	0.882	0.75	0.811	0.825	With NP
	0.9	0.25	0.783	0.9	0.837	0.825	Without NP
	0.825	0.175	0.832	0.825	0.824	0.825	Average
Naive Bayes	0.85	0.1	0.895	0.85	0.872	0.92	With NP
	0.9	0.15	0.857	0.9	0.878	0.92	Without NP
	0.875	0.125	0.876	0.875	0.875	0.92	Average
Logistic	0.8	0.1	0.889	0.8	0.842	0.898	With NP
	0.9	0.2	0.818	0.9	0.857	0.898	Without NP
	0.85	0.15	0.854	0.85	0.85	0.898	Average
RBF Network	0.85	0.1	0.895	0.85	0.872	0.878	With NP
	0.9	0.15	0.857	0.9	0.878	0.878	Without NP
	0.875	0.125	0.876	0.875	0.875	0.877	Average
Simple Logistic	0.75	0.1	0.882	0.75	0.811	0.9	With NP
	0.9	0.25	0.783	0.9	0.837	0.9	Without NP
	0.825	0.175	0.832	0.825	0.824	0.9	Average
SMO	0.9	0.15	0.857	0.9	0.878	0.875	With NP
	0.85	0.1	0.895	0.85	0.872	0.875	Without NP
	0.875	0.125	0.876	0.875	0.875	0.875	Average
J48	0.85	0.2	0.81	0.85	0.829	0.97	With NP
	0.8	0.15	0.842	0.8	0.821	0.97	Without NP
	0.825	0.175	0.826	0.825	0.825	0.97	Average

(b)

Table 3.7 gives performance values (TP rate, FP rate, Precision, Recall, F-Measure and ROC values) for both training methods (N-Fold Cross-Validation and Leave-One-Out Cross-Validation-LOOCV) used with the classification models (MLP, Spegasos, NB, Logistic, RBF Network, Simple Logistic, SMO, J48).

Table 3.8: The classification model results in detail.

Using Algorithm	N-fold cross-validation (N=10)	
	Correctly Classified Instances	Incorrectly Classified Instances
Multilayer Perceptron	90%	10%
Spegasos	85%	15%
Naive Bayes	92.5%	7.5%
Logistic	85%	15%
RBF Network	82.5%	17.5%
Simple Logistic	87.5%	12.5%
SMO	90%	10%
J48	90%	10%

(a)

Using Algorithm	Leave-One-Out Cross-Validation (LOOCV)	
	Correctly Classified Instances	Incorrectly Classified Instances
Multilayer Perceptron	87.5%	12.5%
Spegasos	82.5%	17.5%
Naive Bayes	87.5%	12.5%
Logistic	85%	15%
RBF Network	87.5%	12.5%
Simple Logistic	82.5%	17.5%
SMO	87.5%	12.5%
J48	82.5%	17.5%

(b)

Table 3.8 shows the success rate of the classification models for correct classification. When the results were examined, it was seen that the most successful method was the model in which Naive Bayes classification algorithm and N-fold cross validation (while N=10) training method were applied to dataset together. In this model, 92.5% classification success was achieved and 37 out of 40 data were successfully detected.

At the end of this classification process, ROC (Receiver Operating Characteristic) curves were drawn to show the classification performance obtained.

The ROC curve has a False Positive Rate on the x-axis and a True Positive Rate on the y-axis. The larger the area under the curve in the ROC curves, the better the classification done. Figures 3.15 and 3.16 show the ROC curve for the Naive Bayes and N-fold cross validation (N=10) method with the highest percentage success rate.

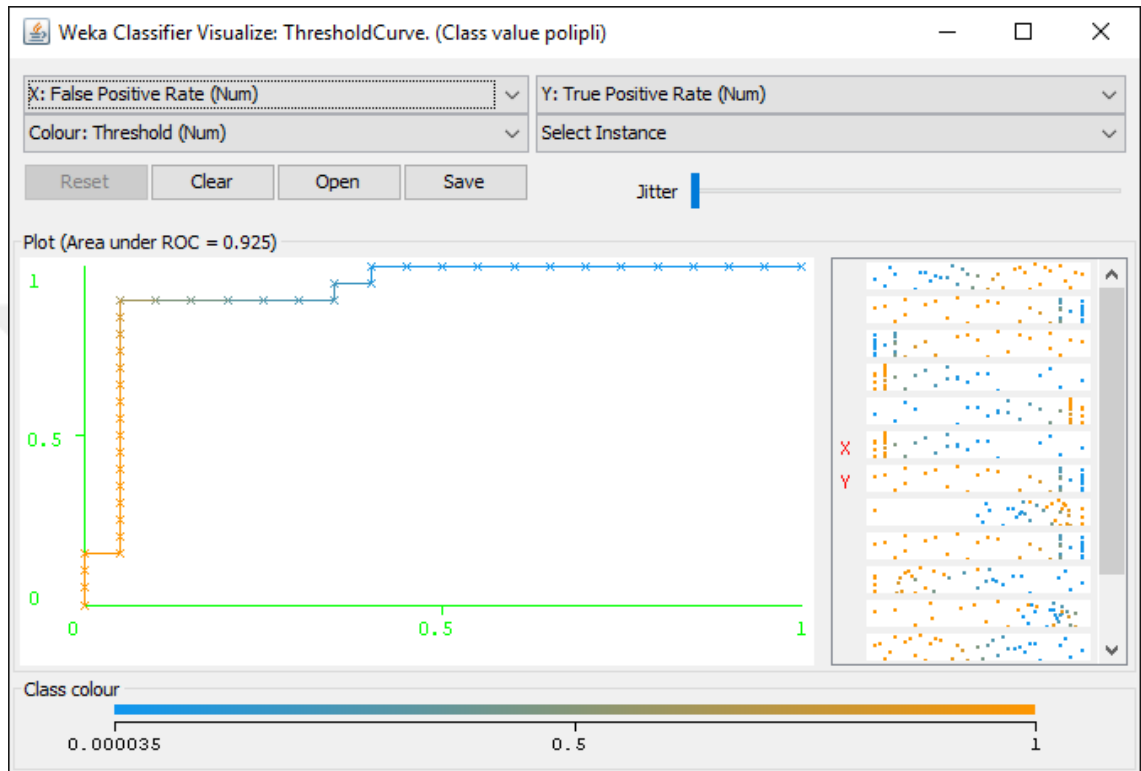


Figure 3.15: ROC Curve, Naive Bayes and N-fold cross validation

(N=10, Class value= with NP).

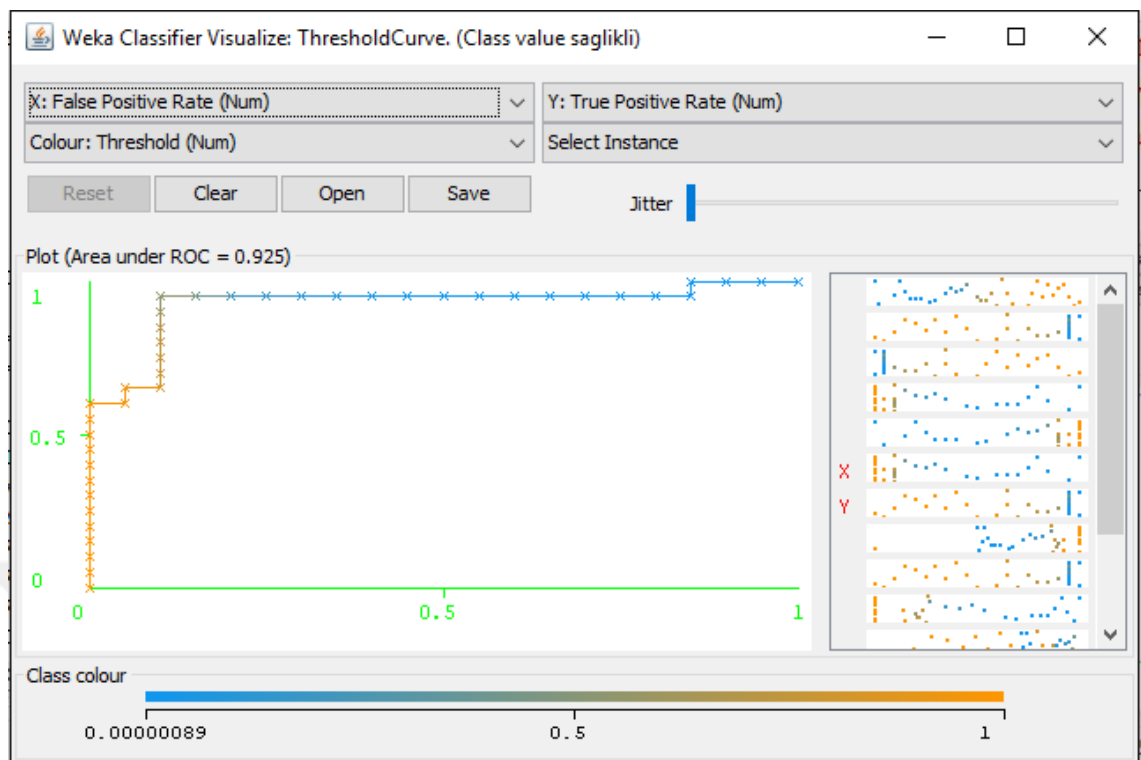


Figure 3.16: ROC Curve, Naive Bayes and N-fold cross validation
(N=10, Class value=without NP).

4. DISCUSSION

The main purpose of this study is to obtain the voice recordings of patients with the Nasal Polyp (NP) and to analyze the changes in voice formation through these recordings. Sound occurs when the air from the lungs passes through the larynx region (source region) where the vocal cords are located and filtered in the resonant region of the mouth, nasal cavities, and paranasal sinuses. In this study which is also referred to as the source-filter model, It is aimed to reveal the effects of the nasal polyps which causes the change in the filter region. Sound recordings obtained for this purpose were classified by using mathematical properties with classification algorithms and the results were interpreted.

Today, nasal polyps are known to have affected speech due to the changes which it caused in the resonance region [3, 4]. The medical effects of nasal polyps have been the focus of many studies [5]. In a study conducted on children aged 4-12, the effect of barriers in the nasal region on the sound quality has been examined and nasal polyps as one of these barriers have been discussed. As a result, it has been revealed that the barriers in the nasal region affect the sound quality [6]. In another study, nasometer measures and audio spectrum of patients with nasal polyps have been studied [7]. In addition, there are studies examining the effects changes in nasal resonance on sound [8, 9]. To sum up, the general view we encountered in the literature review is that there are effects of nasal polyps on human life or the effects of changes in nasal resonance on sound.

As the fundamental frequency includes the features of the person's speech, it's been used in many academic studies. In one study, fundamental frequency of men and woman compared and analyzed. In one study, the voice sensitivity of 18-45 males and females were examined by the basic frequency data [14]. In another study, fundamental frequency is used to analyse and classify mutational dysphonia disease [15]. Also for the gender perception and male-female voice comparison, fundamental frequency has been used in some studies [16, 17].

As shown in Figure 4.1, speech is a complex system of many subcomponents. As explained in the relevant sections of the thesis, these components have contributed separately to the sound formation. Therefore, the failure which will occur in the sub-components of the sound will affect the sound. The main purpose of our thesis is to examine the effect of the fault occurred in a part of the speech complex system on other parts of the system. With this motivation, we tried to reveal the effects of nasal polyposis speech on various measurements and analyzes.

In this study, normal sounds of the volunteers were recorded with a microphone. In addition, EGG signals (fundamental frequency) from the glottis part of the larynx region, where the vocal tracts are located, were recorded with an Electroglottograph. The nasal vibrations at the time of speech were recorded with a vibration transducer attached on the nose. In the study, volunteers were told to say words containing nasal sounds (m, n). Fundamental frequency (F0) for each word were calculated. Loudness for the whole recording and nasal vibrations with a vibration transducer attached on the nose were determined as percentage.

As a result, two different analyses were made using the obtained data. First, it was revealed if nasal polyps affect source region and thus fundamental frequency in sound formation by using fundamental frequency and loudness. Second, the effects of nasal polyps on speech were revealed assessing the whole speech from the obtained nasal vibration data.

In our first analysis in the study, we tried to reveal the effects of nasal polyps on the fundamental frequency and source region. The approach in the literature regarding the source-filter model is as follows:

The idea of the air from the lungs as a source and the vocal tract as a filter is called the source-filter model for sound production. Goldberg and Riek describe the process of designing speech formation as a generally concept of modelling any type of sound wave. “Sound waves are pressure variations that propagate through air (or any other medium) by the vibrations of the air particles. Modelling these waves and their propagation through the vocal tract provides a framework for characterizing how the vocal tract shapes the frequency content of the excitation signal” [84].

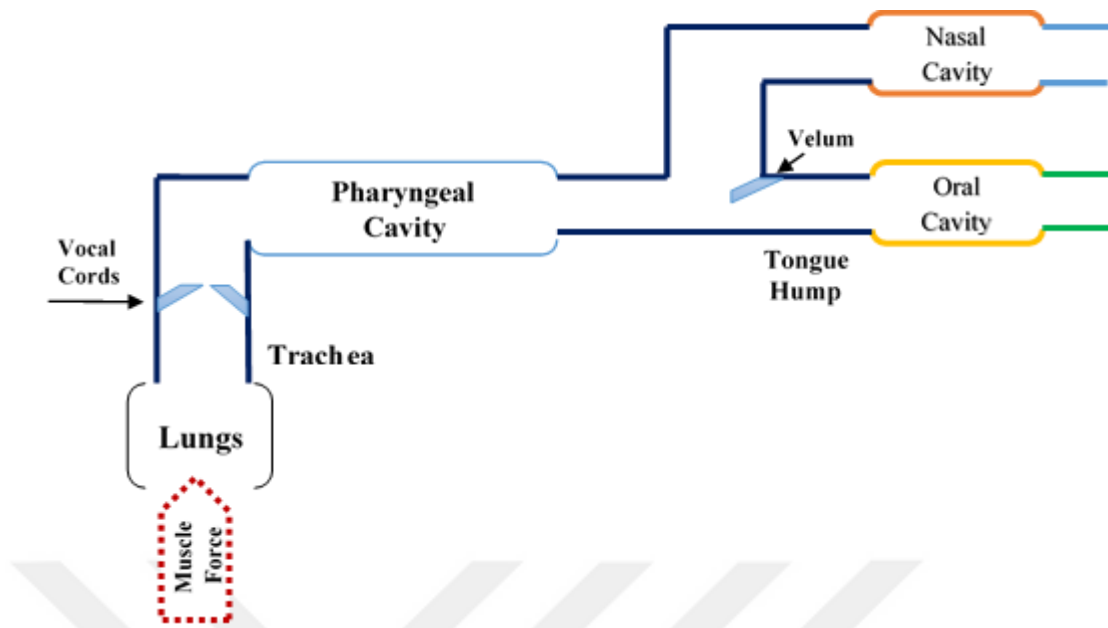


Figure 4.1: The formation of speech as a complete system.

Today, nasal polyps are known to have affected speech due to the changes which they caused in the resonance region [3]. In a similar paper, Acar investigated the changes in acoustic parameters by using multi-dimensional voice program in patients with nasal polyp after the surgery and found statistically significant changes only in shimmer values [85]. The medical effects of nasal polyps have been the focus of many studies [5]. In Figure 4.1, we consider the formation of speech as a complete system, a problem that occurs in Nasal Cavity, also shows would lead to changes in larynx. Thus, Owens states in his book *Signal Processing of Speech*: “Besides, the source-filter model suppose that the source is linearly separable from the filter and that there is no interaction between them. This opinion is not exactly correct, since the vibration of the vocal cords is affected by the sound pressure inside the vocal tract. There is coupling between the vocal tract and the lungs during the period when the glottis is open, thereby modifying the filter characteristics every cycle of the excitation” [48].

This opinion of Owens is seen as true. The results obtained indicate that, during the formation of the speech source region affects all kinds of changes in the filter region. The result of 92.50 % is quite high result.

In our study, when we evaluate the analyses we made with the data obtained from the vibrations detected with the help of a transducer attached on nose:

Nasal polyp is an unpleasant disease with tremendous adverse effects on the quality of life of patients. The diagnosis of the disease is easy; however, the high recurrence rate and insufficient elimination of the disease by treatment are the major problems for otorhinologists. Besides, the etiopathogenesis of the disease is not fully understood [86].

There are several studies investigating the effects of nasal pathologies on nasal resonance (nasality) by using nasometer [87, 7]. These studies revealed that sinuses have an important effect on voice resonance. However the effects of nasal pathologies on the larynx and the vocal folds are unclear.

In voice production vocal folds (vocal cords) are the primary sound source. Generally speaking, the mechanism for generating the human voice can be subdivided into three parts; the lungs (airflow), the vocal folds within the larynx (oscillator), and the articulators (the supraglottic larynx, the pharynx, the oral cavity the nasal cavity, the sinuses, and the head) Any disorder in this entire system may effect the quality of voice [48].

Acar investigated the changes in acoustic parameters by using multi-dimensional voice program in patients with nasal polyp after the surgery and found statistically significant changes only in shimmer values [85].

Our study is the first study analyzing the voice features in nasal polyp patients using EGG-D400 device. In this study we found statistically significant differences between the patients with nasal polyp and the control group.

The main problem in nasal polyp treatment is the high recurrence rate and it cannot be predicted when the polyp will reappear again. Besides, patients generally become symptomatic when the polyps go outside the boundaries of the sinuses, so it is important to make the early diagnosis of the NPs before they spread to the nasal cavity. If we can make the early diagnosis of the recurrence, intranasal glucocorticoids can be sufficient for the treatment, and, as a result, we can reduce the morbidity of systemic steroid treatment. The early diagnosis of the polyps confined in the sinuses can be done by paranasal CT, but this technique cannot be performed very often because of the radiation exposure, and it is also not cost-effective.

In this study, eight different classification methods successfully distinguishes the individuals with or without NP.

Voice analysis by the developed model can be an alternative method to detect the early recurrences of the nasal polyps but further studies with larger series are warranted in order to establish a new diagnostic method by using voice analysis in nasal polyp patients.



5. CONCLUSION AND RECOMMENDATIONS

Speech is a complex system of many subcomponents. Considering this system, first the effect of the source region on the nasal cavity region has been investigated. The results obtained indicate that this method is successful, showed that it supported Owens's suggestions and the changes in the nasal cavity region affected the source region. As a result of the applied method, the classification rate of 92.5%, has shown that this work is a reliable and successful method.

Then, the effects of changes in the nasal cavity region on speech have been examined. In the classification of data obtained in this context, 92.5% success rate was achieved. This successful result makes it possible to be used in the diagnosis of nasal polyposis to make a preliminary diagnosis by analyzing patients' speech.

As a result, the experimental results obtained in the thesis will contribute to the scientific work to be related to the speech hereinafter, especially research on fundamental frequency and nasal cavity.

REFERENCES

- [1]. Guner, L., 1999, *Türkçe konuşma ve konuşmacı tanıma* yönelik veri tabanı örnekleri, Thesis (M.Sc.), Ankara University.
- [2]. Nabiyev, V., 2010, *Yapay Zeka insan-bilgisayar etkileşimi*, 1st edition, Seçkin Publishing, Ankara.
- [3]. Demirsoy, S., Koç, C., Memiş, L., 2002, *Nazal Polipler*, 1st edition, Çizgi Tıp Publishing, Ankara.
- [4]. www.irishhealth.com/article.html?con=482 [Access date: 03 April 2014]
- [5]. Valerie, J.L., 1995, Diagnosis and treatment of nasal polyps, *British Medical Journal (BMJ)*, 311(7017), 1411–1414.
- [6]. Labio, R.B., Tavares, E.L.M., Alvarado, R.C., Martins, R.H.G., 2011, Consequences of Chronic Nasal Obstruction on the Laryngeal Mucosa and Voice Quality of 4- to 12-Year-Old Children, *Journal of Voice*, 26(4), 488-492.
- [7]. Hong, K.H., Kwon, S., Jung, S.S., 1997, The assessment of nasality with a nasometer and sound spectrography in patients with nasal polyposis, *Otolaryngol Head Neck Surg*, 117, 343–348.
- [8]. Jiang, R.S., Huang, H.T., 2006, Changes in nasal resonance after functional endoscopic sinus surgery, *Amerikan Journal of Rhinology*, 20(4), 432-437.
- [9]. Kumar, S., Basu, S., Sinha, A., Chatterjee, I., 2012, The Effect of Menstrual Cycle on Nasal Resonance Characteristics in Females, *Communications Disorders Quarterly(Online First)*, 1(1), 1-5.
- [10]. MUT, O., 2005, *Konuşmacı Tanıma*, Thesis (M.Sc.), GYTE.
- [11]. Harnsberger, J.D., Shrivastav, R., Brown, W.S., 2008, Speaking Rate and Fundamental Frequency as Speech Cues to Perceived Age, *Journal of Voice*, 22(1), 58-69.
- [12]. Ting, H.N., Chia, S.Y., Manap, H.H., Ho, A.H., Tiu, K.Y., Badruzaman, A., 2011, Fundamental Frequency and Perturbation Measures of Sustained Vowels in Malaysian Malay Children Between 7 and 12 Years Old, *Journal of Voice*, 26(4), 425-430.

- [13]. Lee, G.S., 2012, Variability in Voice Fundamental Frequency of Sustained Vowels in Speakers with Sensorineural Hearing Loss, *Journal of Voice*, 26(1), 24-29.
- [14]. Izadi, F., Mohseni, R., Daneshi, A., Sandughdar, N., 2012, Determination of Fundamental Frequency and Voice Intensity in Iranian Men and Women Aged Between 18 and 45 Years, *Journal of Voice*, 26(3), 336-340.
- [15]. Lim, J.Y., Lim, S.E., Choi, S.H., Kim, J.H., Kim, K.M., Choi, H.S., 2007, Clinical Characteristics and Voice Analysis of Patients with Mutational Dysphonia: Clinical Significance of Diplophonia and Closed Quotients, *Journal of Voice*, 21(1), 12-19.
- [16]. Gelfer, M.P., Bennett, Q.E., 2013, Speaking fundamental frequency and vowel formant frequencies: Effects on perception of gender, *Journal of Voice*, 27(5), 556-566.
- [17]. Ting, H.N., Chia, S.Y., Kim, K.S., Sim, S.L., Hamid, B.A., 2011, Vocal fundamental frequency and perturbation measurements of vowels by normal Malaysian Chinese adults, *Journal of Voice*, 25(6), e311-e317.
- [18]. Enderle, J.D., Bronzino, J.D., 2012, *Introduction to biomedical engineering*, 3rd edition, Elsevier Inc, Oxford, UK.
- [19]. Bronzino, J.D., 2006, *Biomedical Engineering Fundamentals*, 3rd edition, CRC Press Taylor & Francis Group, New York, USA.
- [20]. Cerutti, S., Marchesi, C., 2011, *Advanced Methods of Biomedical Signal Processing*, 1st edition, John Wiley & Sons Inc, New Jersey, USA.
- [21]. Pınarbaşı, M.O., 2009, *Nazal Polipli Hastalarda Ameliyat Öncesi Ve Sonrası Yaşam Kalitesi*, Thesis (M.D), Osmangazi University.
- [22]. Bachert, C., Gevaert, P., Holtappels, G., 2000, Nasal polyposis: From cytokines to growth, *Am J Rhinol*, 14, 279-290.
- [23]. Bateman, N.D., Fahy, C., Woolford, T.J., 2003, Nasal polyps: Still more questions than answers, *J Laryngol Otol*, 117(1), 1-9.
- [24]. Yuksel, A., 2006, *Yaygın Nazal Polipozis Hastalarında Biyofilm Oluşumunun Araştırılması*, Thesis (M.D), Gazi University.
- [25]. Yalçın, Ş., Keleş, E., 2007, *Nazal polipoziste tanı ve ayırıcı tanı*, Nazal Polipler, In: İleri, F. (ed.), Deomed Yayıncılık, İstanbul, 41-56.
- [26]. Yalçın, Ş., Keleş, E., 2007, *Nazal polipoziste medikal tedavinin rolü*, Nazal Polipler, In: İleri, F. (ed.), Deomed Yayıncılık, İstanbul, 121-129.

- [27]. Damm, M., 1999, Effects of systemic steroid treatment in chronic polypoid rhinosinusitis evaluated with magnetic resonance imaging, *Otolaryngol Head Neck Surg*, 120, 517-523.
- [28]. Van, C.C., Clement, P.A., 1994, Results of oral steroid treatment in nazal polyposis, *Rhinology*, 32, 5-9.
- [29]. Assanasen, P., 2001, Medical and surgical management of nazal polyps, *Otolaryngol Head Neck Surg*, 9, 27-36.
- [30]. Rasp, G., Schuck, A., Kastenbauer, E.R., 2000, A new grading system for nazal polyps of the ethmoid bone: Influence of a combined systemic and local steroid treatment, *Laryngorhinootologie*, 79, 266-272.
- [31]. Parnes, S.M., 2000, Acute effects of antileucotrienes on sinonasal polyposis and sinusitis, *Ear Nose Throat J*, 79, 18-20.
- [32]. Stammberger, H., Posawetz, W., 1990, Functional endoskopik sinus surgery. Concept, indications and results of the Messerklinger technique, *Eur Arch Otorhinolaryngol*, 247, 63-76.
- [33]. Onerci, M., 2006, *Diffüz nazal polipozis cerrahisi*, Nazal Polipozis, Hacettepe Üniversitesi Hastaneleri Basımevi, Ankara, 150-158.
- [34]. Sahin, E., 2012, *Esneme-İç Geçirme, Dudak Trili Ve Dil Rahatlatma Egzerszilerinin Ses Eğitiminde Kullanılabilirliği*, Thesis (PhD), Marmara University.
- [35]. Devge, C., Oguz, A., 1996, *Ses ve Ses Bozuklukları*, Ekin Tıbbi Yayıncılık, Ankara.
- [36]. Okten, İ., 1980, *Ultrases dalgalarının elektrolitlerde elektriksel iletkenliğe etkisi*, Thesis (PhD), Ege University.
- [37]. Ogut, F., Kalaycı, T., Uluoz, B., 1997, Ses analizinde son gelişmeler, 24. *Ulusal Türk Otorinolarengoloji ve Baş Boyun Cerrahisi Kongre Kitabı*.
- [38]. Kara, O.K., 2011, *Adli Amaçlı Ses Analizinde Otomatik Konuşmacı Tanıma Yazılımı Kullanılarak Kardeşler Arası Ses Benzerliklerinin Ve Farklılıklarının Belirlenmesi*, Thesis (M.Sc.), Istanbul University.
- [39]. Vennard, W., 1967, *Sarkı Söyleme Teknigi ve Mekanizması*, Bogaziçi Üniversitesi Yayınevi (Çeviri), Istanbul.
- [40]. Beninger, M.S., Murry, T., 2006, *The Singer's Voice*, Plural Publishing, UK.
- [41]. Akmes, P.P., 2008, *Çocuk Seslerinin Akustik Özelliklerinin Belirlenmesi*, Thesis (M.Sc.), Hacettepe University.
- [42]. <http://tr.wikipedia.org/wiki/Yutak> [Access date: 05 March 2015]

- [43]. Sarac, E.T., 2010, *Türkçe Fonemlerin ünlü-ünsüz kombinasyonlarında nazalite değerlendirmesi*, Thesis (M.Sc.), Hacettepe University.
- [44]. Omori, K., Kacker, A., Carroll, M., 1996, Singing Power Ratio: Quantitative Evaluation of Singing Voice Quality, *Journal of Voice*, 10(3), 228-235.
- [45]. Koca, O., Boyacı, Z., 1996, *Spectrogram ve Fonetik Analiz*, Ses ve Ses Hastalıkları, Ekin Tıbbi Yayıncılık, İstanbul.
- [46]. Jiang, J., Lin, E., Hanson, E.G., 2000, Vocal Fold Physiology, *The Otolaryngologic Clinics of North America*, 33(4):699-718.
- [47]. http://clas.mq.edu.au/speech/acoustics/frequency/source_filter.html [Access date: 15 March 2015]
- [48]. Owens, F.J., 1993, *Signal Processing of Speech*, 1st edition, The Macmillan Press Ltd, London.
- [49]. Koç, C., 2009, *Temel Rinoloji*, Güneş Tıp Kitabevleri, Ankara.
- [50]. Tor, G., 2004, *Mersin Ağzı Sözlüğü*, Türk Dilleri Araştırmalar Dizisi, Simurg Publishing, İstanbul.
- [51]. Lecluse, F.L.E., Brocaar, M.P., Verschurre, J., 1975, The electoglottography and its relation to glottal activity, *Fol Phoniatr*, 27, 215-24.
- [52]. Fourcin, A.J., Abberton, E., 1971, First applications of a new laryngograph, *Med Biol Illus*, 21, 172-82.
- [53]. <http://www.laryngograph.com/pdfdocs/paper3.pdf> [Access date: 14 May 2016]
- [54]. <http://www.lpl-aix.fr/~ghio/pedago-EggUK.ht> [Access date: 14 May 2016]
- [55]. <https://en.wikipedia.org/wiki/Electroglottograph> [Access date: 14 May 2016]
- [56]. Herbst, C.T., Fitch, W.T., Svec, J.G., 2010, Electroglottographic wavegrams: a technique for visualizing vocal fold dynamics noninvasively, *J Acoust Soc Am* 128, 3070-3078.
- [57]. Miller, D.G., 2008, *Resonance in Singing*, Inside View Press, USA.
- [58]. www.wevosys.com/products/lingwaves/lingwaves_electroglottography.html [Access date: 14 May 2016]
- [59]. Kazi, R.A., Prasad, V.M., Kanagalingam, J., et.al., 2007, Assessment of the Formant Frequencies in Normal and Laryngectomized Individuals Using Linear Predictive Coding, *Journal of Voice*, 21(6), 661–668.
- [60]. Holmes, G., Donkin, A., Witten, I.H., 1994, WEKA: a machine learning workbench, *2nd Australian and New Zealand Conference on Intelligent Information Systems*, 357 – 361.

- [61]. Jenitha, G., Vennila, V., 2014, Comparing the partitional and density based clustering algorithms by using WEKA tool, *2nd International Conference on Current Trends in Engineering and Technology (ICCTET)*, 328 – 331.
- [62]. Mallios, N., Papageorgiou, E., Samarinas, M., 2011, Comparison of Machine Learning Techniques Using the WEKA Environment for Prostate Cancer Therapy Plan, *20th IEEE International Workshops on Enabling Technologies: Infrastructure for Collaborative Enterprises (WETICE)*, 151 – 155.
- [63]. Turhal, U., Babur, S., Avci, C., Akbas, A., 2013, Performance Improvement for Diagnosis of Colon Cancer by Using Ensemble Classification Methods, *International Conference on Technological Advances in Electrical Electronics and Computer Engineering (TAECE)*, 271 – 275.
- [64]. Chokkalingam, S.P., Komathy, K., 2013, Comparison of Different Classifier in WEKA for Rheumatoid Arthritis, *International Conference on Human Computer Interactions (ICHCI)*, 1 – 6.
- [65]. Liu, H., Setiono, R.A., 1996, Probabilistic Approach to Feature Selection: A Filter Solution, *13th International Conference on Machine Learning*, 319–327.
- [66]. Bishop, C.M., 1995, *Neural Networks for Pattern Recognition*, Oxford University Press, USA.
- [67]. John, G.H., Langley, P., 1995, Estimating Continuous Distribution Bayesian Classifiers, *11th Conference on Uncertainty in Artificial Intelligence*, 338–345.
- [68]. Breiman, L., 2001, Random Forest, *Machine Learning*, 45(1), 5–32.
- [69]. Fan, R.E., Chang, K.W., Hsieh, C.J., Wang, X.R., Lin, C.J., 2008, LIBLINEAR: A Library for Large Linear Classification, *Journal of Machine Learning*, 9, 1871–1874.
- [70]. Witten, I.H., Frank, E., 2011, *Data Mining: Practical Machine Learning Tools and Techniques*, Elsevier Inc, USA.
- [71]. Tufan, K., 2011, *Noninvasive Detection Of Stress Urinary Incontinence Using Doppler Ultrasound*, Thesis (PhD), Fatih University.
- [72]. <http://weka.sourceforge.net/doc.stable/weka/classifiers/functions/SPegasos.html> [Access date: 14 May 2016]
- [73]. Lakoumentas, J., Drakos, J., Karakantza, M., Sakellaropoulos, G., Megalooikonomou, V., Nikiforidis, G., 2012, Optimizations of the naive-bayes classifier for the prognosis of b-chronic lymphocytic leukemia incorporating flow cytometry data, *Computer Methods and Programs in Biomedicine*, 108 (1), 158–167.

- [74]. Wiggins, M., Saad, A., Litt, B., Vachtsevanos, G., 2007, Evolving a bayesian classifier for ecg-based age classification in medical applications, *Applied Soft Computing*, 8 (1), 599–608.
- [75]. Wang, Z., Hope, RM., Wang, Z., Ji, Q., Gray, WD., 2011, An EEG workload classifier for multiple subjects, *33rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 1-4.
- [76]. Wajs, W., Stoch, P., & Kruczek, P., 2007, Radial Basis Networks and Logistic Regression Method for Prediction of Broncho Pulmonary Dysplasia, *In Intelligent Systems Design and Applications Seventh International Conference on IEEE*, 551-555.
- [77]. <http://www.biostathandbook.com/simplelogistic.html> [Access date: 02 Feb 2017]
- [78]. https://en.wikipedia.org/wiki/Sequential_minimal_optimization [Access date: 03 Feb 2017]
- [79]. Muralidharan, V., Sugumaran, V., 2013, Feature extraction using wavelets and classification through decision tree algorithm for fault diagnosis of mono-block centrifugal pump, *Measurement*, 46(1), 353-359.
- [80]. http://scikit-learn.org/stable/modules/cross_validation.html [Access date: 10 March 2016]
- [81]. Sokolova, M., Japkowicz, N., Szpakowicz, S., 2006, Beyond accuracy, F-score and ROC: a family of discriminant measures for performance evaluation, *In AI 2006: Advances in Artificial Intelligence*, 19th Australian Joint Conference on Artificial Intelligence, Hobart, Australia, December 4-8, 2006, Springer Berlin Heidelberg, 1015-1021.
- [82]. Hu, S., 2012, *Speech Studio User Manual*, Laryngograph Ltd, UK.
- [83]. Kara, O.K., 2011, *Adli amaçlı ses analizinde otomatik konuşmacı tanıma yazılımı kullanılarak kardeşler arası ses benzerliklerinin ve farklılıklarının belirlenmesi*, Thesis (M.Sc.), Istanbul University.
- [84]. Randy, G., Lance, R., 1999, *A practical handbook of speech coders*, CRC Press, USA.
- [85]. Acar, A., Cayonu, M., Ozman, M., Eryilmaz, A., 2014, Changes in acoustic parameters of voice after endoscopic sinus surgery in patients with nasal polyposis, *Indian J Otolaryngol Head Neck Surgery*, 66 (4), 381-385.
- [86]. Rajguru, R., 2014, Nasal Polyposis: Current Trends, *Indian J Otolaryngol Head Neck Surgery*, 66 (1), 16–21.
- [87]. Soneghet, R., Santos, R.P., Behlau, M., Habermann, W., Friedrich, G., Stammberger, H., 2002, Nasalance Change Safter Functional Endoscopic Sinus Surgery, *Journal of Voice*, 16(3), 392–397.

APPENDICES

APPENDIX A

BİLGİLENDİRİLMİŞ OLUR FORMU-1

Turgut Özal Üniversitesi Tıp Fakültesi KBB Kliniği olarak “**Nazal Polip hastalığının konuşma üzerine etkilerinin incelenmesi ve konuşma üzerinden nazal polip tanısı**” isimli bir araştırma yapmaktayız. Sizde nazal polip tespit edildikten sonra sizden Laryngograph microProcessor EGG-D400 cihazı ile ses kayıtları ve nazal kavite ölçümü alınacaktır. Daha sonra bu kayıtlar ile sağlıklı bireylerin kayıtları karşılaştırılacaktır. Sizin de bu araştırmaya katılmanızı öneriyoruz. Araştırmaya katılmakla yeni kullanılmaya başlanan bir ilaç sizin üzerinizde denenmeyecektir.

Araştırmaya katılmayı kabul ederseniz bu durumun size ek bir maliyeti olmayacaktır. Bu çalışmaya katılmanız için sizden herhangi bir ücret istenmeyecektir. Çalışmaya katıldığınız için size ek bir ödeme de yapılmayacaktır.

Sizin de bu araştırmaya katılmanızı öneriyoruz. Bu araştırmaya katılıp katılmamakta serbestsiniz. Çalışmaya katılım gönüllülük esasına dayalıdır. Kararınızdan önce araştırma hakkında sizi bilgilendirmek istiyoruz. Bu bilgileri okuyup anladıktan sonra araştırmaya katılmak isterseniz formu imzalayınız.

Bu çalışmaya katılmayı reddedebilirsiniz. Bu araştırmaya katılmak tamamen isteğe bağlıdır ve reddettiğiniz takdirde size uygulanan tedavide ya da bundan sonra kliniğimizde size karşı davranışlarımızda herhangi bir değişiklik olmayacaktır. Yine çalışmanın herhangi bir aşamasında çalışmadan onayınızı çekmek hakkına da sahipsiniz.

Tarih :

Doktor :

Adres :

Tel :

İmza :

BİLGİLENDİRİLMİŞ OLUR FORMU-2

Sayın Doç.Dr. Nebil ARK ve arkadaşları tarafından Turgut Özal Üniversitesi Tıp Fakültesi KBB Kliniğinde tıbbi bir araştırma yapılacağı belirtilerek bu araştırma ile ilgili yukarıdaki bilgiler bana aktarıldı. Bu bilgilerden sonra böyle bir araştırmaya “katılımcı” (denek) olarak davet edildim.

Eğer bu araştırmaya katılırsam hekim ile aramda kalması gereken bana ait bilgilerin gizliliğine, bu araştırma sırasında da büyük özen ve saygı ile yaklaşılacağına inanıyorum. Araştırma sonuçlarının eğitim ve bilimsel amaçlarla kullanılması sırasında kişisel bilgilerimin ihtimamla korunacağı konusunda bana yeterli güven verildi.

Projenin yürütülmesi sırasında herhangi bir sebep göstermeden araştırmadan çekilebilirim. (Ancak araştırmacıları zor durumda bırakmamak için araştırmadan çekileceğimi önceden bildirmemim uygun olacağının bilincindeyim). Ayrıca tıbbi durumuma herhangi bir zarar verilmemesi şartıyla araştırmacı tarafından araştırma dışı da tutulabilirim.

Araştırma için yapılacak harcamalarla ilgili herhangi bir parasal sorumluluk altına girmiyorum. Bana da bir ödeme yapılmayacaktır.

İster doğrudan, ister dolaylı olsun araştırma uygulamasından kaynaklanan nedenlerle meydana gelebilecek herhangi bir sağlık sorununun ortaya çıkması halinde, her türlü tıbbi müdahalenin sağlanacağı konusunda gerekli güvence verildi. (Bu tıbbi müdahalelerle ilgili olarak da parasal bir yük altına girmeyeceğim).

Araştırma sırasında bir sağlık sorunu ile karşılaştığımda; herhangi bir saatte, hangi araştırmacıyı, hangi telefon ve adresten arayabileceğimi biliyorum.

Bu araştırmaya katılmak zorunda değilim ve katılmayabilirim. Araştırmaya katılmam konusunda zorlayıcı bir davranışla karşılaşmış değilim. Eğer katılmayı reddedersem, bu durumun tıbbi bakımına ve hekim ile olan ilişkiye herhangi bir zarar getirmeyeceğini de biliyorum.

Bana yapılan tüm açıklamaları ayrıntılarıyla anlamış bulunmaktayım. Kendi başıma belli bir düşünme süresi sonunda adı geçen bu araştırma projesinde “katılımcı” (denek) olarak yer alma kararını aldım. Bu konuda yapılan daveti memnunlukla ve gönüllü olarak kabul ediyorum.

İmzalı bu form kâğıdının bir kopyası bana verilecektir.

KATILIMCI GÖRÜŞME TANIĞI

Tarih :
Doktor :
Adres :
Tel :
İmza :

KATILIMCI İLE GÖRÜŞEN HEKİM

Tarih :

Doktor :

Adres :

Tel :

İmza :



APPENDIX B

EGG D400 Device Specifications

Laryngograph Ltd, 1 Foundry Mews, *London NW1 2PR*

**European Communities Council Directive 93/42/EEC
Concerning Medical Devices**

DECLARATION OF CONFORMITY

Certificate No: DC-7

Issue C; October 2009

The undersigned declares that the products named in this certificate meet the provisions of this Council Directive and the CE Mark may be affixed.

General Product Name:	Laryngograph microProcessor and Software; 3 variations: 1 Laryngograph microprocessors (original model, model EGG-D200, model EGG-D400); microphone and electrodes; 2 Laryngograph microProcessor with: integral USB interface & Lx/Tx board; microphone, neck electrodes; optional additional sensors and software providing PC/laptop based speech and swallowing related displays and analyses; 3 As in variation 2 above but additionally as part of: a precision stroboscopic system for the imaging of the vocal folds; and, or FEES (oro-pharyngo-laryngeal constant light source) imaging.
Manufacturer:	Laryngograph Ltd. 78 Manor Road Wallington SM6 0AB Greater London, UK
Intended Use:	Windows compatible USB interface to provide and also to respond to: vocal fold contact and swallowing correlated signals derived from superficial skin contact electrodes; acoustic speech signals from a microphone; and also to provide triggering pulses – these may be used for voice analysis, as a trigger input to a stroboscopic system and for neural stimulation. Designed to operate with Windows based software programs for use by ENT and GI surgeons, Laryngologists, Speech and Language Therapists for data recording, creating real-time displays, instantaneous quantitative analysis, speech and swallowing therapy, monitoring and training.

Notified Body: British Standards Institution Medical Device Product Certification
Maylands Avenue, Hemel Hempstead, HP2 4SQ UK.
Notified Body No. 0086

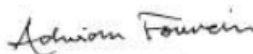
Sterile: No

Measuring Function: Yes

Conforming to harmonised standards: EN 60601-1: 1990 plus amendments A1: 1993, A11: 1993, A12: 1993, A2: 1995 and A13: 1996
EN60601-1-2: 2001

Directive Classification No: II(a)

Signed:



Date: 2 October 2009

Name: Mr Adrian Fourcin

Who are the natural and legal individuals with responsibility for the design, manufacture, packaging and labelling before the device is placed on the market under his own name, regardless of whether these operations are carried out by the Manufacturer or on his behalf by a third party.



Specifications

Laryngograph Processor

Microphone	Omnidirectional (pressure sensitive) electret, +/- 2dB 100Hz to 10kHz noise level 26dB (SPLA), dynamic range 88dB
Laryngograph	Gold plated electrodes in small, medium and large sizes
Bandwidth	+/- 1dB, 1Hz to 10kHz
Gain	0-22.5dB, software adjustable

USB Interface

Analog inputs	4 channel, +/- 5V, 16-bit A to D, 90dB dynamic range
Sampling rate	24, 16, 12kHz
Analog outputs	Speech and Lx waveforms, 16 bit D to A, speaker or headphone compatible
PC Interface	USB 2.0

Speech Studio Software

- record speech and laryngograph and optionally up to two more waveforms to hard disk at 12000 or 16000Hz
- real time display of waveforms, fundamental frequency (Fx), amplitude (Ax), frication, contact quotient (Qx), spectrogram and LPC spectrum.
- high quality playback
- display of fundamental frequency (Fx) and/or closed quotient (Qx)
- pattern display combining Fx, amplitude (Ax), frication and optionally nasality
- realtime acoustic spectrogram with narrowband (40Hz) or broadband (200Hz) resolution
- realtime display of formants via LPC spectrum
- Sustained vowel analysis
 - Minimum, maximum, average and SD for Fx and Qx
 - Jitter (%)
 - Shimmer (% and dB)
 - HNR (harmonic to noise ratio)
 - NNE (normalised noise equivalent)
 - RAP (relative amplitude perturbation)
- Connected speech analysis (QA)
 - First and second order frequency distribution (DFx1 and 2)
 - Pitch crossplot (CFx) with irregularity score (%)
 - First and second order closed quotient distribution (DQx1 and 2)
 - Closed quotient crossplot (CQx) with irregularity score (%)
 - First and second order amplitude quotient distribution (DAX1 and 2)
 - Amplitude crossplot (CAX) with irregularity score (%)
 - Speech pattern elements – time spent in voice, non voice, friction and nasalance (with optional Nasality Processor) (%)
 - Dynamic phonetogram – Ax vs Fx, first and second order
 - Qx vs Fx, first and second order
 - Statistics

Minimum PC specification

P4 or Pentium-M (Centrino) processor, 512 MB RAM, 80GB IDE HDD, CD-RW, USB 2.0 interface, Windows XP Professional.

Laryngograph

ISO9001:2000
ISO13485:2003
FM73968

Laryngograph Ltd. • 1 Foundry Mews, London, NW1 2PR
Tel: +44 (0)20 7387 7793 • Fax: +44 (0)20 7383 2039
E-mail: lx@laryngograph.com • www.laryngograph.com

CURRICULUM VITAE

Personal Information	
Name Surname	Emrah ORUN
Place of Birth	Menemen
Date of Birth	January 1981
Nationality	☒ T.C. ☐ Other:
Phone Number	+90553534638
Email	e.orun@outlook.com
Web Page	--

Educational Information	
B. Sc.	
University	Gazi University
Faculty	Technical Education Faculty
Department	Electronics and Computer Education / Electronic Teaching
Graduation Year	2003

M. Sc.	
University	Gazi University
Institute	Informatics Institute
Department	Health Informatics
Programme	Health Informatics
Graduation Year	2009

Ph. D.	
University	İstanbul University
Institute	Institute of Graduate Studies in Science and Engineering
Department	Department of Biomedical Engineering
Programme	Biomedical Engineering Programme
Graduation Year	2017

Publications	
Orun E, Akkoz A, Yuksel A, Ark N, (2013). “The Classification of Nasal Polyps Disease over Voice Recording of individuals by Using Support Vector Machine”, National Congress of Medical Technologies, 31 October – 2 November 2013, Antalya, Turkey.	