People's Democratic Republic of Algeria

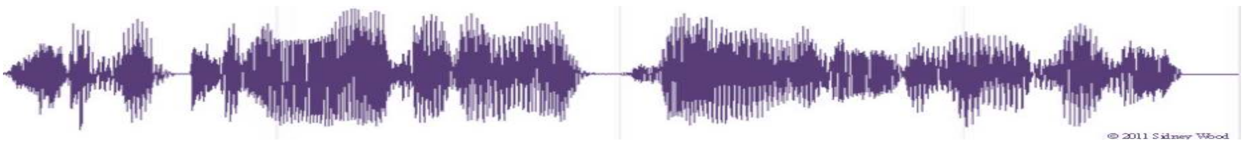
Ministry of Higher Education and Scientific Research

University of Batna II

Faculty of Arts and Languages

Department of English Language

Innovations in Practical Pronunciation Teaching: A Case Study of the Department of English at Batna University



Thesis Submitted in Part Fulfilment of the Requirement for the es Science Doctorate Degree in Language Sciences/Speciality: Phonetics

Submitted by: **Mrs. Badache Linda** Supervised by: **Pr. Beghoul Youcef**

Board of Examiners

Chairman: Dr. Kaouli Nadhir (Batna II University)

Supervisor: Pr.Beghoul Youcef (Mentouri University, Constantine)

Member: Dr. Maâouche Salima (Abderrahmane Mira University, Bejaia)

Member: Dr. Benboulaid Charif (Batna II University)

Member: Dr. Chelli Saliha (Mohammed Kheider University, Biskra)

2017/2018

Dedication

To my late father and my beloved mother I dedicate this work;

To my endless love, my husband Amir;

To God's gift, my son Alexandre Aylan

And to my lovely parents-in-law and all my family for their love and unwavering support

Acknowledgments

First, I would like to give special thanks to my supervisor Pr. Youcef Beghoul for reviewing the draft and making invaluable guiding comments about its various aspects.

I am very grateful to the examining members of the jury. Undoubtedly their priceless comments and observations will efficiently contribute to polishing the research work.

I am so grateful for the guidance that I received from many teachers and researchers at The Sorbonne University (Paris3), who offered their time and expertise without reservation in helping me design and write this thesis. I am also indebted to Pr.Vaissière Jacqueline, for having taught me different cues of reading spectrograms. Mrs.Claire Tardieu, Pr.Cédric Gendrot, who taught me how to use Praat. Pr. Phillipe Martin from Université Paris Diderot (Paris7), for teaching me experimental phonetics (Acoustic Phonetics), Pr. Jean Yves Dommergues from the Université Saint Denis (Paris8) for teaching me Psycholinguistics and for opening my mind to statistical analyses. I would also like to express my appreciation to Cecile Fougéron for her assistance in rating the pronunciation samples. I would like to single out Anna Mary Mattson, who generously offered of her precious time for recording the sample of American pronunciation, and Mr Mike Butler, a broadcast trainer, for his reading a British Pronunciation.

I will not forget to thank all my colleagues and students who have contributed to this study.

Abstract

This work aims at investigating a recent technology and exploring the ways by which it can be used to enhance pronunciation teaching and learning. This technology consists of speech analysis software which provides students with both visual and audio data and gives them an immediate and authentic feedback on their production using Praat computer software of phonetic analysis. The hypothesis is that if students were introduced and trained using this technology, the training provided by these instruments would considerably improve their sound production. The research is based on a pilot study conducted at the Sorbonne University which was undertaken to implement this pronunciation teaching method by testing the effectiveness of using this computer based visual feedback systems to help students better and improve their pronunciation. To test the hypothesis, we opt for the use of different data gathering tools namely the questionnaire, the recording of students' production and the teachers'interview. The first results show that Praat can enable students to refine their pronunciation and remedy their errors. By having instant, visual and automatic feedback displayed in spectrograms, students are able to actually observe the errors that they might not otherwise notice through listening alone. The work also demonstrates that the training from Praat is transferred to sound production as the students are able to produce the utterances more clearly and naturally in a subsequent reading task. The final results of this study show how teachers and students can benefit greatly from this free and readily available open-source, educational software, which helps students, at least, gain confidence when they speak.

List of Abbreviations and Acronyms

ASR: Automatic Speech Recognition

BE: British English

CAPT: Computer Assisted Pronunciation Teaching

CALL: Computer Assisted Language Learning

CLL: Community Language Learning

CPS: Cycle per Second

CSL: Computerized Speech Laboratory

DARPA: Defence Advanced Research Projects Agency

dB: Decibel

DLL: Digital Language Learning

ESL: English as Second Language

EFL: English as a Foreign Language

Hz: Hertz

IPA: International Phonetic Association 1886, International Phonetic Alphabet 1897

F0: Fundamental Frequency

F1: Formant one

F2: Formant Two

F3: Formant Three

F4: Formant Four

FCE: First Certificate in English

FFT: Fast Fourier Transformation

KET: Key English Test

L1: First Language

L2: Second Language

LTL: Language Teaching and Learning

MRI: Magnetic Resonance Imaging

Ms: Millisecond

NNSs: Non Native Speakers

NAE: North American English

PC: Personal Computer

PET: Preliminary English Test

SHM: Sound Harmonic Motion

SPE: Sound Pattern of English

TGG: Transformational Generative Grammar

TESOL: Teaching English to Speakers of Other Languages

ToBi: Tone Break Indices

UCL: University College London

VOT: Voice Onset Time

List of Tables

List of Figures

Table of Contents

General Introduction

Statement of the Problem

Learning a foreign language requires mainly training students through the four skills: listening, reading, writing and speaking. Speaking as a skill, usually presents difficulties to BA students of English at Batna II University. What is of concern in this research is one aspect of the speaking skill which is pronunciation.

In fourteen years of teaching at the department of English, Batna II University, we have noticed that students not only encounter great difficulties in pronouncing but also in applying the phonetic rules they are taught. Even though some students speak rather fluently, their pronunciation is not always correct. This may result from not only lack of practice but also lack of interest in phonetics as a module or simply lack of motivation for learning pronunciation and performing correctly. Moreover, the teachers tend to neglect correcting students' pronunciation errors.

Aims of the Study

Through this work, our intention is to discuss the main reasons of mispronunciation, to show the common mispronunciations and the students' weaknesses in this aspect of speaking. We attempt to provide teachers with innovative, helpful and suitable devices to better teach pronunciation, putting into practice the theoretical lectures we have been teaching in phonetics by using software.

We aim to make students perceive the importance of accurate pronunciation in learning English, motivate them to try to make more efforts in order to improve their pronunciation. Our work will hopefully make students aware of the paramount importance of pronunciation in learning any foreign language, encourage them to make more efforts to improve their pronunciation of English and reduce their performance errors.

Research Questions

Pronunciation is neglected in our department and is given little attention though it is of a major importance. Based on our interest in the use of computer technology for the teaching of pronunciation, the following questions have been asked and are sought to be answered in this research. In particular, our concern is the great need not only to teach pronunciation but also to find out new methods and techniques which enable our students not only to practice and develop their speaking skill but also to improve their pronunciation.

1. What are the major problems that students encounter in the pronunciation skill?
2. How are they taught pronunciation?
3. What are the reasons behind their lack of interest in this aspect of foreign language?
4. How can students, in particular shy or reticent ones, become more confident and motivated to learn pronunciation?
5. How can innovative techniques and devices to teach pronunciation make phonetics teaching more interactive and more interesting?
6. To what extent would the use of Praat software improve students' pronunciation?

Basic Assumptions

It is assumed in this study that:

- Students, when working in groups, will obtain better results in pronunciation exercises.

- Making the teaching of phonetics more lively and more interesting enhances students' learning.
- Innovative techniques and devices to teach pronunciation certainly spark students to become eager and motivated to learn mainly because they can easily download software on their laptops and freely exercise at home.
- Shy, reticent students will certainly feel at ease, secure and more confident to pronounce even if their pronunciation may sound wrong.
- Students' pronunciation improves thanks to students' training mainly as they can work at ease at home without the presence of the teacher.

Hypotheses

- If the origin of the students' failure to pronounce correctly such as the lack of interest and practice as well as language interference (Arabic and French) were known, it would help teachers solve their students' problems of mispronunciation.
- If students paid more attention to their pronunciation, they would take into account the importance of putting into practice the phonetic rules that they have been taught.
- If teachers found suitable devices to teach pronunciation, this would certainly encourage students to improve and better their pronunciation.

Methodology

The teaching of the pronunciation of any foreign language must encompass both segmental and suprasegmental aspects of speech. Thus investigating the problem of mispronunciation requires the use of an experimental method. The experiment will display students' errors and through an experimental method, we can classify the errors according to their origins.

The experimental method is the most appropriate and adequate one in our work for it enables us to control all the essential factors. It is mainly used in natural and physical sciences, and recently it has been of great use in social and human sciences, and hence, in the field of education.

Experimental methods have obtained satisfactory results when the design is carefully used, and since it includes experimentations, the researcher is supposed to obtain the same results in case the experiment is repeated.

Definitions of Key Terms

Below is a list the definitions of key terms that will help the reader in understanding better the various aspects of this research.

1. **Pronunciation:** The way in which a language or a particular word or sound is spoken.
2. **Innovation**: An innovation is something original, new, and important - Innovation differs from invention in that innovation refers to the use of a better and, as a result, novel idea or method, whereas invention refers more directly to the creation of the idea or method itself.
3. **Communicative Competence**: The ability to accomplish communication goals through correct use of language.
4. **Feedback:** A function of an interlocutor’s ability to understand and cognitively process the pronunciation of a speaker.
5. **Pausing:** A brief suspension of the voice to indicate the limits and relations of sounds, words, and sentences (adapted from Webster's Collegiate Dictionary, 1975).
6. **Practicing**: The act of producing sounds either in isolation or in communicative contexts.
7. **Segmentals**: Discrete units of speech that can be identified physically or auditory as vowels and consonants

8. **PRAAT Software:** Praat from the Dutch word *Talk* is used to obtain a real representation of the students' pronunciation. It is free scientific software, which enables visualizing, annotating, playing and analysing sound objects in terms of acoustic properties (frequency, intensity, pitch...). It is a useful tool to teach aspects of suprasegmental pronunciation, such as intonation, as well as segments such as vowel sounds.
9. **Spectrogram:** A spectrogram, or sonogram, is a visual representation of speech. It has two axes: The vertical axis represents the frequency (Hz) and the horizontal axis represents the time measured in milliseconds (ms). Spectrograms are used to identify spoken words phonetically.

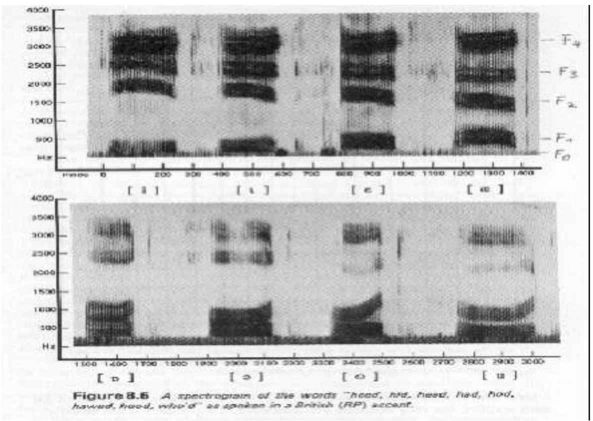


Figure1: The Spectrogram (Adapted from Ladefoged, 2006)

10. **Suprasegmentals**: Elements of stress, rhythm, and intonation of native speech (prosody)
11. **Mispronunciation**: Is defined in Oxford English dictionary as bad pronunciation.

Sampling

The population of this study includes all first, second and third year BA students of English at Batna University during the academic year 2011/2012 .Yet, it is neither possible, nor desirable to study the entire population, since, according to Deldime&Demoulin (1975), sufficient data can be obtained through the study of a proportion of the population i.e. *a sample*.

We avoided the use of a random sampling for it is likely to make us fall onto subjects of the same features and characteristics, which are not essential in our work. Moreover, random sampling cannot be used because the subjects should be selected according to certain variables like their previous grades in the modules of phonetics and oral expression, motivation to learn and so on…Almost all of these students were our students in their second year.

A systematic sampling technique in selecting learners gives us the opportunity to have a group of mixed levels. This can raise our chance to have a heterogeneous group of diverse characteristics and abilities and, thus, diminish the risk of bias resulting from selecting students randomly.

The selected group, then, consists of students from the third year level, classified according to their grades in phonetics. The selected sample, which comprises students with different levels, is then, sub-divided into a control group and an experimental one. Mixing students with different levels, different abilities, different genders, various ages, cultural backgrounds and different ways of pronunciations forms our designed sample.

Data Gathering Tools

In addition to the information which is obtained from the results, other ways of collecting data about students' pronunciation is used: a questionnaire and recordings and analysis of the students' pronunciation. The questionnaire helps us collect the students' own perceptions and opinions about pronunciation and the ways of teaching it. It helps us to know the origins of mispronunciation. The questionnaire was given to students at the end of the academic year 2011.

Recording students helps us obtain an immediate feedback of their speech thanks to the spectrogram and enables the students -mainly those having difficulties- to correct their mispronunciation.

These two tools are especially selected for their usefulness; their practical use and because they complement each other. The first questionnaire helps us assess the students' perceptions of pronunciation skills and the second one enables us to analyse their productions.

To support our data collection, some teachers, mainly those of oral expression and phonetics are interviewed and recorded in order to explore in depth their attitudes to pronunciation. The guiding questions for the interviews are provided in the appendices.

As far as recording is concerned, various multimedia programs and software are mentioned and explained in this study.

Software tools like Praat, Winsnoori, which are more effective, educational and multimedia visualization tools for teaching pronunciation and prosody when compared to similar software programs. Only Praat, a free software, was used in this study

The two groups began shortly after the beginning of the academic year and continued throughout the whole year. In the experimental group, students were asked to pronounce single words, in order to check the pronunciation of vowels and consonants (segmental level).

After that, they were asked to pronounce short sentences, they were checked on some of the main aspects of phonetics for instance: assimilation, elision, linking...etc. (aspects of connected speech).

Hence, the first two levels of language learning activities can be decomposed at least into phonemic aspects, which include the correct pronunciation of single phonemes and the co-articulation of phonemes into higher phonological units (intonation), these aspects include:

- The correct position of stress at word level.
- The alternation of stressed and unstressed syllables in terms of compensation and vowel reduction.
- The generation of adequate intonational pattern for each utterance related to communicative function.
- The generation of adequate rhythm from the interleaving of stress, accent and phonological rules.

Focus of this resource is on rhythm as well. Attention to sounds is integrated with rhythm practice. The units are divided into four sections: *sounds, words, sentences* and *conversations*. Within these sections is instruction that deals with syllable stress, sentence stress, linking and intonation. Practice exercises include dictations, short dialogues, limericks 1 and listening discrimination activities. Each unit ends with a review. Practice

[1]**Limerick** (*plural***limericks**) A humorous, often bawdy verse of five anapestic lines, with the rhyme scheme*aabba*,(https://en.wiktionary.org/wiki/limerick#English)

activities progress from individual words to words within sentences, i.e. within context. (Connected speech)

As it appears above, for a student to communicate intelligibly and as close as possible to native speakers' pronunciation, the application of these aspects is very important and required.

After that, we conduct a final post-test, for which students are recorded reading the following corpus:

> *All human beings are born free and equal in dignity and rights. They are endowed with reason and conscience and should act towards one another in a spirit of brotherhood.*

The Universal Declaration of Human Rights (Article one)

We even carry out a recording of the same article by a female American PhD student living in Paris; the recording is held in an anechoic chamber at the Laboratory ofPhonetics and Phonology (LPP). The second recording is of an English broadcast trainer named Mike Butler.

Students' pronunciation is compared with the one of the native speakers (American and British) as their speech is immediately displayed via the spectrogram.

Procedures for Treating Data

The activities accomplished by both groups are recorded regularly in order to continuously assess students' pronunciation and progress and, of course, compare the results of both groups.

To obtain reliable results, we administer a pre-test right at the beginning of the experiment. After three months, a first post-test is used to check the first results because three months are quite enough to teach at least two units.

After this, a second post-test is conducted in order to draw the first conclusions of the whole work and try to predict what the final results of this research are. A final post-test is then administered to ensure a complete analysis and point out the possible differences between the two groups (control & experimental). The results obtained allow us to make suggestions to improve the teaching of phonetics in general and pronunciation in particular using computer software, for instance Praat, which can be very useful and effective as a modern way of teaching as compared to the traditional one, which is teacher-centred and students have no role except sitting and listening to the teacher's explanations.

Structure of the Thesis

The thesis is divided into two major parts: the theoretical part and the empirical one.

The theoretical part is intended to present important theoretical background information about phonetics, phonology and pronunciation. It is divided into six chapters. Chapter one explores the two tightly linked linguistic subfields, which are phonetics and phonology, setting first the differences between them, providing the basic elements of phonology in general and in particular the sound distinctive features. It also delves into an assortment of branches of phonetics such as psycho-phonetics, orthophonic phonetics, developmental phonetics, neuro-phonetics, clinical phonetics, statistical or computational phonetics, articulatory phonetics, auditory phonetics and acoustic phonetics. Yet, the focus is on the two lastly mentioned disciplines which rely on the speech perception and transmission since the purpose of our research is to better our students' pronunciation using software which involves them in practising and mastering sound recording, sound perception and sound analysis. Students receive a good and indispensable instruction in acoustic phonetics to enable them to be familiar with its fundamental enquiries. For instance what is a sound and what are its characteristics? This software enables students to see their speech represented on the screen in a form of a spectrogram. What a spectrogram is, what types of spectrogram and cues for reading a spectrogram are, will be treated in this chapter. Chapter two is a detailed description of the speech production and speech mechanism. After the speech is produced, a chain is being formed between the speaker and the listener, taking into account different speech levels, which are the linguistic, the physiological and the acoustic one in order to convey any message. Chapter three examines some of the major themes in pronunciation teaching as discussed by pronunciation specialists and teachers. Answers to various questions related to pronunciation teaching are supplied, for instance: do we need to teach pronunciation? And if so, how to teach it? What is meant by pronunciation and intelligibility? Do we focus on intelligibility or correctness in dealing with pronunciation? Moreover, it goes through the two different but mandatory aspects of pronunciation which are the segmental and the suprasegmental ones, with reference to the different aspects of connected speech.Chapter four is a detailed overview, basically about the history of pronunciation teaching and its evolution though time and a research on the teaching and acquisition of pronunciation skills referring to factors influencing the learning of pronunciation.Chapter five is a general review of the relevant related literature about some practical aspect of pronunciation teaching. It deals with the pedagogical goals of teaching pronunciation using software and the improvements and advantages of using CAPT for enhancing English learners' pronunciation. After that, it outlines how a visual and immediate feedback is provided with automatic diagnosis of pronunciation errors. Chapter six scrutinizes pronunciation Teaching Programs and Software, in which we draw attention to the close interrelationship between how pronunciation teaching could be better using various programs and software.

The empirical part consists of two chapters.

Chapter seven is an account of the research methodology underlying the study. We discuss the procedures used to conduct this work, describe the population, the experimental method used, which best suits our aims, the variables, the instruments used to collect the data which form the backbone of the research and, finally, the procedure of the data analyses used in this study. Chapter eight reports on the results and findings that help us argue, discuss our claims, answer our questions and test our hypotheses. In the conclusion we draw out and state clearly what our major findings imply and identify potential limitations to the study. Finally, we indicate some directions for future research and perspectives.

PART ONE: THEORETICAL ISSUES

Chapter One

Chapter One: Phonetics and Phonology

Introduction

Two terms are often loosely used to refer to two linguistic disciplines studying that part of the linguistic sign which de Saussure (1910) called the acoustic image: *phonetics* and *phonology*. The importance of sounds as vehicles of meaning is something of which people have been aware for thousands of years. However, systematic studies of the speech sounds only appeared with the development of modern sciences. The term phonetics used in connection with such studies comes from Greek and its origins can be traced back to the verb phōnein, which means to speak, in its turn related to phōnē, which is a sound. The end of the 18th century witnessed a revival of interest in the studying of the sounds of various languages and the introduction of the term phonology. The latter comes to be, however, distinguished from the former only more than a century later with the development of structuralism which emphasizes the essential contrastive role of classes of sounds which are labeled phonemes. The terms continue to be used, however, indiscriminately until the prestige of phonology as a distinct discipline is finally established in the first half of the 20th century. Though there is no universally accepted point of view about a clear-cut border line between the respective domains of phonetics and phonology as, indeed, we cannot talk about a phonological system ignoring the phonetic aspects it involves and, on the other hand, any phonetic approach should take into account the phonological system that is represented by any language. Most linguists will agree about some fundamental distinctions between the two.

1.1. Definition of Phonetics

Phonetics will be almost unanimously acknowledged to be the linguistic science which studies speech sounds: the way in which they are *produced* (uttered or articulated), the way in which they are *perceived*, their *physical characteristics*, etc. Therefore, it is these more *palpable*, measurable aspects of the phonic aspects of language that constitute the domain of phonetics. On the other hand, it is obvious, however, even for those whose perception of linguistic phenomena is rather of an empirical and not of a very scholarly kind, that when communicating verbally, though they are producing a wide variety of sounds, people are actually *aware* of using a comparatively drastically limited set of sounds; in other words, they tend to disregard the obvious (more or less important) differences between the way in which sounds are uttered and have in mind only classes of sounds that perform a certain function in language. From this new perspective, it is not the sounds as such that are important, but rather the role they have in linguistic communication.

We have already said that phonetics is concerned with various aspects relevant for the physical characteristics of sounds (Mateescu, 2003.p.33).

Several branches of phonetics will be further distinguished in this chapter.

1.2. Definition of Phonology

Jakobson (1956) went to say that the question of how language uses sound matter, selection certain of its elements and adapting them to its various ends, is the field of a special linguistic discipline. This discipline is often called phonemics (or, puristically, phonematics) since among the functions of sound in language, the primary one is to serve as a distinctive vehicle and since the basic vehicle for this function is the *phoneme* with its components. The prevailingly continental term *phonology* was launched in 1923 based on the suggestions of the Geneva School. (p.7)

It is precisely this aspect of sounds that is of interest for phonology, which is thus understood to study not so much the sounds as such, but rather *classes of sounds* that have a certain function in the structure of a given language. This distinction will be further analyzed in the chapter dealing with *the phoneme*.

1.2.1.Elements of Phonology

There are phonological basic elements, which are compulsory for everyone -who is interested to learn about this field- to study and to be aware of their use to construct the sound system of a given language or languages in general. These are: Distinctive Features, Syllable, Syllabification, Phoneme, Allophone, Phone and Mora. Gussenhoven, (2004).

1.2.1.1. Distinctive Features

When we speak about distinctive, the first thing that comes to our minds is to distinguish, and generally when we distinguish, we state differences and similarities that may exist between two or more things. That is to say, we classify, we organize, we arrange…etc… and this is exactly what is done when dealing with sounds in phonology. But, in order to know more about how sounds can be distinguished from each other, it would be preferable to provide a general and brief overview of the history and foundations of these distinctive features to understand their system.

A point which continues to be an issue in current feature theory is the relationship between phonetics and phonology. In particular, it concerns the relationship between possible binary features on a phonological level and binary features on a phonetic level and the way such features are mapped into each other. Studies in feature analysis have shown that many of the phonological features involve the specification of more than a single phonetic parameter, a fact that had not been considered seriously enough in early investigations. Still with the question of the relationship between phonetics and phonology, the issue of whether the same set of features should be used on the phonological and the phonetic level also becomes of distinctive feature theory, since the distinctive feature developed as a purely phonological functional concept. The inclusion of phonetics in the generative model therefore also occasioned substantial changes in the feature concept. An additional issue in current feature theory is the type of correlate to be sought for each other. In their inception, features tended to be described in terms of articulatory correlates. But early mentalistic and psychological influences also brought into practice such notions as a speaker's linguistic consciousness and the way a sound was felt or perceived by the listener. A semi- concerted effort appears to have been made in the early Prague School to find acoustic or perceptual correlates for the individual features then being discovered. It is to be noted that in their early use the attributes '*perceptual*' and '*acoustic'* were used quite impressionistically and were essentially interchangeable in meaning (Baltaxe, 1978).

It is obvious that the concept of distinctive features did not develop all at once. Several steps can be identified in its development, and several related notions were significant and need to be taken into consideration. The first in the unfolding of the concept was the distinction between correlations and disjunctions. Even though this distinction had to be ultimately abandoned, it was significant in early development. Associated with it was the recognition of the existence of different types of contrastive relationships within a phonological system. The concept of phoneme content followed this discovery of appositive relationships. The exploration of phoneme content in turn ultimately led to the definition of the phoneme as a bundle of distinctive features. Even though earlier definitions of the phoneme within Prague School Phonology had included considerations of the entire system and had incorporated the notions of relativity and opposition, the development of the distinctive feature concept would not have been possible without such an exploration.

These developments, in turn, subsequently required a shift of the notion of opposition from the phoneme to the feature level. Inherent in the development of the distinctive feature was the Praguian recognition that there was coherence between the sounds of a language as a system, and that the value of each sound depended on its relationship with other sounds of that system. Hence, the value of each sound or phoneme was relative. Actually, the reality of the phoneme came first, and then that of the distinctive feature, which depended on the structure of a given phonemic system. In this sense, both were *abstract constructs*.

The recognition that phonemes constituted further divisible units can be viewed as a factor in the abandonment of the earlier mentalistic framework, since it seemed to make their study more concrete. A shift away from mentalism can also be viewed as reflecting the general trends of the times, which showed their influence in other sciences, for instance, in physiology, where the earlier mentalistic approaches were replaced by a new model, that of behaviorism (Baltaxe 1987 p, 15).

Even though, distinctive features had been gradually developed thanks to the respective roles of Trubetzkoy and Jacobson,Chomsky and Halle, in their book entitled *The Sound Pattern of English* published in 1968, gave abroad view of the subject, combining generally applicable theoretical contributions with analysis of the details of a single language. *The Sound Pattern of English* is considered by Hill and Nessly to be one of the great contributions to phonological theory in the history of linguistics and the most comprehensive and ambitious single statement on generative phonology.

McCawley (1974) said that:

> *A book which no linguist can afford not to read and reread with great care …the value of the Sound Pattern of English lies especially in the kinds of questions which Chomsky and Halle ask, the kinds of answers they formulate, and the ways in which they justify their answers, all of which lead one deeper into the inner working of the phonology of a language than any previous work. (P.40)*

In *the Sound Pattern of English*, 1968 a book which represented a turning point in the development of phonological theory in the 20th century by Chomsky and Halle, who defined the distinctive feature as "*…minimal elements of which phonetic, lexical, and phonological transcriptions are composed by combination and concatenation*" (1968, P.6).

1.2.1.1.1. History and Foundations

We will introduce the basic elements of Phonology tracing back not only the history of distinctive features but also their evolution, taking into consideration Trubetzkoy's classifications system compared with other systems developed by Chomsky, Halle, Fant and Jakobson. But before dealing with the history and foundations of the distinctive features, let us define what a distinctive feature is. Various definitions or descriptions are quoted as follows:

Distinctive features are those indispensable attributes of a phoneme that are required to differentiate one phoneme from another in a language.(Singh and Singh, 1976, p.177)

Distinctive features are really distinctive categories or classes within a linguistic system but just like in accepted analysis it is required that they are consistent with the phonetic facts and these phonetic facts on various levels[articulatory, acoustic, perceptual, or linguistic] have lent their names to the features. (Fant, 1973, p. 152)

A theory of distinctive features constructs category systems for phonemes; the categories are intended to cover all languages. Each feature category is derived by joint consideration of three levels of linguistic analysis: the perceptual level, the acoustic level, and the articulatory level. Thus, the purpose of distinctive feature theory is to provide a single consistent framework for specifying the phonology, i.e., the communicative sound structure, of any language.(Picket, 1980, p. 103)

From these different definitions, where distinctive features were said to be pertinent to phonemic and not phonetic contrasts and that the names assigned to individual features may not always come from the same level of analysis (articulatory, acoustic, perceptual, or linguistic).Nonetheless, distinctive features have a general relevance to the study of phonetic and phonology. It should be of a great priority to trace back what is the history and foundations of distinctive features.

1.2.1.1.2. Distinctive Oppositions

The theory of *Distinctive Oppositions* goes back to Trubetzkoy (1939) a Russian phonologist, who was one of the core members of the Prague school. He attempted to classify the properties of distinctive *contrasts* operative in natural languages. Trubetzkoy (1939) defined a taxonomy of phonological contrasts of meaning changing oppositions. Hence he stated that there are three opposite members which are: Oppositions within a System, Oppositional Relations and Types of Oppositional Extent. He provided a framework for alternative views '*Contrasts*'. We will examine his types of oppositions that are relevant to phonological features:

- **Bilateral**: Refers to pair sounds that share a set of features which no other sound shares fully. For instance, voiceless labial obstruents =/p,f/.
- **Multilateral**: Are oppositions of segments which have more than one phonetic property that keeps them apart. e.g., in the Korean language, we can have:

/p^h/ Aspirated

/p/ Unaspirated +/- Voice A multilateral opposition means '*many'*

/b/ Voiced +/– aspiration

- **Proportional**: If the relation between its members is identical, *the same,* for instance: [t, d] [k. g] the difference between /t/ and /d/ in English is the same difference between /k/ and /g/ thus we have proportional relations.
- **Isolated**: If no other segment stands in the same relation in the language.eg; /r/≠ /l/ that is to say:/r/ is an alveolar approximant and /l/ an alveolar lateral.
- **Privative**: Means that one member carries a phonetic property while the other does not or lacks. Privative oppositions are relationships of the 'presence versus the absence of a feature'. e.g.: voicing: [t, d] [+/– voiced]
- **Gradual**: The members of a class of sounds possess different degrees or graduations of a feature or property. They are oppositions which are characterized by different degrees or graduations are defined as '*gradual*.' This type of relationship holds for *vowels*. For example, the three short front unrounded vowels in English /ɪ, e, æ/ where the property *tongue height* is a matter of *degree* rather than *binarity*.

- **Equipollent**: If the relationship between the members of an opposition can be considered logically equivalent, it is then defined as equipollent. A typical case is the oppositions in between English [p vs. t] and [t vs. k]; unlike with vowels, it is not possible to speak of a continuum from labial to velar and distinguish the consonants by means of degrees of *backness*, rather the different places of articulation constitute an equipollent relation.
- **Constant**: The extent of an opposition can also be defined in Trubetzkoy's taxonomy. When two members can occur in all positions, they are *constant* e.g.: /t/ /d/
- **Neutralizable**: The oppositions between /t/and /d/ may disappear; it is neutralized.eg: when we hear *Stick* in a recording we will hear afterwards *sdick*. After /s/ the opposition between /t/and /d/ is *neutralized*, disappears and no longer exists. For instance, in the German language the opposition between /t/ and /d/in final positions disappears; *und* is pronounced *unt*.

1.2.1.1.3. Acoustic Features

Jakobson was also a member of the Prague school of linguistics and worked closely with Trubetzkoy. Distinctive feature theory, based on his own work and the work of Trubetzkoy, was first formulated by Jakobson in 1941 and remains one of the most significant contributions to phonology.

Defined by Jacobson, a term presented initially in a book written with the collaboration of *Fant and Halle*, entitled *Preliminaries to Speech Analysis in 1952.*The book was the first major attempt by the structural school of linguistics to give a comprehensive and articulate, coherent picture of the distinctive features in language. Four years later, he co-authored by Halle *Fundamentals of Language* in 1956. After Jacobson went to the United States, he adopted the term distinctive feature from the Bloomfield's *Language*1933. However, Bloomfield had used the term in the sense of a phoneme.

It was mentioned earlier that the use of the terms *perceptual and acoustic* were interchangeable. It appears, however, that *acoustic* was generally intended to mean *perceptual*. Early feature correlates were primarily articulatory and secondarily *perceptual*. The first consistent use of acoustic (in the sense of instrumentally verifiable) correlates in Jacobson's *Preliminaries* aimed at the description of the oppositions operative in languages, together with his Swedish colleague Fant, and the American phonologist Halle, he proposed a system of binary features now called *distinctive*; they capture phonological contrasts rather than describe the phonetic segments themselves. But what exactly does binary mean?

1.2.1.1.4. The Concept of Binary: A feature is defined as binary if it has two *values*, one of which is designated as **[+]** the value and the other is **[–]** the particular value. In many cases a binary approach is only phonologically significant in those oppositions, which Trubetzkoy termed *privative*, for example, the feature *nasal* is binary logically; however, phonetically, it is a matter of degree, since some speech sounds can be more or less heavily nasalized. Probably due to the availability of the *Sonograph* and later referred to as the *Spectrograph*, which enabled phonologists to analyze the speech sounds acoustically. Jacobson's features were acoustic in character, and the +/–notation in phonology was introduced for the first time, by Jacobson in 1949 in his book *On the Identification of phonemic Entities* (1949/1962, pp.418–425) and later in *Notes on the French Phonemic Pattern* (1951/1962, pp.426–434). Thus, he defined these three feature groups as: Major Class Feature, Features for vowels, Features for consonants and Secondary Articulation Features.

Jacobson (1938), by a change in parameters for the classification of consonants along the horizontal axis, that is by changing in the previous point of classification in articulation. Instead of classifying the consonants in terms of points of articulation, he found a common and differential denominator in the shape and volume of the oral resonance chambers. He then successfully combined these newly found articulatory parameters for the consonants with certain acoustic observations, and thus established correlates for his newly found division as well.

Starting from the observation that in many languages a close relationship could be noticed between *velars and labials*, Jacobson found that what these sounds had in common was their *long, undivided oral resonator* as opposed to the palatals and dentals, which had two small resonance chambers. Furthermore, in the production of the velars and labials the pharynx was retracted, but this retraction did not occur with the palatals and the dentals:

> *Velars and labials take their specific character from a long, undivided oral resonator; while for the palatals and dentals the tongue divides the oral cavity into two short resonance chambers…the pharynx is retracted for velars and labials, while it is expanded for the corresponding palatals and dentals.* (Jacobson, 1938/1962, p.274)

Having grouped the velars and labials – as contracted to be palatals and dentals- as *grave* versus *acute* consonants, Jacobson then further grouped velars and palatals (including sibilants) together, uniting them under the term *back consonants* (postérieurs) and opposing them to the labials and dentals , which he classed as *front consonants* (antérieurs). The specific feature separating the *back* from the *front* consonants was their point of articulation. For the back consonants, the point of articulation was behind and for the front consonants it was in front of the only or dominant resonance chamber. In this regard, he said:

> *Similarly, the velar and palatals, including the sibilants, are opposed to the labials and dentals by means of a specific characteristic. By grouping velars and palatals (including sibilants) under "back consonants" and labials and dentals under "front consonants" the*

> *following rule can be stated: for the back consonants the point of articulation is behind the only or major resonance chamber and for the front consonants it is in front of it.*(1939/1962, p. 27)

Jacobson (1938) thus articulatorily reduced the velars, palatals, dentals, and labials to two oppositions, based on *shape* and *volume* of resonance chamber. Actually, it was not Jacobson who introduced this new parameter into phonology. It had already been present with respect to the vowels in the work of Trubetzkoy (1931) (***Zurallgem einen theorie***). However, Jacobson did extend it to the consonants. Similarly, with respect to the opposition *grave/acute*, the terms *grave* and *acute* had already been used for *vowels* by Jacobson in his book entitled *Phonemic Notes on Standard Slovak*, (1931/1962, p. 221)

Jacobson then examined the above oppositions from an acoustic point of view and found that for the back consonants, as opposed to the front consonants, the distinction was one of a higher degree of perceptibility, often accompanied by a greater degree of duration. He pointed that "… *the back consonants are contrasted with the corresponding front consonants by means of a greater degree of perceptibility frequently accompanied, ceteris paribus, by a longer duration"*(1938/1962, p.274).

The distinction between the *grave* class of labials and velars, on the one hand, and the class of *acute* dentals and palatals, on the other, was that the former had a relatively *dark timbre* and the latter had a relatively *clear timbre*.

> *A long and undivided resonator as well as a retracted back orifice lends a characteristic and relatively dark timbre to the velar and labial consonants. It corresponds to the timbre of the velar vowels and contrasts with that of the palatals and dental consonants is relatively clear and almost corresponds to the characteristic timbre of the palatal vowels* (1938/1962, p. 274)

What is termed here *acoustic* is not necessarily acoustic as it is used in Jacobson's book *Preliminaries to Speech Analysis* and *Fundamentals of Language.* The term as used in *Observations* should be interpreted as *impressionistically* acoustic. In contrast, the term was used in the sense of *instrumentally* acoustic in *Preliminaries* and *Fundamentals.* But already in *Observations (1938)*, Jacobson was attempting to relate this classification to acoustic experiments.

> *Thus, in Preliminaries and Fundamentals the terms back and front consonants were replaced by the terms compact and diffuse. The opposition was now described in instrumental acoustic terms:…compact phonemes are characterized by the relative predominance of one centrally located formant region (or formant). They are opposed to diffuse phonemes in which one or more non-central formants or formant regions predominate* (Jacobson, Fant, and Halle, 1952, p.27)

In *Fundamentals,* the distinction was acoustically refined: "*…higher (vs. lower) concentration of energy in a relatively narrow, central region of the spectrum, accompanied by an increase (v. decrease) of the total amount of energy*. (Jacobson and Halle, 1956, p.29).

In *preliminaries* the opposition *compact/diffuse* was differentiated in articulatory terms.

> *The essential articulatory difference between the compact and diffuse phonemes lies in the relation between the volume of the resonating cavities in front of the narrowest stricture and those behind this stricture. The ratio of the former to the latter is higher for the compact than for the corresponding diffuse phonemes. Hence the consonants articulated against the hard or soft palate (velars and palatals) are more compact than the consonants articulated in the front part of the mouth.* (Jacobson, Fant, and Halle, 1952, p.27)

In *Fundamentals, compact* versus *diffuse* were differentiated in articulatory terms:

> *…forward-flanged versus backward-flanged. The difference lies in the relation between the volume of the resonance chamber in front of the narrowest stricture and behind this stricture. The ratio of the former to the latter is higher for forward- flanged phonemes (wide vowels, and velar and palatals, including post-alveolar consonants) than for the corresponding backward-flanged phonemes (narrow vowels, and labial, dental, and including alveolar consonants).* (Jacobson and Halle, 1956, p.29)

A similar change and sharpening of definition can be seen with respect to the *grave/acute* dichotomy. In *Preliminaries* the opposition *grave/acute* was differentiated in instrumental acoustic terms:

> *Acoustically this feature means predominance of one side of the significant part of the spectrum over the other. When the lower side of the spectrum predominates, the phoneme is labeled grave; when the upper side predominates, we term the phoneme acute.* (Jacobson, Fant, and Halle, 1952, p.29)

In *Fundamentals* the acoustic correlate was given as*"…concentration of energy in the lower (vs. upper) frequencies of the spectrum"* (Jacobson and Halle, 1956, p.31).

In articulatory terms, the opposition was differentiated in *Preliminaries* in the following way:

> *The gravity of a consonants or vowel is generated by a larger and less compared mouth cavity, while acuteness originates in a smaller and more divided cavity. Hence gravity characterizes labial consonants as against dentals, as well as velars versus palatals.* (Jacobson, Fant, and Halle, 1952, p.30)

In *fundamentals,* the articulatory distinction was given as:

> *…peripheral vs medial: peripheral phonemes (velar and labials) have an ampler and less compartmented resonator than the corresponding medial phonemes (palatal and dental)*(Jacobson and Halle, 1956, p.31).

A comparison of the first binary classification of the two oppositions discussed shows that Jacobson, after successfully applying the dichotomous scale in his book *Observations*, together with Halle subsequently refined his articulatory specifications with a slight shift in emphasis. However, what must be considered as much more important is the replacement of *impressionistic acoustic corollaries* by a consistent effort at an *instrumental – acoustic* classification in *Preliminaries* and *Fundamentals.*

Observations, in addition, analyzing the previous point of articulation parameter into two binary oppositions, was further important in that Jacobson likened the acoustic impression related to these consonantal oppositions to those found in the vowel oppositions. Furthermore, he gave instances, of the practical application for these features. For example, by using these features, he was able to relate contextual changes in vowel or consonant to a single feature. Thus a segment which was[+consonant] [+grave] could change to [+consonant], [+acute] in the environment of [+ vocalic], [+acute].

In Trubetzktoy (1931),however, the classificatory status of *nasal* versus *oral* consonants had not yet been decided. However, Jacobson in his book *Observation* proposed to oppose the nasals to the orals as *"the result of a divided passage way"(1938/1962, p.276).*

Jacobson then, considered these three oppositions*:Grave/ Acute, Compact/Diffuse, Oral/Nasal*. All were based on a *different place and structure of the resonators*- as the kernel of any phonological system. At the same time he called attention to their importance in first language acquisition: *"…it is to these three consonant oppositions that some archaic types of primitive language and also child language are limited."(1938/1962, p. 276)*.The liquids, which Trubetzkoy had left unanalyzed vis- à- vis the phonological system were incorporated in *Observations* into the analysis of the feature system. Trubetzkoy had organized the correlative relationship between liquids. However, he was not sure how the liquids should be characterized vis-à-vis the entire system. Jacobson, was able to distinguish the liquids from the rest of the consonants *by the simultaneous opening and closing of the buccal passage* (1938/1962, p.278.). It seems that this definition of the liquids in *Observations* already pointed the way to their classification in terms of a split of the *vocalic/consonantal* features found in *Preliminaries and Fundamentals*. In *Observations,* hence, the liquids had been opposed to each other in terms of the *simultaneous closure* of the *oral cavity:*

> *…for the laterals the two aforesaid simultaneous actions are in fact implemented at the same time, but at two different places, while for the trills opening and closure occurs at the same place, but in succession* (1938/1962, p.278).

Jacobson established yet another binary opposition for the consonants in his book *Observations:* the opposition *strident/mellow*. He accommodated the classification of the *affricates* (such as the oppositions t–ts and p–pf in German). However, Trubetzkoy in his book *Die Phonologischen Systeme* (1931) grouped these under the opposition of *constriction stops/Fricatives*, an opposition which he considered as *gradual*. (For the subsequent Trubetzkoyan framework for stridency in *Grundzüge (1938*, p.112)

> *The stridents in Observations by the expiring air resulting in a sharp tone… An additional obstacle participating in this friction thus*

> *distinguishes the articulation of the strident fricatives from that of the mellow fricatives* (1938/1962, p.277).

Acoustically, *strident* sounds are sounds that have *irregular wave forms*…as opposed to *mellow* sounds that have *regular wave forms* (*Preliminaries* p.23). In *Fundamentals*, however, the distinction between *strident* and *mellow* sounds is presented as acoustically related to *higher intensity* versus *lower intensity noise* and genetically to *rough-edged* versus *smooth edged*.

Let us look at Jacobson's acoustic features in detail. The first group of features he defined are referred to as *Major class features*: they have been introduced to distinguish the major classes of *segments*: vowels, consonants and segments somewhere in between nasals, liquids and glides. However, there is disagreement about the number and type of *major classes* features. Jacobson introduced the features *consonantal* and *nasal*. Later, further features were added. Examples:

/u/ [–cons][–nasal] **/l/ [+ cons][–nasal]** **/j/ [–cons][–nasal]**

/p/ [+cons][–nasal] **/m/ [+cons][+nasal]**

So in Jacobson's approach, all vowels are assigned the major class feature [–consonantal]. To account for the parameters *tongue height*, *tongue position* and *lip rounding*, additional features based on acoustic measurements came in, and here they are, they look complicated at first sight but they are not.

- **[Diffuse]** relates to all segments whose energy concentration can primarily be found in a narrow region central to the frequency spectrum. This defines vowels such as /i/ and /ʊ/ as well as most non alveolar consonants as [+ diffuse]

- **[Compact]** relates to all segments with low energy concentration in the narrow central regions of the spectrum. This defines the low vowels as [+compact]
- **[Grave]** relates to all segments whose energy concentration can primarily be found in a narrow region of the lower frequencies of the frequency spectrum. This defines all back vowels as [+ grave]
- **[Flat]** relates to all segments with weak or low energy concentration in the upper regions of the frequency spectrum. This defines most back vowels as [+ flat]

In Jacobson feature system, all segments specified as [+ consonantal] qualify as a consonant. This feature distinguishes sounds with low energy and substantial obstruction from non-consonantal sounds. Among the features assigned to such segments are:**[+/– diffuse], [+/– grave], [+/– flat].** They were also used for vowels as we have just seen. Additional features were primarily articulatory in character.

- **[Voice]** correlates with vocal folds action, a speech sound is voiced and has a feature [+voice] if the vocal folds *vibrate*. It has the feature [–voice] if the glottis is *open.*
- **[Nasal]** sounds are produced by lowering the soft palate or 'velum' and allowing the air to pass outward through the nasal cavity. Non nasal or *oral* sounds are produced with the velum raised to prevent the passage of the air through the nose.
- **[Continuant]** sounds involve a primary constriction allowing the air to flow in the mid-sagittal region of the vocal tract. Sounds produced with a sustained occlusion are defined as [–continuant] hence vowels, glides, trills and fricatives are [+continuants].Yet, plosives are *interrupted* or [–continuants].
- **[Strident]** or sharpness of noise was introduced by Jacobson to distinguish consonants involving high energy and high frequency from consonants involving low energy and

low frequencies, thus it distinguishes noisy consonants like **/s/** from more *mellow* consonants like**/**ʃ/, so alveolar fricatives are classified as [+strident] whereas dental, labiodentals, post-alveolar ones are defined as[–strident]

- **[Tense]** This characterizes sounds which are articulated with a greater effort. Acoustically they evince a greater spread of energy in the spectrum and have a longer duration, while articulatorily, they require a greater deformation of the vocal tract. Voiceless consonants are thus specified, while voiced ones will be described as [-tense].

1.2.1.1.5. The Sound Pattern of English 1968 'SPE' (Chomsky and Halle)

The most influential approach of our time is the *Sound Pattern of English*, which is essentially a book that was written and published in 1968 by Chomsky and Halle. They defined the distinctive features in two ways: as units that capture phonological contrasts and patterns of languages and as means that describe the phonetic content of the segments. The Sound Pattern of English was subdivided into four general groups of features: *Major Class Features, Cavity Features, Manner of Articulation Features* and *Source Features*. (p.299)

a. Major Class Features dealt with the fundamental *vocalic/ non-vocalic* and *consonantal/ non-consonantal* distinctions. The distinction suggested by Chomsky and Halle was essentially an articulatory one: the uttering of vowels did not involve any major obstruction in the way of the airstream, while a major constriction at some point along the vocal tract was always associated with the articulation of consonants. Like Jacobson , Chomsky and Halle used the feature **[Consonantal],** which is abbreviated as [cons] to distinguish consonants from vowels; they defined all those sounds that are produced with some sort of constriction in the vocal tract, at least equal to that found in fricative consonants, as [+ cons].

Just as with Jakobson, *liquids* were described as being [+consonantal, +vocalic], glides were [-consonantal, -vocalic], a combination of features that also characterized the glottal fricative /h/and the glottal stop/ʔ/.

The distinction *Sonorant/ Obstruent* was introduced. The former being described as sounds allowing spontaneous voicing. Vowels, glides, liquids and nasals were naturally included, though it was not clear why /h/and the glottal stop /ʔ/received the same specification. A refinement of these features was suggested in the epilogue of the book where vowels are described as *syllabic* and *vocoid*, while glides are characterized as *non-syllabic* and *vocoid*. Thus, the articulatory similarity between vowels and glides is captured, the difference being one of distribution in the position of syllable nuclei.

Syllabic/Non-syllabic: sounds that have the feature [syllabic] are those which constitute peaks of syllables. Non syllabic sounds are those which are in the margin of syllables. For instance: the word cattle/ kǽtl/, /k/ which is the onset is defined as [+ cons] [–syllabic], /ǽ/ stands as a peak is [–cons] [+syllabic], /t/ is [+cons] [–syllabic], /l/ is [+ cons] [+syllabic].Consonants are then described as contoids and vowels as vocoids while the same distinction syllabic/non-syllabic that differentiated between vowels and glides operates in the case of consonants too. It keeps apart syllabic consonants (nasals and liquids) and non-syllabic ones (the true consonants or obstruents).

b. Cavity Features were essentially place of articulation features.

Coronal sounds (a new feature originating in Jacobson’s grave/acute opposition) were defined as sounds produced with the blade or the tip of the tongue raised from the neutral position (dental, alveolar, palate-alveolar consonants). The Feature [Coronal] refers to the crown of the tongue.

Anterior sounds, which can be associated to Jacobson's *Compact/Diffuse* one, are sounds produced in front of the palate-alveolar region.

Speech sounds, can be defined as [Sonorant] if the pressure of air behind the constriction is approximately equal to the ambient air pressure according to their amount of sonority, all speech sounds can be ranked where 1 is the highest degree of sonority and 5 is the lowest.

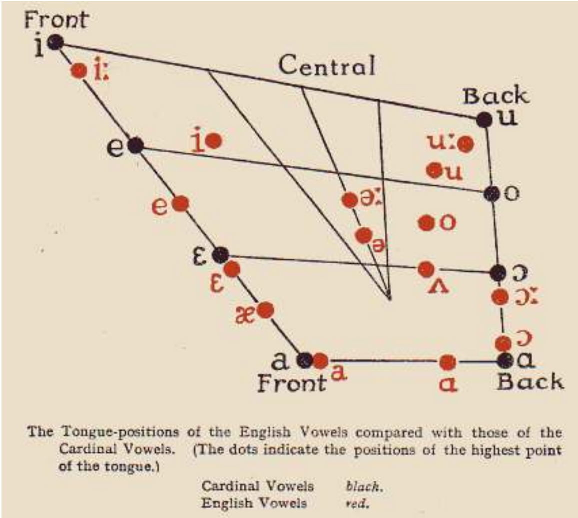
High: **(1)** The three highest (vowels, liquids and nasals) are **[+ sonorant]**

$$\left.\begin{array}{l}\textbf{a}\\ \textbf{m}\\ \textbf{l}\end{array}\right\}$$

$$\left.\begin{array}{l}\textbf{p}\\ \textbf{s}\end{array}\right\}$$

Low **(5)** The two last **(Fricatives and Plosives)** are **[–sonorant]**

Vowels can be defined using four such placement features, if they involve a high tongue position, we refer to them as **[+high]**, if the tongue is in a low position, the feature assigned to them is **[+low]** and if the back of the tongue is involved, this specifies a feature **[+back].** Finally, the distinctive feature **[rounded]** is used for vowels that involve lip rounding .Hence it distinguishes primary cardinal vowels by Jones from the secondary ones by Ladefoged. Here are some examples of the way we proceed.



The Tongue-positions of the English Vowels compared with those of the Cardinal Vowels. (The dots indicate the positions of the highest point of the tongue.)
Cardinal Vowels *black.*
English Vowels *red.*

Figure1.1: Secondary Cardinal Vowels (Ladefoged, 2006p. 44)

/i/ [+high, –low, –back, –round]

/a/ [–high, +low, –back, –round]

/ɒ/ [–high, –low, +back, +round]

We need to add a fifth feature for vowels which is *Advanced Tongue Root* **ATR. [+/–Tense]** this feature mainly applies to the position of the root of the tongue when articulating vowels. **[+Tense]** vowels have an advanced tongue root. In fact, this feature is often referred to as advanced tongue root.

c. Manner of Articulation: features essentially distinguished between stops and fricatives on the one hand and plosives and affricated on the other.

Continuant/Non-continuant: continuant sounds are produced with a primary constriction that does not entirely block the air flow; while the articulation of non-continuant sounds (stops) involve such a complete closure.

Instantaneous Release/Delayed Release is a feature that keeps apart plosives from affricates. It refers then to sounds produced with a complete closure of the tract, but which differ in the manner of the release instantaneous or abrupt in the case of plosives and delayed in the case of affricates. The two features then combine to describe the respective consonant classes. Stops are characterized as [-continuant, +instantaneous release], while fricatives are [+continuants] and affricates are [-continuant, +delayed release]

The feature *Tense/Lax* parallels the feature *Long/Short* in vowels and *Voiceless/Voiced* in consonants. It describes the higher or lower muscular articulatory effort required by the uttering of the respective sound.

d. Source Features: They are related to the source, which is the vocal folds vibration that sustains voiced sounds or a turbulent airstream that sustains voiceless sounds.

Voiced/Unvoiced is a fundamental feature characterizing sounds in any language requiring the vibration or not of the vocal folds.

The *Heightened Subglottal* pressure feature accounts for aspiration in the tense voiceless stops.

The feature *Strident/Non-strident* was described as being marked acoustically by greater (or lower) noisiness and restricted to obstruent continuants and affricates. The dental fricatives of English are non-strident, while the alveolar ones are strident.

Chomsky and Halle made a basic distinction between classificatory and phonetic functions of their distinctive features. Only in their classificatory functions are distinctive features strictly binary. In their phonetic function, distinctive features appear as *non-binary* and receive a physical interpretation. The way in which the conversion from one matrix to the other-from the classificatory to the phonetic, or vice versa- is to be carried out still awaits definitive exploration.

An addition concept important in *The Sound Pattern of English* was a newly introduced concept, that of the *alphabetic symbol.* In accordance with the descriptivist practice, the notation for the alphabetic symbol is written between *slashes* to signify *phonemic status*. Chomsky and Halle defined these *alphabetic symbols* with their conventional interpretation as abbreviation for feature sets loosely as they stated in the SPE. They deemed that they were "*…nothing more than convenient ad hoc abbreviations for feature bundles introduced for ease of printing and reading but without systematic import."*(1968, p.64)

The Chomsky and Halle term, *systematic import*, must be taken as intrinsic to the definition itself. Based on their definition of *alphabetic symbol* and ignoring the use of the notational slashes, the Chomsky–Halle alphabetic symbol seems to resemble more closely the earlier conception of the speech sound, which Trubetzkoy (1938) had defined as "*…the sum of all distinctive as well as non-distinctive phonic properties occurring at a specific point in the sound flow". (p.37)*

As far as the opposition between consonants and vowels is in question, and we have been through different terms used by different phonologists, namely Trubetzkoy and Jacobson. Chomsky and Halle wrote in the Sound Pattern of English:

> *This complete identification of vowel and consonant features [in terms of gravity, compactness, and diffuseness] seems in retrospect to have been too radical a solution…We have therefore made a number of changes in the framework, in particular, with regard to the primary cavity features. The devised framework is quite likely to appear to*

> *depart from the earlier framework much more rapidly than it in fact does.* (1968, p. 303)

Chomsky and Halle continue that:

> *This deceptive impression is the result of the unfortunate need to change terminology once again and to replace the by now reasonably familiar terms compact, diffuse, and grave in part by totally new terms that are a return to the status ante quo.*(1968, p. 303)

Chomsky and Halle (1968) described only the articulatory correlates of their features. Their reason for doing so was not because other aspects (acoustic and perceptual) "*are less interesting or less important, but rather …because it would make the book too long" (*P. 86). It is clear that such an attitude must leave any criticism with respect to the acoustic correlates of compactness and gravity unanswered and in abeyance.

In the Chomsky – Halle feature classification of the SPE, the shape and volume of the resonator cavity used by Jacobson as a base for feature differentiation was replaced by a new parameter. Feature differentiation became based on and related to the body of the tongue and the location of the obstruction in the oral cavity. The point of reference for the tongue was its position in *its neutral state*. However, neutral state of the tongue did not mean its position in its relaxed state, at the floor of the mouth. Chomsky and Halle understood neutral state of the tongue to mean *raised to about the level that it occupies in the articulation of the English vowel /e/.* According to them, their new classification did away with earlier differences between the physical correlates to vowels and consonants for the oppositions *compact/diffuse* and *grave/acute* which had existed in the Jacobson's classification.

The oppositions *compact/diffuse* and *grave/acute* in the Chomsky–Halle classification were partly replaced by the opposition *coronal/non coronal***,** and *anterior/non anterior***.** However, they are not entirely *identical*. These two new Chomsky–Halle features are supplemented by the features *high, low,* and *back*. The table below shows the new Chomsky–Halle feature classification as it relates to the English phonological system, compared to the earlier Jacobson's features as they relate to the phonological features of English.

Chomsky-Halle Classification	Articulatory Correlates	Correspondences in Jacobsonian Classification
Coronal	Blade of the tongue raised	**Acute** –all of these, except palatal sounds, and addition of retroflexes
Alveolar, dentals	From neutral position	
Palatoal veolars, liquids		
produced with the blade of the tongue, retroflexes		
Noncoronal		**Grave** – in part, except that body of tongue also includes palatals, which are acute.
Labials, palatals, uvular	Sounds not produced with the blade of the tongue (lip articulation, body of tongue)	**Diffuse**
R, vowels, glides y and w	Obstruction located in front of palatoal veolar region of mouth	
Anterior	Obstruction not located in front of palatoal veolar	**Compact**, some flat
Labials, dentals,		
Alveolars (front) liquids	Region of the mouth	
Nonanterior		
Vowels, palatoalveolars,		
Retroflexes, palatals,		
Velars, uvulars,		
pharyngeals		

Table1.1: Chomsky–Halle Classification (coronal/noncoronal and anterior/nonanterior) and Correspondences in the Jacobsonian Classification (1978, p.88)

Chomsky and Halle also established the features *high/non-high*, *low/non-low*, and *back/non-back.* All of these are related to the placement of the body of the tongue as well. For the feature high, the tongue is *raised* to a level above its neutral position; for the feature low, it is *lowered* below its neutral position; and for the feature back, it is *retracted* from its neutral position. All five features are used for *consonants* and *vowels*. It needs to be pointed out, however, that Chomsky and Halle did not succeed in establishing a fully systematical system. A relative asymmetry existed between the vowels and the consonants. Vowels were always *non-coronal* and *non-anterior*. The presence of the features [+vocalic], [–consonantal] thus always made the specification non-coronal and non-anterior redundant.

The feature *stridency* was still included in the Chomsky Halle framework of the *Sound Pattern of English*, but its application was restricted.

The feature *distributed* as represented in the *Sound Pattern of English* was, among other things, to provide for differentiation in the pre-palatal region in cases where more than three points of articulation were present. An example of a language where this is the case is Toda, a language of India, which has five stops in the pre-palatal region, /p,˘t, t, t, k/ labial, dental, alveolar, retroflex and palate-alveolar respectively.

Hence, distributed sounds are opposed to non-distributed sounds as:

> *…with a constriction that extends for a considerable distance along the direction of the air-flow, non-distributed sounds are produced with a constriction that extends only a short distance in this direction.* (Chomsky and Halle, 1968, p.312)

Thus, *apical* and *retroflex* stops were opposed to *laminal* and *non-retroflex* stops as [**–distributed]** to **[+distributed].**The feature *delayed release* as opposed to *instantaneous release* Chomsky and Halle described as affecting:

> *…only sounds produced with the closure of the vocal tract…During delayed release turbulence is generated in the vocal tract so that the release phase of the affricates is acoustically quite similar to the cognate fricative*.(1968, p.318)

1.2.1.2.What is a Syllable?

A *syllable* is a unit of organization for a sequence of speech sounds. For example, the word water is composed of two syllables: *wa* and *ter***.** A syllable is typically made up of a syllable nucleus (most often a vowel) with optional *initial* and *final* margins (typically, consonants).

Syllables are often considered the phonological building blocks of words. They can influence the *rhythm* of a language, its prosody, its poetic *meter* and its stress *patterns***.** A word that consists of a single syllable (like English *cat*) is called a *monosyllable* (and is said to be *monosyllabic*). Similar terms include *disyllable* (and *disyllabic***)** for a word of two syllables; *trisyllable* (and *trisyllabic***)** for a word of three syllables; and *polysyllable* (and syllables or to any word of more than one syllable.

1.2.1.2.1. Kinds of Syllables

There are six kinds of syllables in English:

a. **Closed Syllables:** a closed syllable is the one that in spelling contains one vowel and it ends in a consonant. e.g., stretch .ask.
b. **Open Syllables:** an open syllable has in spelling one and only one vowel, and that vowel occurs at the end of the syllable, e.g. no, she, me …etc.
c. **Silent –E -Syllables**: a silent –e syllable ends in spelling in an **e**, has only one consonant before that **e**, and has only one vowel before that consonant. e.g.: ate, ice, tune.

d. **Vowel Combination Syllables:** a vowel combination syllable has in spelling a cluster of two or three vowels or a vowel –consonant unit with a sound or sounds particular to that unit e.g., rain, day, veil, piece, see…etc.

e. **Vowel –R Syllables:** A vowel-r syllable is one which includes in spelling one vowel followed by an 'r', or one vowel followed by an 'r' which is followed by a *silent* e, or a vowel combination followed by an 'r'. e.g. : car, or, care, air, dear…etc.

f. **Consonant-L-Syllables**: In these syllables, a consonant is followed by *le*. The vowel sound in these syllables is the schwa sound that occurs before the/l/.e.g. : words which end with ble, cle, dle, fle, and gle.

1.2.1.3. Syllabification

The term *syllabification* refers to the division of a word into syllables. Syllables are formed by the combination of vowels and consonants. To predict the correct number of syllables within a word, the number of vowels indicates the number of syllables. The syllable structures are of various types and common structure of syllable is the combination of a consonant and a vowel (CV), this syllable structure may be found almost in all languages. For instance in English we can have: CCCVCCCC.

1.2.1.4. Phoneme

A *phoneme* is a basic unit of the phonology of a language, which is combined with other phonemes to form meaningful units such as words or morphemes. The phoneme can be described as "the smallest contrastive linguistic unit which may bring about a change of meaning" (Gimson, ed. Cruttenden, 2008.p. 41). In this way the difference in meaning between the English words *kill* and *kiss* is a result of the exchange of the phoneme /l/ for the phoneme /s/. Two words that differ in meaning through a contrast of a single phoneme form a minimal pair.

1.2.1.5. Allophone

In Phonology, an *Allophone, is a term from* the Greek: *állos*, *other* and φωνή, *phōnē*, 'voice, sound') is one of a set of multiple possible spoken sounds (or phones) used to pronounce the same phoneme. For example, [p^h] (as in *pin*) and [p] (as in *spin*) are allophones for the phoneme /p/ in the English Language. Replacing a sound by another allophone of the same phoneme will usually not change the meaning of a word, although sometimes the result may sound non-native or even unintelligible. Jakobson (1980) states that *"an allophone is the set of phones contained in the intersection of a maximal set of phonetically similar phones and a primary phonetically related set of phones"* (p. 157).

1.2.1.6. Phone

Crystal described the term *phone* as the smallest perceptible discrete segment of sound in stream of speech. From the segmental phonology's angle, phones are the physical realization of phonemes (Crystal 2003, p. 347).

The word *phone* may then refer to any speech sound or gesture considered as a physical event without regard to its place in the phonology of a language: an unanalyzed sound of a language (Loos et al, 1997).This is in contrast to phoneme, which is a set of phones or a set of sound features that are thought of as the same element within the phonology of a particular language (Crystal 1971, p. 180).

In other terms, a phone is a speech segment that possesses distinct physical or perceptual properties and serves as the basic unit of phonetic speech analysis.

1.2.1.7. Mora

Mora (plural *morae* or *moras*; often symbolized as**μ**) is a unit in phonology that determines syllable weight, which in some languages determines stress or timing. The definition of a mora varies. Perhaps the most succinct working definition was provided by the American linguist McCawley (1968) as "*something of which a long syllable consists of two and a short syllable consists of one*". The term comes from the Latin word for *linger*, *delay*, which was also used to translate the Greek word *chronos* (time) in its metrical sense.

Hence, a syllable containing one mora is said to be *monomoraic*; a syllable with two morae is said to be *bimoraic*. Also, in few cases, a syllable with three morae is said to be *trimoraic*.

1.2.1.7.1. Morae Formation

In general, morae are formed as follows:

- A syllable onset (the first consonant or consonants of the syllable) does not represent any mora.
- The syllable nucleus represents one mora in the case of a short vowel, and two morae in the case of a long vowel or diphthong. Consonants serving as syllable nuclei also represent one mora if short and two if long.
- In some languages (for example, Japanese), the coda represents one mora, and in others (for example, Irish) it does not. In English, the codas of stressed syllables represent a mora (thus, the word *cat* is bimoraic), but for unstressed syllables it is not clear whether the codas do so (the second syllable of the word *rabbit* might be monomoraic).

- In some languages, a syllable with a long vowel or diphthong in the nucleus and one or more consonants in the coda is said to be *trimoraic*

In general, monomoraic syllables are said to be *light syllables*, bimoraic syllables are said to be *heavy syllables*, and trimoraic syllables (in languages that have them) are said to be *super heavy syllables*.

1.1.1. Branches of Phonetics

Phonetics, which is definitely a scientific study, and apart from its well-known three branches, which deal with the *production*, *transmission* and *perception* of speech sounds. Several branches of phonetics can further be distinguished, depending on the narrower domain of interest of the respective field. Other branches of paramount importance should not be ignored. According to Vaissière (2007), the branches of phonetics include Psycho-phonetics, Orthophonic phonetics, Developmental phonetics, Neuro-phonetics, Clinical phonetics and Statistical or Computational phonetics.

a. Psycho-phonetics: deals with the sensations evoked by the sounds and the ones coming after. For instance, the /i/would evoke the colour *yellow* and the /r/ would be perceived as the strongest and masculine than /l/ by listeners from diverse languages.

b. Orthophonic Phonetics: reeducation and didactics, studies the means of correcting the pronunciation’s imperfection of children. In France, for instance, 15 to 20% of children visit frequently orthophonists because of their speech pathologies and disorders.

c. Developmental Phonetics: It is very close to psycholinguistics; its main interest is the reactions of the fetus to different sonic or resonant stimulus, to the acquisition process (perception and production) of the segmental and prosodic features related to the mother tongue by the baby , then the infant , whether a monolingual or a bilingual.

d. Clinical Phonetics: clinical phonetics is at the crossroad of linguistics and medicine. Clinical Phonetics' main interests are speech pathology, including voice disorders, speech therapy, speech perception disorders,Cochlear implants and how they work and patients with cancer and who have had surgery.

e. Statistical Phonetics or Computational: Phonetics is the study of speech sounds; it investigates the physical properties of the sound waves constituting the actual speech signal and shares a great deal of notion and tools with *standard* Signal Processing. Computational phonetics is the design of algorithms implementing phonetic analysis methods; the use of software requires a very good mastery of computer sciences to be able to work with scripts and spectrograms.

f. Articulatory Phonetics: Articulatory phonetics is a branch of phonetics which is largely based on data provided by other sciences, the most important among which are human anatomy and physiology. This is a result of the fact that human beings do not possess organs that are exclusively used to produce speech sounds; all organs involved in the uttering of sounds have in fact, primarily, other functions: digestive, respiratory, etc.. This actually raises interesting questions about whether we were born (destined, programmed) to speak or speech developed rather accidentally – anyway, comparatively later – in the evolution of mankind. Therefore, fundamental physiological processes like those mentioned above take place simultaneously or alternatively with the production of speech sounds. We can hardly think of speaking as being separated from the activity of breathing, as the air that is breathed in and out of the lungs has a crucial role in the process of uttering sounds.

Breathing is a rhythmic process including two successive stages: inspiration and expiration. It is during the latter phase that speech production takes place in most languages. Because we speak while we expel the air from our lungs, the sounds that we produce are called *egressive*. The continuous alternation between *inspiration* and e*xpiration* fundamentally shapes the rhythmicity of our speech.

We have already mentioned the fact that oral communication is based on sound waves produced by the human body. The initial moment of this rather complex process is the expelling of the air from our lungs. The lungs can therefore be considered the very place where speech production originates. The airstream follows a road that is called the vocal tract. We will follow this tract of the air that is expelled from the lungs out of the body. As we are going to see, this tract includes segments of the respiratory and digestive tracts and the physiology of speaking is therefore intimately linked to the physiology of the respective vital processes. The lungs are pair organs, situated inside the thoracic cavity (the chest). They are formed of three, respectively two spongy lobes (the left lung is smaller because of the vicinity of the heart within the thoracic cavity). The capacity of the lungs (that is the total amount of air that they can contain is of about 4500-5000 cm (4.5-5 liters) in an adult person, the capacity being generally slightly superior in the case of male persons. The so-called vital capacity (that is the maximal amount of air that can be exchanged with the environment during breathing is of about 3500-4000 cm. In other words, we can never completely empty our lungs of air during expiration. During normal breathing, however, only about 10-15% of the vital capacity is used, that is the quantity of air that is exchanged amounts to about 400-500 cm^3. The act of speaking requires a greater respiratory effort and consequently the amount of air increases to up to 30-80% of the vital capacity (30-40% during expiration and 45-80% during inspiration). Variations are due to different position of the body, to the quality, quantity and intensity (loudness) of the sounds we articulate. Breathing is a complex process that essentially consists in the exchange of air between our body and the environment. It leads to the oxygenation of our body and to the expulsion of the carbon dioxide resulting from the processes of combustion within our body. It is basically achieved by the successive expanding and compressing of the volume of the two lungs, the air being sucked in and pushed out respectively. This happens because the thoracic cavity itself modifies its volume, a complex system of bones (the ribs), muscles (the most important of which are the intercostal ones, that coordinate the movements of the ribs, and the diaphragm, that represents the floor of the thoracic cavity) and membranes (pleurae) being involved in the process. The entire process is controlled by the respiratory canters in the brain.

From each of the lungs a bronchial tube starts. At one end, the ramifications of these tubes spread inside the spongy mass of the pulmonic lobes. They are called *bronchioles* and their role is to distribute and collect the air into and from the innermost recesses of the lungs. These exchanges are made at the level of small air sacs called *alveoli* and represent the ultimate ramifications of the bronchioles. At the other end, the two bronchial tubes are joined at the basis of the *trachea*, or the *windpipe*.

The windpipe has a tubular cartilaginous structure (its components are a number of cartilages held together by membranous tissue) and is about 10 cm long and 2.5 cm in diameter. Its elasticity and the position of the larynx can result in important variations in the actual length of the organ. The latter is an essential segment of the respiratory system but does not play an active role in speech production.

Another organ that has a crucial role in the process of speaking is the larynx. It is a cartilaginous pyramidal organ characterized by structural complexity and situated at the top of the trachea. As all speech organs, it primarily performs a vital role, namely it acts as a valve that closes, thus blocking the entrance to the windpipe and preventing food or drink from entering the respiratory ducts while we are eating. They are instead directed down the pharynx and the esophagus. The larynx is the first speech organ proper along the tract that we are following, as it interferes with the outgoing stream of air (which, so far, has followed its way rather unimpeded) and establishes some of the essential features of the sounds that we produce. However, it is not the larynx proper (that is the organ in its entirety) that performs this important role within the speech mechanism, but two muscular folds inside it, called *the vocal folds*.

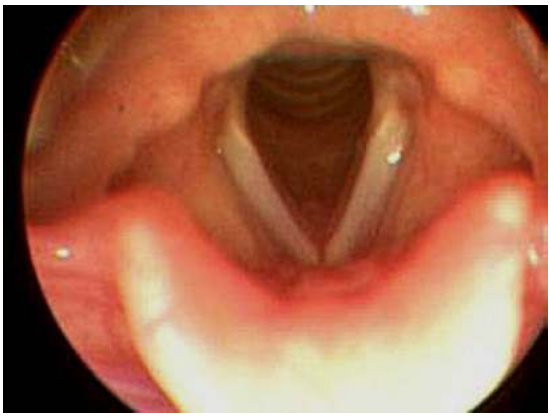


Figure1.2.: **Vocal Folds**[3]

As mentioned above, the larynx consists of a number of cartilaginous structures that interact in an ingenious way enabling the larynx to perform its important respiratory and articulatory functions. The thyroid cartilage is made of two rectangular flat plates (left and right) at form an angle anteriorly, resembling the covers of a book that is not entirely open.

[3] http://www.svas.com.au/normal-voice-function/

The aperture of the angle, oriented posteriorly, varies with gender. It is a right angle in men (90°) while in women it is 120°. The angle is more visible, because more acute, in the former situation and the cartilage is popularly known as *Adam's apple*. Posteriorly, each of the plates has two horns (an inferior and a superior one) called cornua, through which the thyroid cartilage is connected with the cricoid one. The joint that the two cartilages form, resembling a sort of hinges, allows the cricoid one to move anteriorly and posteriorly with respect to the thyroid one, thus controlling the degree of tension in the vocal folds.

One of the main functions of the thyroid cartilage is to protect the larynx and particularly the vocal folds. The cricoid cartilage is made of a ring-shaped structure, situated anteriorly and of a blade situated posteriorly and represents the base of the larynx, controlling communication with the trachea. On top of its blade, on the left and right side respectively, another pair of cartilages is situated: the arytenoid ones. The last important cartilage in the process of phonation or speech production is the epiglottis which is a spoon-shaped cartilage also playing an important role in keeping the food away from the respiratory tract. It is between the arytenoid cartilages and the thyroid cartilage that the two vocal cords mentioned above stretch. The vocal cords are each made of what is known as a vocal ligament and a vocal muscle. They are covered in mucous membranes or skin folds also known as the vocal folds. They connect the lower part of the thyroid cartilage to the anterior part of the arytenoid cartilages. The opening between the folds and the arytenoid cartilages represents the glottal aperture, more commonly called the glottis. The length of the vocal folds varies with age and gender. They become longer at the age of puberty and are longer in men (24-26 mm) than in women (17-20 mm). During breathing, the two folds part, letting the air come into the larynx or go out. During phonation they come closer, having an important role in establishing some of the main characteristics of the articulated sounds. By the very complex action of adjacent anatomical structures (the cartilages described above and a number of laryngeal muscles) the two vocal cords can be brought together or parted. They thus interfere to various extents with the outgoing airstream. They can obstruct the passage completely, as in the case of the glottal stop, or their participation in the uttering of a given sound can be minimal (as in the case of many hissing sounds). The rapid and intermittent opening and closing of the vocal cords, which results in the vibration of the two organs, plays a key role in one of the most important phonetic processes, that of voicing. Thus, vowels and vowel-like sounds, as well as a number of consonants, are produced with the vibration of the vocal folds and are consequently voiced. The absence of vibration in the vocal folds is characteristic for voiceless obstruents. The amplitude of the vibration is also essential for the degree of loudness of the voice: thus the intensity of the sound that is uttered depends on the pressure of the air that is expelled. The rate at which the vocal cords vibrate has also important consequences as far as the pitch of the voice is concerned; this is closely linked to the pressure exerted on the vocal folds. When we produce acute (high-pitched or shrill) sounds the vocal folds come closer to each other, while during the articulation of grave sounds the vocal folds leave a greater space between them. The next organ in the vocal tract is the pharynx, situated at a kind of crossroads along the above-mentioned tract. It does not play an active part in the articulation of sounds; its main role is to link the larynx and the rest of the lower respiratory system to its upper part, thus functioning as an air passage during breathing. It is also an important segment in the digestive apparatus as it plays an essential role in deglutition (the swallowing of food).

The pharynx branches into two cavities that act as resonators for the air stream that the vocal folds make vibrate: the *nasal cavity* and the *oral cavity*.

Before discussing the two cavities, it is important to mention the role played during articulation by the velum or the soft palate. The velum is the continuation of the roof of the mouth also called the palate. The harder, bony structure situated towards the exterior of the mouth continues with the velum into the rear part of the mouth. The latter's position at the back of the mouth can allow the airstream to go out through either the mouth or the nose or through both at the same time. Thus, if the velum is raised, blocking the nasal cavity, the air is directed out through the mouth and the sounds thus produced will be oral sounds. If the velum is lowered, we can articulate either nasal sounds, if the air is expelled exclusively via the nasal cavity, or nasalized sounds if, in spite of the lowered position of the velum, the air is still allowed to go out through the mouth as well as through the nose. If we nip our nostrils or if the nasal cavity is blocked because of a cold, hay fever, etc., we can easily notice the importance of the nasal cavity as a resonator and the way in which its blocking affects normal speech production. The distinction nasal/oral is essential in all languages and it will further be discussed when a detailed analysis of both English consonants and vowels is given.

We have mentioned above the oral cavity as one of the two possible outlets for the airstream that is expelled by our respiratory system. The oral cavity plays an essential role in phonation, as it is here that the main features of the sounds that we articulate are uttered. The cavity itself acts as a resonator, and we can modify its shape and volume, thus modifying the acoustic features of the sounds we produce, while various organs that delimit the oral cavity or are included in it (the tongue) are active or passive participants in the act of phonation. If we follow the airstream out through the mouth (oral cavity) we can easily notice the above-mentioned organs that play an important role in the process of sound articulation.

Undoubtedly, the most important of all is the tongue, which plays a crucial role in oral communication, the very fact that in many languages (Greek, Latin, Romance languages) the same word is used to refer to both the anatomical organ and language as a fundamental human activity showing that in many cultures the two concepts came to be assimilated or at least considered to be inseparable. The tongue is actually involved in the articulation of most speech sounds, either through an active or a comparatively more passive participation. It is a muscular, extremely mobile and versatile organ (by far the most dynamic of all speech organs) and it plays an essential role in the producing of consonants, while its position in the mouth is also very important for differentiating among various classes of vowels. When an articulatory classification of speech sounds is given below, the upper surface of the tongue will be divided, for practical and didactic purposes, into several parts: a) its fore part, made up of the tip (*apex*) and the *blade*; b) the front, and back part (the dorsum) – the label dorsum is often applied to front and back together – and c) the root(*radix*) of the tongue (the rearmost and lowest part of the organ, situated in front of the laryngo-pharynx and the epiglottis. The sides or rims of the tongue also play an important role in the uttering of certain sounds. (As we are going to see, the various parts of the tongue lend their names to the sounds they help produce: thus, sounds uttered with the participation of the tip of the tongue will be called *apical*– from the Latin word *apex*, meaning top or extremity – those in the production of which the blade is involved will be called *laminal*– from the Latin word *lamina* having the same meaning – while the back part of the body of the tongue, *the dorsum*, will give its name to *dorsal sounds*, produced in the velar region.) The tongue is a movable articulator (the term active is usually used) that influences the way in which sounds are produced. But more often than not it does that with the help of other articulators (fixed or mobile i.e. passive or active) as well, like the roof of the mouth (the palate), the lips or the teeth. The palate essentially consists of two parts: the hard palate and the soft palate or the velum. We have shown the important role played by the velum in differentiating between the articulation of oral and nasal sounds. The hard palate in front of it functions as a fixed (passive) articulator. Not less important are, at the other end of the mouth, the teeth and the lips. Just behind the teeth we can notice the alveolar ridge (the ridge of the gums of the upper teeth). While the upper teeth are fixed, the lower jaw (the mandible) is mobile and its constant moving permanently modifies the size and shape of the oral aperture. The lips also play an important role in the articulation of some consonants by interacting with each other or with the upper teeth and their position (rounded or spread) is also relevant for differentiating between two major classes of vowels. They are pretty mobile articulators, though far less so than the tongue. Just like the tongue, they can yield a variety of configurations. The lower lip can cooperate with the upper teeth to produce labio-dental sounds, the two lips can interact to articulate bilabial sounds, while lip position (rounded or spread) is essential in determining one of the basic configurations of vowels.

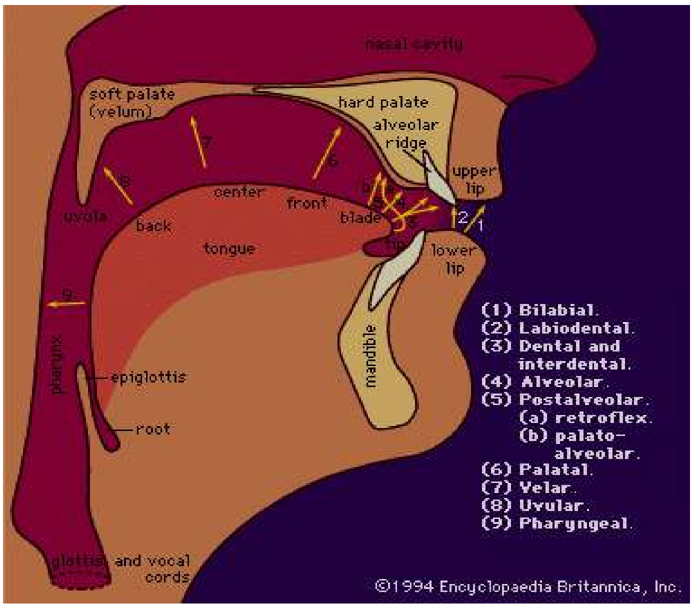


Figure 1.3.: Anatomy of the Vocal Tract (Articulators) (from the Encyclopedia Britannica, 1994)

Our brief schematic presentation of speech production has consciously neglected the essential role the brain plays in the articulation of sounds. Speech production is a process that can be observed quite easily as the major articulators lend themselves to direct and detailed scientific observation. We should not forget, however, that our presentation above is obviously partial, since all articulatory processes are controlled by the brain and we cannot imagine any kind of activity of the articulators without the participation of the brain that actually controls the entire process of speech production. We chose to leave aside the discussion of the part played by the brain in the physiology of articulation only because the complexity of the analysis would have taken us too far away from the purpose of this study.

1.1.1.1. Auditory Phonetics

If articulatory phonetics studies the way in which speech sounds are produced, auditory phonetics focuses on the perception of sounds or the way in which sounds are heard and interpreted. Most of us are able to understand and produce speech so easily that we may not realize how complicated these processes are. In fact, speech is so complicated that over the decades in which it has been studied, no scientist has been able to invent a machine or computer that can recognize speech, as well as we do. Remembering our conventional division of linguistic communication into several stages of a process unfolding between two parties, the sender of the message and its addressee, we may say that while articulatory phonetics is mainly concerned with the speaker, auditory phonetics deals with the other important participant in verbal communication, the listener. It is again, obviously, a field of linguistic study which has to rely heavily on biology and more specifically on anatomy and physiology.

We should say from the very beginning, however, that the mechanism and physiology of sound perception is a much hazier field that the corresponding processes related to the uttering of the respective sounds. This is so because speech production is a process that takes place roughly along the respiratory tract which is, comparatively, much easier to observe and study than the brain where most processes linked to speech perception and analyses occur. We are in fact dealing with two distinct operations which, however, are closely interrelated and influence each other: on the one hand, we can talk about audition proper, that is the perception of sounds by our auditory apparatus and the transforming of the information into a neural sign and its sending to the brain and, on the other hand, we can talk about the analysis of this information by the brain which eventually leads to the decoding of the message, the understanding of the verbal message.[4]

When discussing the auditory system we can consequently talk about its peripheral and its central part, respectively. We shall have a closer look at both these processes and try to show why they are both clearly distinct and at the same time they are closely related.

Before the sounds we perceive are processed and interpreted by the brain, the first anatomical organ they encounter is *the ear* (See below). The ear has a complex structure and its basic auditory functions include the perception of auditory stimuli, their analysis and their transmission further on to the brain. We can identify three components: the outer, the middle and the inner ear. The outer ear is mainly represented by the auricle or the pinna and the auditory meatus or the outer ear canal. The auricle is the only visible part of the ear,

[4]It is obvious, even at an intuitive level, that hearing someone's words and understanding them are two quite different processes. They are chronologically distinct, since we can talk about two successive stages, and they also differ in nature since they involve different operations of the brain. We should not make the mistake, however, of separating them completely since, as we are going to see below, our understanding of what we hear essentially influences the process of hearing itself.

constituting its outermost part, the segment of the organ projecting outside the skull. It does not play an essential role in audition, which is proved by the fact that the removing of the pinna does not substantially damage our auditory capacity. The auricle rather plays a protective role for the rest of the ear and it also helps us localize sounds. The meatus or the outer ear canal is a tubular structure playing a double role: it, too, protects the next segments of the ear – particularly the middle ear – and it also functions as a resonator for the sound waves that enter our auditory system. The middle ear is a cavity within the skull including a number of little anatomical structures that have an important role in audition. One of them is the eardrum. This is a diaphragm or membrane to which sound waves are directed from outside and which vibrates, acting as both a filter and a transmitter of the incoming sounds. The middle ear also contains a few tiny bones: the mallet, the anvil and the stirrup. The pressure of the air entering our auditory system is converted by the vibration of the membrane (the eardrum) and the elaborate movement of the little bones that act as some sort of lever system into mechanical movement which is further conveyed to the oval window, a structure placed at the interface of the middle and inner ear. As pointed out above, the middle ear plays an important protection role.

The muscles associated with the three little bones mentioned above contract in a reflex movement when sounds having a too high intensity reach the ear. Thus the impact of the too loud sounds is reduced and the mechanism diminishes the force with which the movement is transmitted to the structures of the inner ear. It is in the middle ear too, that a narrow duct or tube opens. Known as the *Eustachian tube* it connects the middle ear to the pharynx. Its main role is to act as an outlet permitting the air to circulate between the pharynx and the ear, thus helping preserve the required amount of air pressure inside the middle ear. The next segment is the inner ear, the main element of which is the cochlea, a cavity filled with liquid. The inner ear also includes the vestibule of the ear and the semicircular canals. The vestibule represents the central part of the labyrinth of the ear and it gives access to the cochlea. The cochlea is a coil-like organ, looking like the shell of a snail. At each of the two ends of the cochlea there is an oval window, while the organ itself contains a liquid.

Inside the cochlea there are two membranes: the vestibular membrane and the basilar membrane. It is the latter that plays a central role in the act of audition. Also essential in the process of hearing is the organ of Corti, inside the cochlea, a structure that is the real auditory receptor. Simplifying a lot, we can describe the physiology of audition inside the inner ear as follows: the mechanical movement of the little bony structures of the middle ear (the mallet, the anvil and the stirrup) is transmitted through the oval window to the liquid inside the snail-like structure of the cochlea; this causes the basilar membrane to vibrate: the membrane is stiffer at one end than at the other, which makes it vibrate differently, depending on the pitch of the sounds that are received. Thus, low-frequency sounds will make vibrate the membrane at the less stiff (upper) end, while high-frequency (acute) sounds will cause the lower and stiffer end of the membrane to vibrate. The cells of the organ of Corti, a highly sensitive structure because it includes many ciliate cells that detect the slightest vibrating movement, convert these vibrations into neural signals that are transmitted via the auditory nerves to the central receptor and controller of the entire process, the brain.

The way in which the human brain processes auditory information and, in general, the mental processes linked to speech perception and production are still largely unknown. What is clear, however, regarding the perception of sounds by man's auditory system, is that the human ear can only hear sounds having certain amplitudes and frequencies. If the amplitudes and frequencies of the respective sound waves are lower than the range perceptible by the ear, they are simply not heard. If, on the contrary, they are higher, the sensation they give is one of pain, the pressure exerted on the eardrums being too great. These aspects are going to be discussed below when the physical properties of sounds are analyzed. As to the psychological processes involved by the interpretation of the sounds we hear, our knowledge is even more limited. It is obvious that hearing proper goes hand in hand with the understanding of the sounds we perceive in the sense of organizing them according to patterns already existing in our mind and distributing them into the famous acoustic images of which Saussure spoke. It is at this level that audition proper intermingles with psychological processes because our brain decodes, interprets, classifies and arranges the respective sounds according to the linguistic (*phonological*) patterns already existing in our mind.[5]It is intuitively obvious that if we listen to someone speaking an unknown language it will be very difficult for us not only to understand what they say (this is out of the question given the premise we started from) but we will have great, often insurmountable difficulties in identifying the actual sounds the person produced. The immediate, reflex reaction of our brain will be to assimilate the respective sounds to the ones whose mental images already exist in our brain, according to a very common cognitive reaction of humans that always have the tendency to relate, compare and contrast new information to already known information.

[5]Listening comprehension is, after all, an essential part of our mastering a certain language. In order to actually understand what someone says in a given language we need to be accustomed to the phonological structure of the respective language. This will enable us to correctly interpret – acoustically speaking – even phonological structures whose meaning is unknown to us. Conversely, when someone mispronounces a word we know – say, in our own mother tongue – our mind will automatically correct the mistake and we will be able to understand the word the speaker meant to utter in spite of its actual faulty pronunciation.

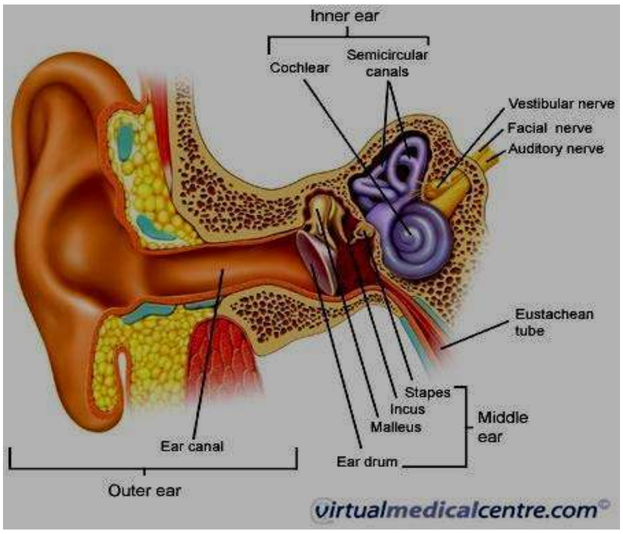


Figure 1.4: Anatomy of the Ear (from: virtualmedicalcentre.com)

1.1.1.7.1. Hearing vs. Listening

Although hearing is a complex process, it is essentially an automatic, passive activity. It is possible to hear sounds without consciously engaging in the process. Below is a step-by-step description of how we hear:

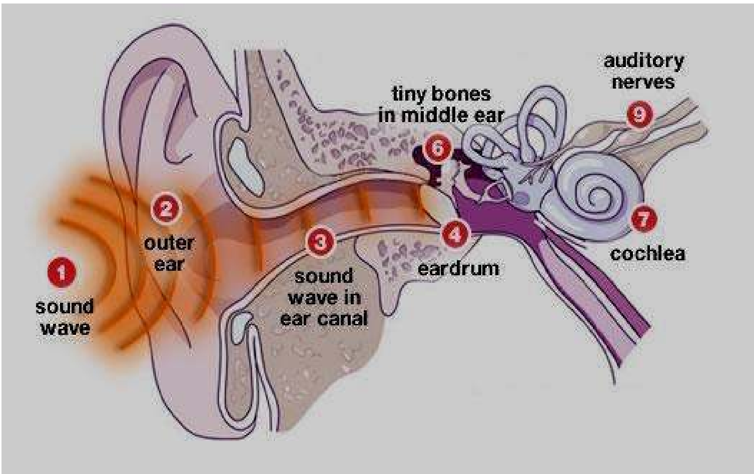


Figure1.5.: Parts of the Ear (virtualmedicalcentre.com)

i. Something vibrates and creates a sound wave. Because a sound comes out of a vibrating object.
ii. The sound wave penetrates the ear and is collected by the outer ear or pinna is called Auricle as well.
iii. The sound wave then moves into the ear canal.
iv. When it reaches the end of the ear canal, the sound waves bump up against the eardrum.
v. The ear drum vibrates with these sound waves.
vi. The vibration moves tiny bones in the middle ear.
vii. These bones carry vibrations into the inner ear to a fluid-filled tube called the cochlea.
viii. The fluid inside the cochlea starts to move quickly, vibrates a series of tiny hairs called cilia, which are attached to auditory nerves.
ix. The movement of these cilia stimulates the nerve cells, and they send some electrical impulses signals to the brain via the auditory nerve.
x. The brain processes these signals into the sounds we hear.

Step ten is where the brain's role is to decode and identifies those sound vibrations as familiar sounds or words. The brain doesn't automatically translate these words into the message they are conveying. That is essentially what listening is – determining the meaning and the message of the sounds or words. It is an active process that involves much more than assigning labels to sounds or words.

The study of this aspect of language comprehension, known as *speech perception*, draws from many disciplines. Major contributions have come from such diverse fields as linguistics, acoustics, speech science, psychology…etc.. Although a complete theory of speech perception is not at hand, much has been learned about the physical properties of speech and the way in which these physical properties are processed by human listeners.

One of the major findings has been that the recognition of speech sounds is far from being simple and straightforward, before moving to the different theories of speech perception; let us briefly explain the way speech takes place in the brain.

1.1.1.7.2. Speech in the Brain

Localization or modularity of function refers to the idea that certain areas of the brain have specific designated functions. For example, the left hemisphere seems to be involved in analytic thinking, whereas the right hemisphere is more concerned with holistic processing. For the majority of people, most language functions are processed in the left hemisphere with secondary language functions (understanding figurative language) processed in the right hemisphere.

While brain localization of function for speech suggests that human speech may be special, there are areas in the brains of marmosets that respond more to these primates' social cries. Other species may have sensitivities to their own form of social communication.

But localization of function can get much more specific than just global functional differences between the right and left hemispheres, especially when it comes to speech perception. It has been found that when patients have brain damage in a certain area toward the back of the left hemisphere in the temporal lobe (*Wernicke's Area*), they have problems comprehending speech. Patients with damage further forward in the left frontal lobe (*Broca's Area*) tend to have problems with speech production. These patients are said to have *aphasia* - impairments in either comprehension (*Wernicke's Aphasia*) or speech production (*Broca's Aphasia*). The interaction below shows the location of the brain damage and the speech characteristics for these two types of patients.

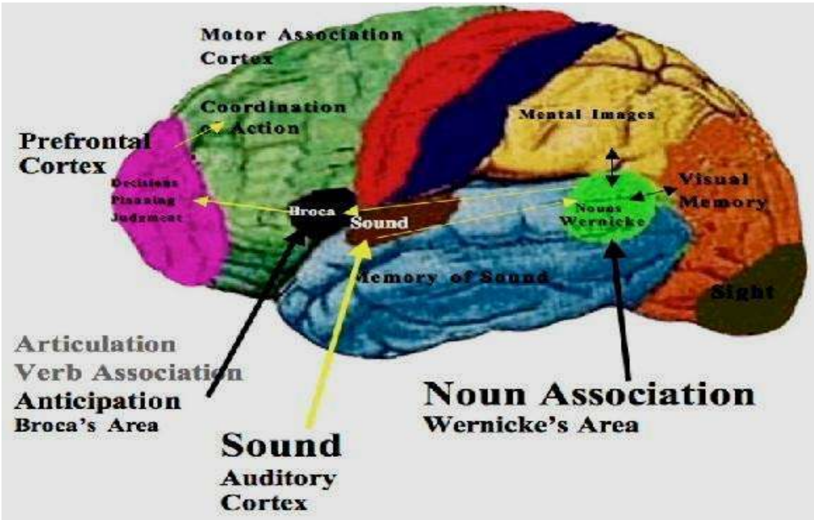


Figure 1.6.: Wernicke and Broca's Areas[6]

1.1.1.7.2.1. Wernicke's Area

Wernicke was a German doctor who, in the 1870s, reported that damage to this part of the brain was found among patients who had speech comprehension difficulties. This finding confirmed the left hemisphere location of language ability and led to the view that Wernicke's area is part of the brain crucially involved in the understanding of speech (cited in Yule, 2006, p.159)

People with Wernicke's aphasia have trouble with language comprehension or speech perception. Hence, they speak fluently, but what they say lacks content and makes little sense. If very severe they may only produce jargon- unintelligible language with adult like *intonation*.

[6]https://shilorives.wordpress.com/2014/03/30/brain-anatomy-featuring-the-brocas-and-wernickes-areas/s

1.1.1.7.2.2. Broca's Area

Broca, a French surgeon, reported in the 1860s that damage to this specific part of the brain was related to extreme difficulty in producing speech. It was noted that damage to the corresponding area on the right hemisphere had no such effect. This finding was first used to argue that language ability must be located in the left hemisphere and since then has been treated as an indication that Broca's area is crucially involved in the production of speech.

People with Broca's aphasia have damage to the frontal lobe. They have difficulty with speech production, producing slow labored speech and having poor articulation.

Speech perception is a term broadly used to refer to how an individual understands what others are saying. More precisely, speech perception is viewed as the way a listener can interpret the sound that a speaker produces as a sequence of discrete linguistic categories such as phonemes, syllables, or words. In dealing with the perception of speech, some interesting questions can be raised for instance:

- Is human speech perception different from other forms of communication?
- Is speech the defining feature that sets humans apart from computers and animals?

To answer these questions, we would need to determine if speech is perceived differently from other auditory stimuli. If so, it suggests a special mechanism in humans for speech perception, since no other animal has speech.

Since speech is a continuous stream of ideas, and when uttering a given word, such as *suit*, we do not produce each sound or segment independently; that is, first [s], then [ʊ], and then [t], with the next sound, beginning only after the previous sound has been completed. This would in fact make speech production very slow and laborious and looks like the one of robots. Instead, when producing the [s] sound, we are already shaping our articulators (tongue, jaw, lips, and so forth) in preparation for the [ʊ] sound and we are even preparing to produce the last segment [t]. This means that the articulatory movements for the different sounds within a word overlap one another in time. That is, the phonetic segments are not individually articulated but are rather coarticulated. Due to this coarticulation, individual phonetic segments do not correspond in a simple way to single distinctive acoustic segments or properties in the speech signal. Therefore, there are two basic ways in which coarticulation complicates the mapping between the phonetic segments intended by the speaker and the speech signal. These are the invariance and the segmentation.

1.1.1.7.3. Acoustic Phonetic Invariance

Acoustic phonetic invariance is the idea that a part of the acoustic signal for individual phonemes remains consistent across various contexts and/or different speakers. This part of the signal would be the invariant acoustic cue and is what we would supposedly process to perceive speech. Researchers have been searching for invariant acoustic cues. The acoustic signal can also be represented as a short term spectrum which graphs frequencies for a brief period of time.

1.1.1.7.4. Segmentation

A major tool in the study of speech perception is the sound spectrograph, a device that provides a visual representation of speech, called spectrogram. In a spectrogram the frequency composition of the sound is displayed across time, frequency is represented on the ordinate (y-axis) and time on the abscissa (x-axis). As an example, a spectrogram of the consonant vowel syllable [sʊ] below, produced by an English male speaker, clearly illustrates the changes in energy pattern across the duration of the syllable. What is most striking is that there are two quite distinct acoustic segments, labeled on the figure below as A and B which correspond in a straightforward way to the two phonetic segments, [s]and [ʊ] respectively.

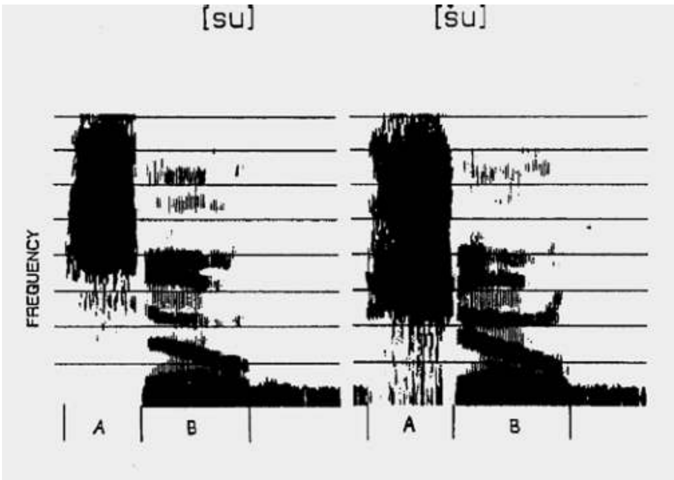


Figure1.7: Spectrogram of [su] and [ʃ̥u],Produced by a Male Speaker of English (Miller, 1990, P.72)

1.1.1.7.5. Theories of Speech Perception

One prominent theoretical position, most closely identified to Liberman (1985) and his colleagues at Haskins Laboratories, is that phonetic perception is accomplished by a processing system that is specialized for the perception of speech. That is to say, listeners possess this specialized system by virtue of their biological status as human, just as an example, bats possess specialized systems for echo location because they are blind, owls possess specialized systems for localizing sound in space. This view has been called the

Motor theory of speech perception, which has three major components, let us explain each one respectively

a. Perception is based on production
b. Perception is species- specific
c. Perception could be innate

According to Liberman's *Motor theory*, the *perception is based on production*, that is to say, there is a close link between the system responsible for perceiving and the system responsible for producing speech. This special link permits the listener, upon hearing the speech signal, to determine which articulatory gestures the speaker has made, and by doing so, to determine which phonetic segments (vowels and consonants) have been produced. The main idea is that we perceive speech by virtue of our tacit (unconscious) knowledge of how speech is produced. We unconsciously know how coarticulation takes place between the phonetic segments intended to be produced.

The second component in Liberman's Motor theory is that *perception is species – specific*; this major component follows directly the first, since only human beings produce speech and have knowledge about how articulation works, and since perception operates in terms of this knowledge of articulation, only humans can actually perceive the phonetic structure of an utterance. Animals will hear speech as a series of noises of some sort; since they do not possess the specialized speech processing system, therefore they will not perceive speech as a structured series of consonants and vowels.

The third major component of the theory is that *Perception is Innate*; that is to say that it is part of our biological heritage as humans. It is further mentioned that this innately given processing system is early operative in infancy. According to the theory, infants come to the world with tacit knowledge of articulation and the production perception link. This enables them to perceive the phonetic structure of an utterance (the sequence of segments) long before they can actually produce speech, and long before language acquisition occurs. In other words, we do not have to *learn* to perceive the sounds of speech because we perceive them innately.

On the other hand, we have a completely opposite and striking contrast theory to the Motor theory, which is the auditory theory of speech perception. Unlike the Motor Theory, this view is not identified with any single researcher or laboratory. Let us consider the basic components of the auditory theory by contrasting it with the motor theory on the three points previously mentioned.

i. *Perception is not based on production*: according to the auditory theory of speech perception, we do not perceive speech via a specialized system that refers to tacit knowledge of articulation. Instead, the auditory system itself does the task. Somehow, because of the way in which our auditory system processes sound, all types of sound (speech and non-speech) the complications caused by coarticulation are' automatically resolved', such that we are able to perceive the ordered sequence of phonetic segments intended by the speaker.

ii. *Perception is not species specific*: since the human auditory system is very similar to the auditory system of many other animals, these animals should perceive speech in much the same way we do. Thus, according to the auditory theory, the ability to perceive speech is not a unique accomplishment which is specific to human species.

iii. *Perception could be innate:* this is to the extent that the auditory system is well developed in infancy; this view suggests that the speech perception abilities should be operative early in infancy. Thus, the auditory theory and the motor theory can both accommodate early speech processing abilities in infancy, though for very different reasons. We will see some evidences showing the experiments used as far as speech perception is concerned. In this scope, we need to deal as well with the categorical perception, which is the core of speech perception.

1.1.1.7.6. Categorical Perception

It turns out that we perceive speech in categories and not continuously. This means that we place sounds that are similar enough into a single category and have a hard time distinguishing between them. For example, /p/ and /b/ share the same place of articulation, manner of articulation, but differ in that /p/ is *voiceless* and /b/ is *voiced*. Both sounds have a different Voice Onset Time (VOT), which refers to the time from the start of a sound until voicing begins or what we refer to as the vocal folds' vibration. The different VOT can be clearly seen on a spectrogram. Generally, listeners identify all these sounds as either /pa/ or /ba/ with sounds closer to the left (which means shorter VOT's) identified as /ba/and those to the right (longer VOT's) identified as /pa/. Thus by increasing the VOT starting with /ba/ the sound becomes gradually /pa/.

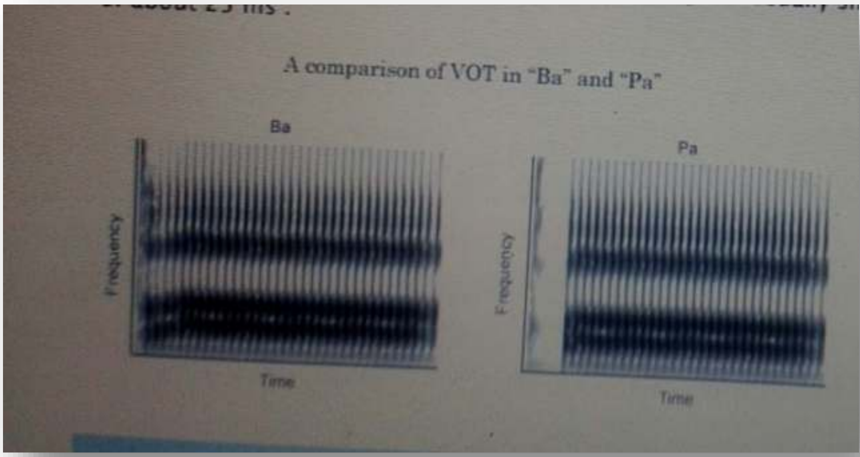


Figure1.8: VOT in /ba/ and /pa/ (Voice Onset Time)

The **VOT,** in which people shift from reporting **/ba/** to **/pa/**, is known as the *phonetic boundary*. Even though there is a continuous change in the sound presented, they are being categorically perceived as one of two different sounds. The same findings occur on a *discrimination test* when two stimuli with different VOT's are presented and listeners are asked to identify the syllables as the *same* or *different*. Listeners report all sounds that are on the same side of the phonetic boundary as the same, while sounds on the other side of the boundary are reported as different.

1.1.1.7.6.1. Is Categorical Perception Innate?

An experimental design was conducted by Eimas and his colleagues (1971) adapted a technique for use with speech, which came to be known as the high amplitude sucking (HAS) procedure. The technique is based on the young infant's proclivity to seek novel stimulation, and worked as follows: the infant was placed in an infant seat in front of a loudspeaker and was given a pacifier to suck on. Once a baseline reading of high amplitude sucking responses was measured, one stimulus from the pair to be tested was presented to the infant, contingent upon sucking, -whenever the enfant sucked, the speech sound was presented. Typically, infants responded to the contingency by increasing their rate of sucking. After some period of time, however, the sound lost its reinforcing properties and the rate of sucking declined. When a predetermined criterion of decline had been reached, the second stimulus of the air was introduced, with no break in procedure. A renewed increase in sucking behaviourto the new sound, compared to no increase in control infants who received the first sound throughout testing, was taken as evidence of discrimination .The experiment was tested on three groups of infants. One group was tested on the *different* pair, one was tested on the *same* pairs, and one was a control group. The results were very similar for both age groups. Therefore, the infants showed clear evidence of discriminating the '*different*' pair but not the '*same*' pairs. Thus, infants could distinguish stimuli that differed by as little as 20 milliseconds VOT, but only if the stimuli were drawn from two different adult categories.

These data provided strong evidence that the young infant was highly sensitive to acoustic properties of speech (such as VOT) and, furthermore, that very early on infants perceived speech in terms of categories (such as those corresponding to [b] and [p]). And this phenomenon was not limited to the voicing distinction. Subsequent research on numerous phonetic contrasts (for instance, contrasts in place of articulation) demonstrated that such categorization by infants was the rule, not the exception. Thus, the basic phonetic categorization abilities required for language were something the infant brought to the task of the language acquisition-not something that was learnt in the course of language acquisition.

The existence of such highly developed abilities in early infancy accorded with both the motor theory and the auditory theory of speech perception, but the two theories accounted for these abilities in very different ways. According to the motor theory, the infant's abilities were due to the operation of specialized, articulatory-based mechanism; while according to the auditory theory, they were due to the operation of the auditory system itself, which was developed at birth (Miller 1990 p. 88.89).

1.1.1.7.6.2 Is Categorical Perception Human Species Specific?

Since the second major component of the motor theory involved the claim that the articulatory-based mechanism for speech perception was species-specific, the argument was simple: if speech perception was accomplished by a specialized mechanism based on articulatory principles, then nonhuman animals, which would not have such a mechanism; should perceive speech quite differently from humans. On the other hand, if speech perception was accomplished by the auditory system itself, as the auditory theory of speech perception claimed, then nonhuman animals should show very similar processing to that of humans-especially if their basic auditory systems were comparable to those of human listeners.

Many such studies were conducted, most notably by Kuhl and her colleagues. We will explore the basic procedures and findings by looking at a study that examined how Chinchillas perceive speech (Kulhl and J. D. Miller 1978). Chinchillas were chosen because, as far as we know, their basic auditory sensitivity is very similar to that of humans.

The study focused on the perception of two consonants, [b] and [p]. Both are stop consonants and both are produced with occlusion at the lips; that is, they have a labial place of articulation. They differ from each other in that [b] is a voiced consonant whereas[p]is a voiceless consonant; the vocal folds vibrate during the production of the former but not the latter. For the voiced consonant [b], just as the consonant is released following the period of occlusion at the lips, the vocal folds start vibrating and they remain vibrating as the [b] is fully articulated. For the voiceless consonant [p], on the other hand, there is a delay between the initial release of the consonant and the beginning of the vocal fold vibration. The time interval between the release of the consonant and the onset of the vocal fold vibration is called *Voice Onset Time*, or VOT. The VOT turns out to be a very important property used by listeners in identifying whether a stop consonant is voiced or voiceless. If the VOT value is relatively short, listeners hear the consonant as voiced, and if it is relatively long, they hear it as voiceless.

The phenomenon was tested on chinchillas concerned the way in which listeners divide or categorise the stimuli in the [ba]- [pa] series. Many studies on human listeners have demonstrated that when stimuli from such a series are randomized and presented for identification, human perceived stimuli with low VOT values as[ba] and those with high VOT values as [pa], with a rather sharp break, or phonetic boundary, between categories at approximately 25milliseconds VOT. According to the motor theory, the stimuli were categorized in that way, because the location of the phonetic category boundary was a consequence of the operation of a species-specific perceptual mechanism that accounted that in production,[b]was typically produced with VOT values less than 25 milliseconds, whereas [b] was typically produced with VOT values greater than 25 milliseconds. Since chinchillas do not have such a species- specific mechanism, they should not categorize the stimuli in that way. According to the auditory theory, the break at the 25 milliseconds for human listeners was a consequence of the way in which the auditory system processes complex acoustic patterns; it had nothing to do with the fact that the stimuli are speech. It just happened that the perceptual break along the series was 25 milliseconds. Chinchillas, should categorize the stimuli as we do, since they have an auditory sensitivity very much like our own.

To test these alternatives, Kuhl and Miller (1978) conducted identification tests on humans and chinchillas. The humans were tested with a typical identification procedure, in which the stimuli from the nine member [ba]-[pa] series were randomly presented for identification as [ba] or [pa].The investigators, on the basis of previous research, expected the human listeners to divide the stimuli into two categories, with the boundary between categories at about 25 milliseconds VOT. It was not possible to test chinchillas in the same way. Since we cannot ask chinchillas to tell us what they hear, whether they hear [ba] or [pa] on each trial. The followings are to sum up the experiment and the result drawn from it:

- Chinchillas exhibit categorical perception as well
- Train on end-point /ba/ (good), /pa/ (bad)
- Test on intermediate stimuli

The result was that Chinchillas switched over from staying to running at about the same location as the English b/p phoneme boundary.

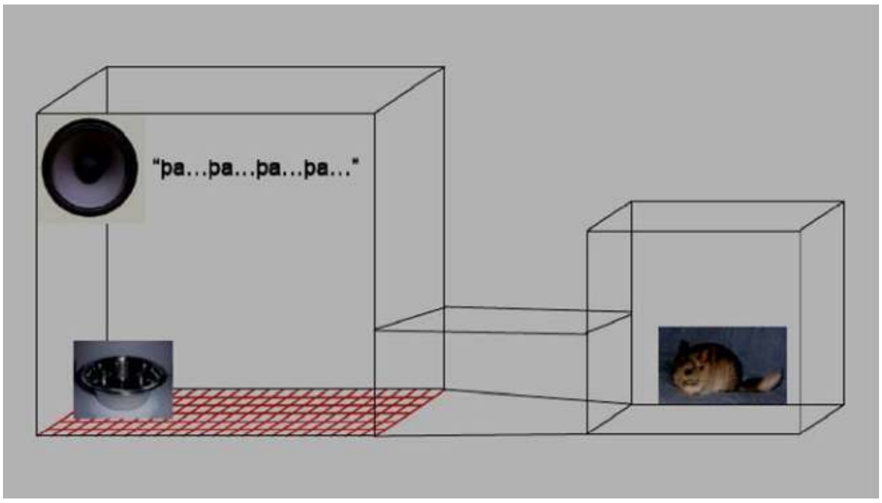
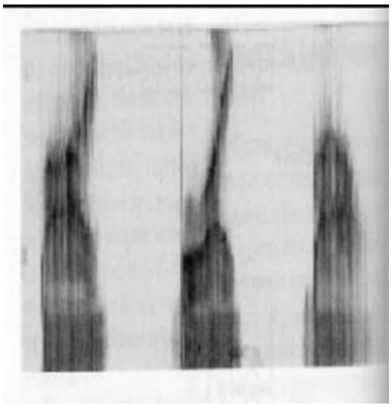


Figure 1.9.: Chinchilla Experiment (Kuhl& Miller,1978)

1.1.1.7.6.3. McGurk and MacDonald Effect (1976) (Vision and Speech Perception)

It has been found that speech perception can be affected by input from other senses such as touch or vision. We will examine this use of visual information by considering a study conducted by Massaro and Cohen (1983).The McGurk effect demonstrated how *visual information* can affect *speech perception*. When you watch the video while listening to what is being said. Most people will hear /da/ when asked what the speaker is saying. You will hear the speaker is saying /ba/.The speaker's lips are pronouncing /ga/ but the sound coming from the speakers is /ba/.If you turn down the volume and watch the speaker's mouth closely, you'll be able to see this. When your eyes are open and the sound clearly audible, the auditory and visual information were integrated to perceive /da/ and is therefore called audiovisual speech perception. It showed that speech perception is susceptible to influences from other senses.



Lips say /ba/ Sound signal /ga/

/ba/Bilabial, /ga/ Velar, /da/Dental

Subjects hear /da/

Figure 1.10.: The McGurkEffect (McGurk and MacDonald, 1976)

1.1.1.8. Acoustic Phonetics

The branch of phonetics that studies the physical parameters or properties of speech sounds is called acoustic phonetics. We have seen earlier that articulatory phonetics deals with speech production, auditory phonetics deals with the speech perception, acoustic phonetics, therefore, deals with the speech transmission. In other words, how speech is transmitted in the air from the speaker’s mouth to the listener’s ear, taking into account their physical properties. It is the most technical of all disciplines that are concerned with the study of verbal communication. The data it handles are the most concrete, palpable, easily measurable ones that can be encountered in the domain of phonetics in general[11].The most important principle of physics on which verbal communication is based is that vibrating bodies send waves that are propagated in the environment.

1.1.1.8.1. Why Should Linguistic and Phonetic Teachers Study Acoustic Phonetics?

- Physical acoustic signal is the primary means through which messages are being transmitted from the speaker to the listener. We communicate using sounds – not movements of the tongue or other articulators.
- Analysis of acoustic signal can provide insight into both the physiological phenomena underlying the production of speech and the perceptual mechanisms by which speech is perceived by listeners.
- Meaningful differences between words are encoded as sound differences. Some natural classes (e.g. labials and velars; retroflex) make sense only in acoustic terms. Some aspects of speech (e.g. friction) can only be properly defined in acoustic terms.
- Acoustic analyses have a wide range of applications, e.g. speech synthesis, automatic speech recognition, speaker identification, communication aids, clinical speech pathology and rehabilitation programs, etc...

1.1.1.8.2. What is a Sound?

Sound is variation in air pressure detectable by the human ear. It is a result of an action, and the sound is not produced unless there is a vibration. Many vibrating objects

[11]If it is difficult to analyze in detail, to carefully observe the speech organs in the process of producing various sounds; the acoustic features of sounds are more easily observable. The sounds we produce can be recorded, their features can be analyzed, and we can even provide graphics representing the sounds we articulate by means of special machines.

create sounds, for instance, a rubber band, a tuning folk, and for human beings, we have the vocal folds. A sound is a wave characterized by the transmission of energy in the form of compression (increased pressure) and rarefaction (decreased pressure) through the medium, which is the air.

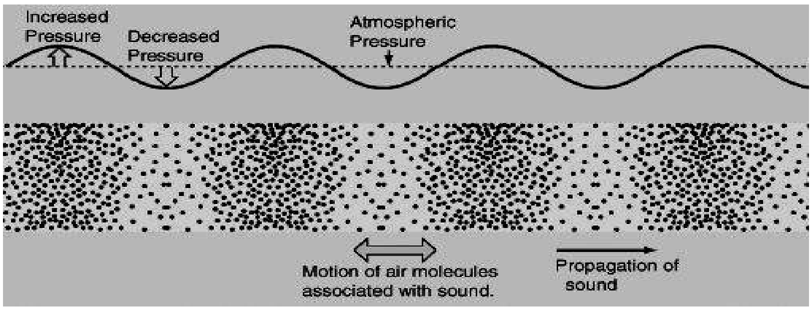


Figure 1.11: Sound Propagation

1.1.1.8.2.1. What are the Basic Properties of Sound Waves?

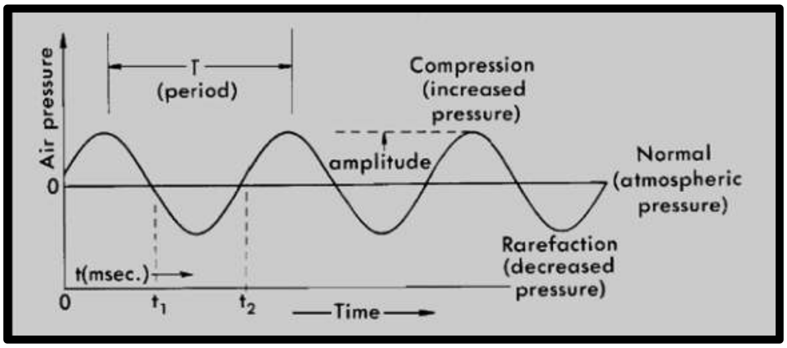


Figure 1.12: Period (T)

a. Period (T) is the amount of time needed to complete one cycle of movement. It is measured in seconds (s) and its fractions (millisecond, nanosecond, etc..).

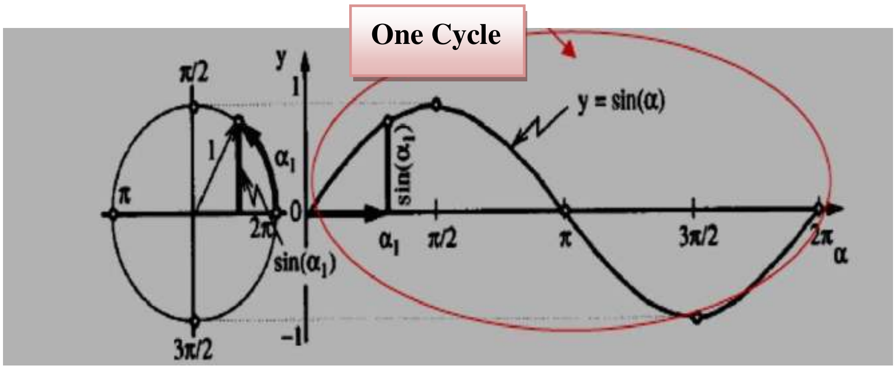


Figure 1.13.: Sinusoidal sound waveform and its properties (**One cycle)**

b. Frequency (f) is the number of cycles completed in one second. It is measured in cycles per second (cps), a unit of measurement known as *Hertz* (Hz) in honour to the German physicist *Hertz***.** Sounds having a constant period (in other words, sounds displaying a regular vibration) are called *Periodic sounds***.** Human ear can pick up frequencies in the range between 20 and 20,000 Hz. The perceptual correlate of frequency is pitch.

However, *Aperiodic* waves (random, non-repetitive) do not have a regular repeating pattern and are perceived as noises, like doors slamming, or wind blowing. They do not have a harmonic basis, i.e. the component frequencies, of which they are made up, are not related to each other (*component frequencies are not integer multiples of fundamental frequency*).It is thus important to mention that relation between frequency and period is:

F= *1/T* E.g. a wave with a period of **1/100th** of a second has a frequency of **1/0.01=100Hz.** Thus, frequency is the inverse of the period duration.

c. Amplitude (A) is the maximum displacement from the equilibrium position. In the case of sound waves, it is the extent of the maximum variation in air pressure from normal atmospheric pressure.

Amplitude is related to *loudness*, to the speech height, when amplitude decreases the sound becomes less loud.

If the distance from the point of rest is greater, we say the amplitude of the vibration is higher. This is related to the amount of energy that is transmitted through the air by means of the respective sound wave. The higher the amplitude is, the louder the sound. The conventional way in which we refer to the intensity or loudness or amplitude of sounds is that of using the *Decibel Scale*. Thus, if we want to compare the intensity of two sounds, we take the logarithm to the base 10 of their ratio and multiply it by 10.

For instance, if a sound is 1000 times more intense than another, it means that 10has to be raised to the power of 3 to get the ratio between them. If we multiply 3 by 10 we get 30, therefore the difference between the two sounds is of 30 decibels (dB). If a sound is a billion times more intense than another, this means that their ratio is 10 raised to the power of 9, so the difference between them is of 9x10 multiplied by 10, that is of 90 decibels (dB). The reference value for the decibel scale is the standard intensity of a sound which has a fixed value close to the audible limit of sound.

d. Intensity*:* represents the amount of energy carried by a sound wave as measured in decibels (dB); the acoustic correlate of perceptual auditory correlate is loudness too.

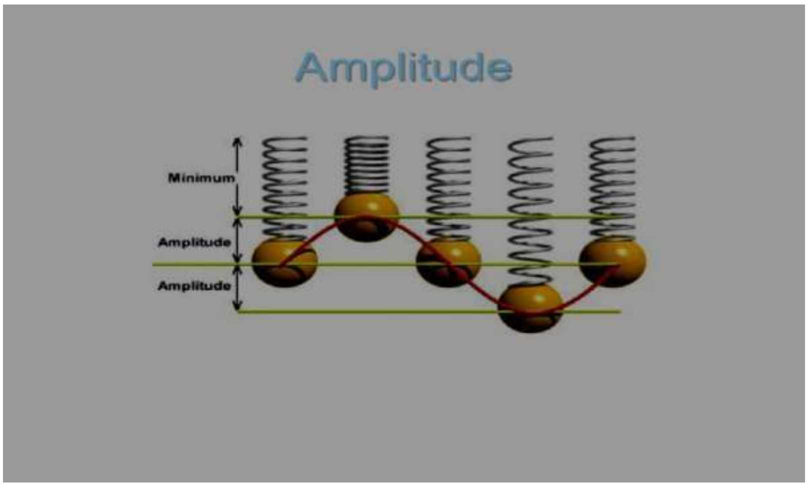


Figure 1.14: Amplitude

NB: Frequency and amplitude are not connected. Two waves can have the same frequency but different amplitude and vice versa (see Figures below).

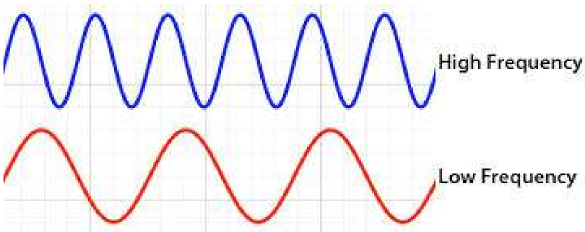


Figure1.15: Frequency (High and Low)

1.1.1.8.2.2. What Types of Waveforms can we encounter?

There are three basic properties or dimensions by which sound waveforms can be classified: complexity, periodicity and duration.

a. Simple (Sinusoidal) vs. Complex Waves

Simple waves are also called *sinusoidal/sine waves*; they result from *Simple Harmonic Motion* (SHM) and are made up of a single frequency component. SHM is a regular motion, in which each cycle of the movement takes exactly the same time, e.g. vibration of a tuning fork. SHM can be represented as a uniform circular motion projected on a plane (hence the terms cycle, sine wave).

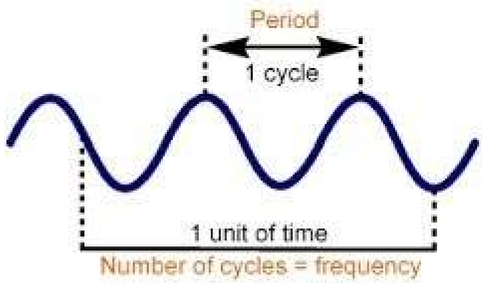


Figure 1.16: Sine wave. One cycle

Addition of sine waves of the same frequency produces another sine wave of greater amplitude (imagine listening to the sound produced by two tuning forks of the same frequency – it will sound louder than just one fork).

Addition of sine waves of different frequencies results in a *complex wave*, therefore any complex wave always has more than one frequency component. It is possible to construct an infinite variety of complex waves by combining sine waves of difference amplitudes and frequencies as the figure below shows.

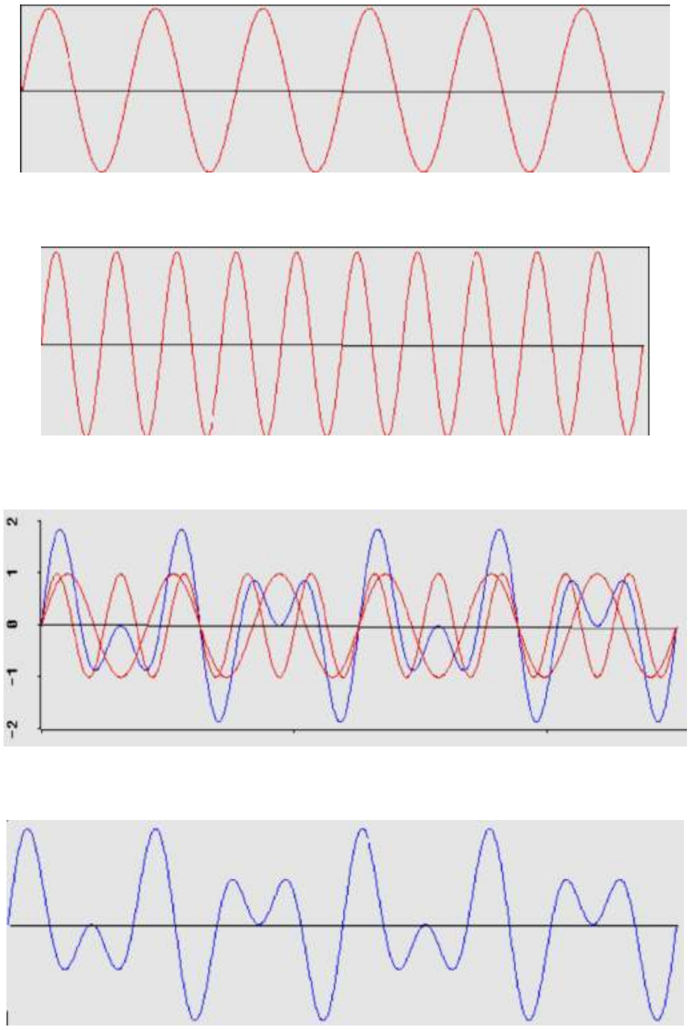


Figure1.17.: Complex Wave Resulting from adding two Sine Waves: 300 Hz and 500 Hz

All complex waves can be regarded as a sum of a (possibly infinite) number of sine waves **FFT** (Fast Fourier Transformation). *Fourier Analysis* is a mathematical technique for decomposing a complex wave into its component sine waves. Result is the *spectrum* of sound. The horizontal axis of a spectrum corresponds to frequency and the vertical axis corresponds to amplitude of the individual components in the complex wave. Figures show spectra of a sine wave and a complex wave respectively.

Fourier's technique is very powerful. Sine waves are easy to deal with mathematically and the representation of complex functions as sine waves often makes analysis much less difficult.

As we have said, acoustic phonetics is the branch of phonetics where data are most liable to measurements, quantification, etc.. If we can hardly think of apparatuses being used in other linguistic fields like syntax or semantics, for instance, the situation is different in the case of phonetics, as scientists have devised various instruments that are used to provide *an image* of the way in which people speak and graphics representing the sounds we produce. Such an instrument is the acoustic spectrograph, an appliance similar in many ways to a seismograph, or to an electrocardiograph (devices that record seismic and heart activity respectively). It marks on paper the vibrations caused by speech sound production. The graphs they produce are called *spectrograms.*

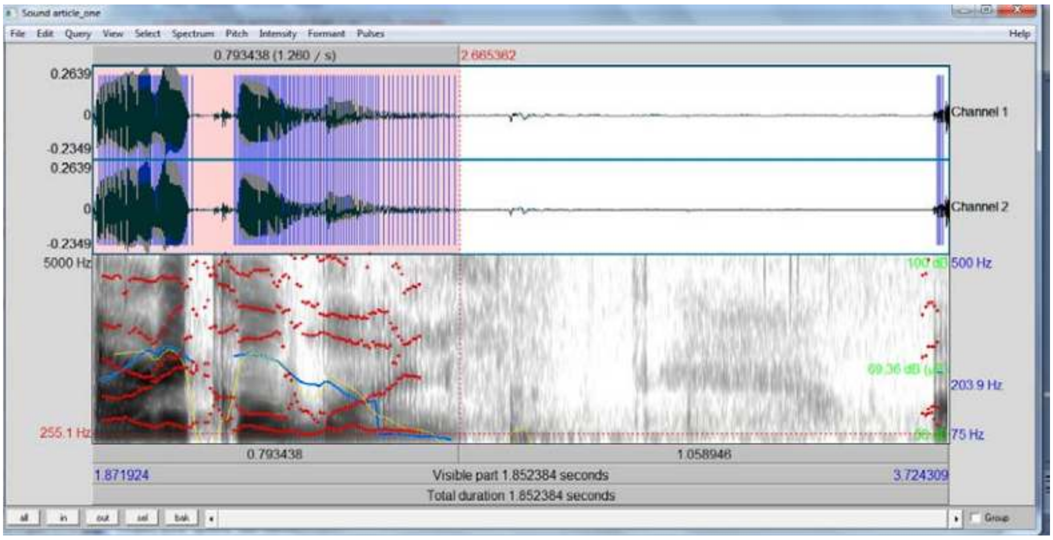


Figure1.18.: Spectrogram with Praat Software

1.1.1.8.3. What is a Spectrogram, Anyway?

A sound spectrogram (or sonogram) is an immediate visual representation of speech in terms of an acoustic signal. To oversimplify things a fair amount, a *Fast Fourier transform* is applied to an electronically recorded sound. This analysis essentially separates the frequencies and amplitudes of its component simple waves. The result can then be displayed visually, with degrees of amplitude (represented *light*-to-*dark*, as in *white*=*no energy, black*=*lots of energy*), at various *frequencies* (usually on the vertical axis) by *time* (horizontal). The vertical axis of the spectrogram, which as mentioned, represented the frequency measured in Hertz, is sub-divided into three frequencies, which are the higher frequency, the medial frequency and the lower frequency.

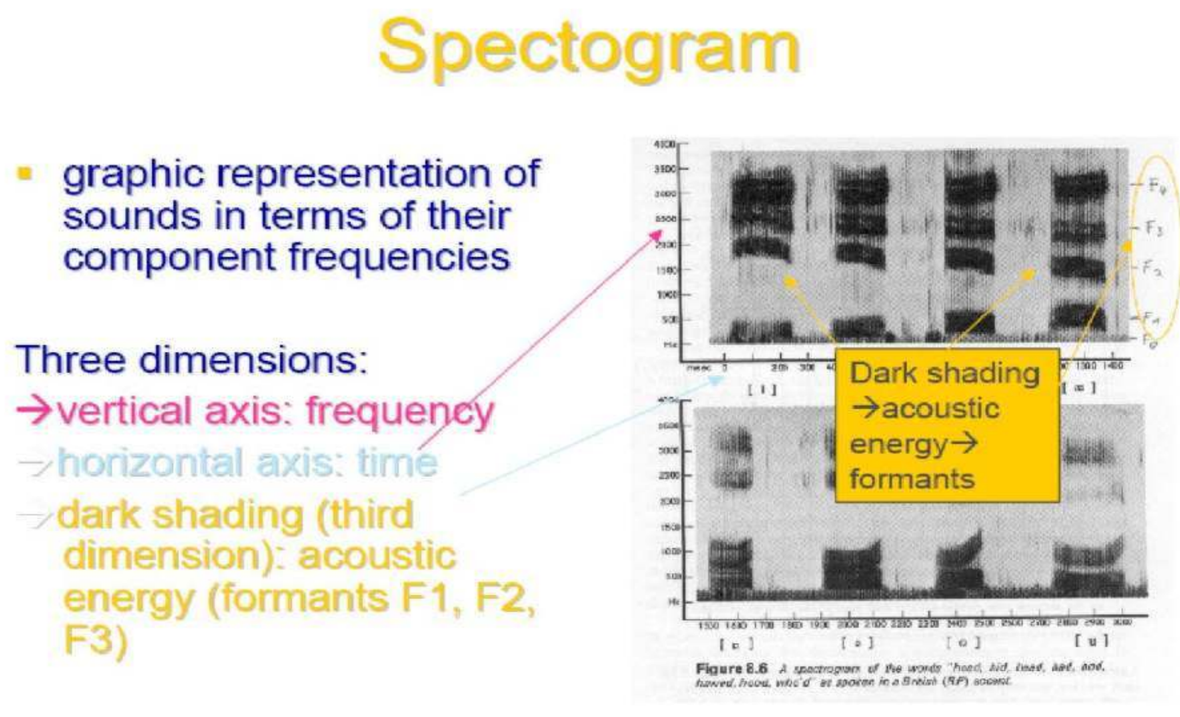


Figure 1.19.: Spectrogram Axes

1.1.1.8.3.1.1. Types of Spectrogram

There are two types of spectrograms, which are used according to what we seek to analyze, whether formants or harmonics.

1.1.1.8.3.1.1.1. Wide Band Spectrogram

A *Wide Band Spectrogram* shows the *formants* (darker bands running horizontally across the spectrogram). Generally, wide band spectrograms are used in spectrogram reading because they give us more information about what is going on in the vocal tract.

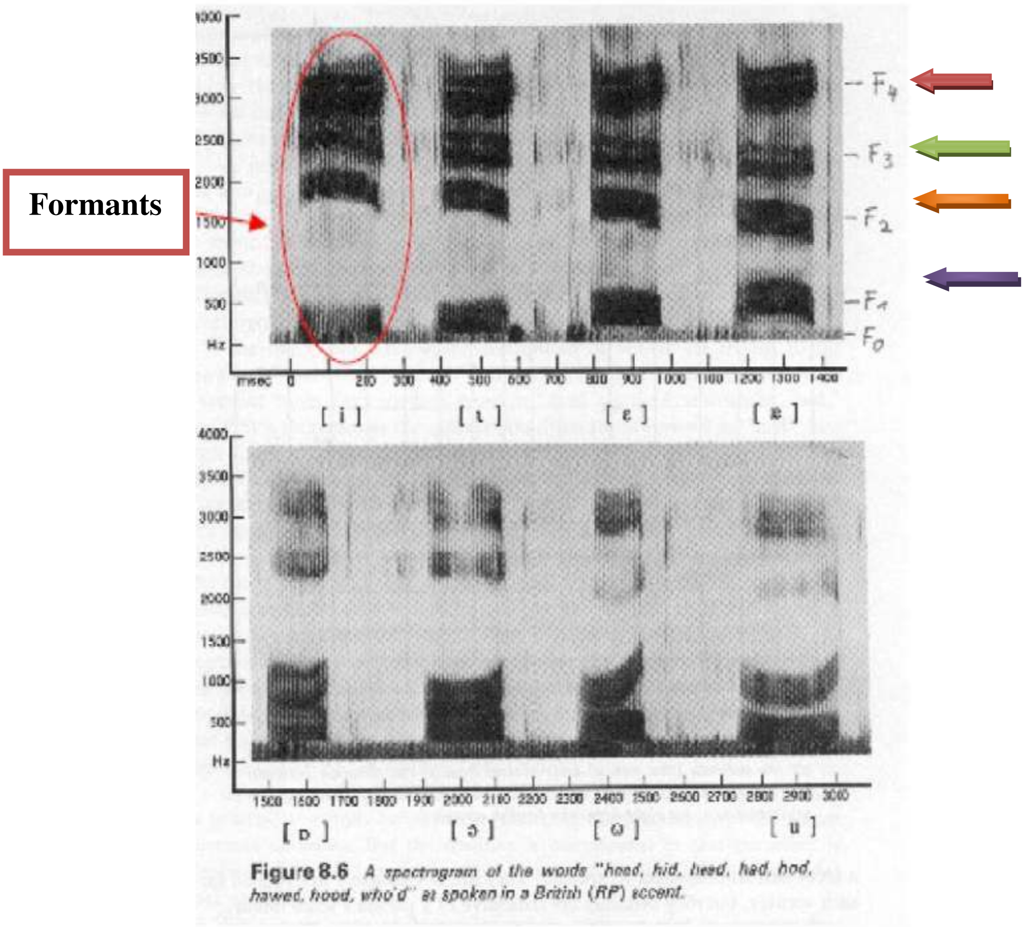


Figure1.20: Formants (Ladefoged, 2006)

Formants: when we speak about the formants, we refer to the vowels, which have a set of dark bars at various frequency bands, different bands for each vowel, amplified by the vocal tract. Since different vowels are produced with the vocal tract in different positions, they will produce different kinds of amplifications or what we refer to as resonances. Technically, formants represent a set of adjacent harmonics which are boosted by a resonance in some part of the vocal tract. Thus, different vocal tract shapes will produce different formant patterns, regardless of what the source is doing. In other terms, the darker bands in the spectrogram are called the *formants* of the respective sounds and they represent the frequencies at which a greater amount of energy is spent. Normally, two or three formants at the most are used to describe a certain sound. Formants are essential for the acoustic representation of sounds and all voiced sounds have a formant structure.

Formants are created by the different tongue postures that are used to create different vowels, that is why, we want to precise that only vowels have formants but not consonants. Let us see what the three formants of vowels represent.

The *first formant* (**F1**) tells us how high or low the tongue posture is, as it extends or shortens the longest part of the vocal tract, which is your throat from the larynx to the roof of your mouth.

The *second formant* (**F2**), and its relationship to F1, tell us how front the articulation is. A fronter articulation will have greater difference between F1 and F2. Essentially, a higher F2 means a fronter vowel, as the front cavity is shortened.

The third formant **(F3)** tells us about rounding, that is whether we extend the front-most part of the cavity by protruding our lips.

1.1.1.8.3.1.1.2. Narrow Band Spectrogram

Depending on the size of the Fourier analysis window, different levels of frequency/time resolution are achieved. A long window resolves frequency at the expense of time—the result is a *narrow band spectrogram***,** which reveals individual *harmonics* (*component frequencies*), but smears together adjacent 'moments'. If a short analysis window is used, adjacent harmonics are smeared together, but with *better time resolution***.** The frequency components of sine waves that make up a complex periodic wave are called *harmonics*. The lowest harmonic is *fundamental frequency*F0, which represents the vibration of the vocal folds. The Fundamental frequency of males is about **120 Hz**, females are around **255 Hz** and the children are estimated to **265 Hz**. All higher harmonics – also called *overtones or partials*– are integer multiples of the fundamental frequency.

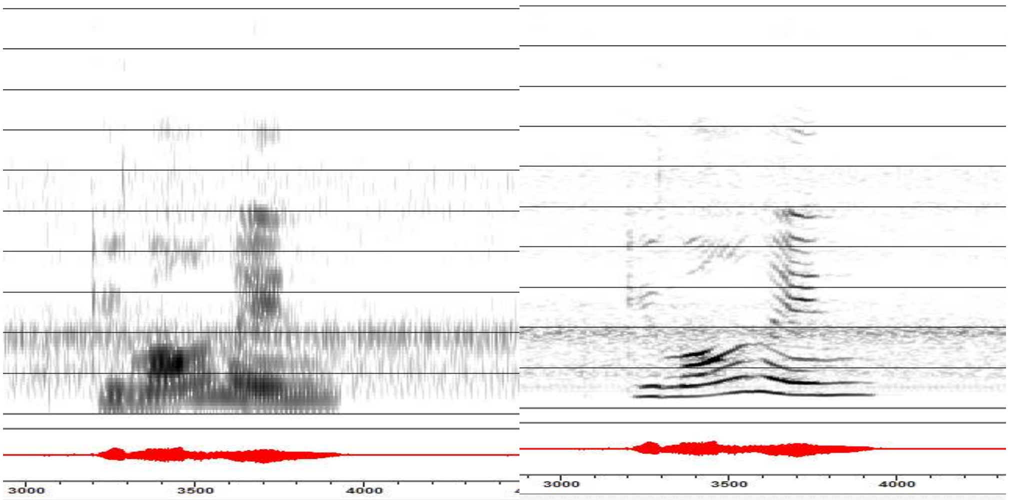


Figure 1.21: Wide Band Spectrogram (Left) and Narrow Band Spectrogram (Right) Winsnoori Software.

1.1.1.8.3.1.2. Cues for Reading a Spectrogram

Sounds have different classes and different acoustic parameters. The two major classes of sounds are vowels and consonants. Vowels in their turn have distinct acoustic features. Front vowels, for instance, are acute sounds, displaying higher frequencies in their second formant (between 1800 and 2300 cps), while back vowels are, comparatively, graver sounds, their second formant ranging between 800 and 1000 cps. We can also distinguish between compact and diffuse vowels, depending on the way in which the main formants are close to each other or are wider apart in the spectrum of the sound. Thus, low or open vowels have their formants grouped towards in the middle of the spectrum and are consequently compact, while high or close vowels are diffuse, the distance between their formants being greater. Consonants, on the other hand, can be clearly distinguished on the basis of their acoustic features. Non-peripheral (dental, alveolar, palate-alveolar, palatal) sounds are acute, as their formants are situated among the upper frequencies of the spectrum, while peripheral consonants are grave, as their formants are situated among the lower frequencies of the spectrum.[1] . We will try to explain some cues, which help us, read a spectrogram and identify the sounds; of course, it would seem as a difficult task at the beginning, but everything needs practice and training.

1.1.1.8.4. Sources and Filters

We often talk about speech in terms of *source-filter theory*. Put simply, we can view the vocal tract like a musical instrument. There is a part that actually makes sound (e.g. the string, the reed, or the vocal folds), and the part that 'shapes' the sound (e.g. the body of the violin, the horn of the clarinet, or the supra-laryngeal articulators). In speech, this *source* of sound is provided primarily by the vibration of *the vocal folds*. From a mathematical standpoint, vocal fold vibration is complex, consisting of both a *fundamental frequency* and *harmonics*. Because the harmonics always occur as integral multiples of the fundamental (*x1, x2, x3*, etc…—which phenomenon was mathematically proven by *Fourier*, hence *Fourier's Theorem* and *Fourier Transform*), it turns out that the sensation of *pitch* of voice is correlated to both the fundamental frequency, and the distance between harmonics.

Voicing is represented on a wide band spectrogram by vertical *striations*, especially in the lowest frequencies. Each vertical 'line' represents a single pulse of the vocal folds, a single puff of air moving through the glottis. We sometimes refer to a 'voicing bar', i.e. a row of striated energy in the very low frequencies, corresponding to the energy in the first and second harmonics (typically the strongest harmonics in speech). For men, this is about 100-150 Hz, for women it can be anywhere between 150-250 Hz, and of course there is lots of variation both within and between individuals. In a narrow band spectrogram, voicing results in harmonics, with again the lowest one or two being the strongest.

Non-voicing is basically silence, and does not show up as anything in a spectrogram. So while there is not a lot going on during *silence* that we can see in a spectrogram, we can still tell the difference between voiced sounds (with a striated voicing bar) and voiceless sounds (without). And usually there is still air moving through the vocal tract, which can provide an alternative source of acoustic energy, via turbulence or 'noise'.

On the other hand, it is worth distinguishing several glottal states that lead to non-voicing. Typically, *active devoicing*, results from vocal fold abduction. The vocal folds are held wide apart and thus movement of air through the glottis does not cause the folds to vibrate. If the vocal folds are tightly adducted (brought together in the midline) and stiffened, the result is no air movement through the glottis, due to *glottal closure*. Ideally, this is how a 'glottal stop' is produced. Finally, the vocal folds may be in 'voicing position', loosely adducted and relatively slack. But if there is insufficient pressure below the glottis (or too much above the glottis) the air movement through the glottis will not be enough to drive vibration, and *passive devoicing* occurs.

1.1.1.8.5. Place and Manner of Articulation

Plosives (oral stops) involve a total occlusion of the vocal tract, and thus a 'complete' filter, i.e. no resonances being contributed by the vocal tract. The result is a period of silence in the spectrogram, known as a 'gap'. A voiced plosive may have a low-frequency *voicing bar* of striations, usually thought of as the sound of voicing being transmitted through the flesh of the vocal tract. However, due to *passive devoicing*, it may not. And due to *perseverative voicing* even a 'voiceless' plosive may show some vibration as the pressures equalize and before the vocal folds fully separate.

1.1.1.8.5.1. Acoustic Properties of Fricatives

Fricative consonants are aperiodic sounds but we can still use source-filter theory to account for their production. Their turbulent noise source is generated as the result of the airflow travelling through a narrow constriction. Then this aperiodic continuous source is filtered by the vocal tract. Here as some examples of fricatives displayed on a spectrogram.

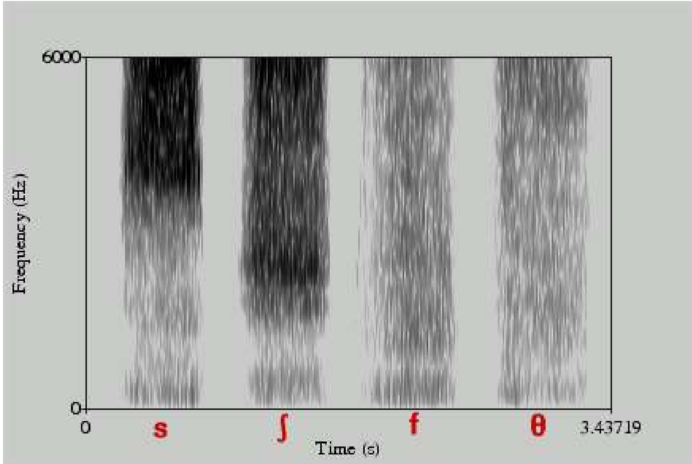


Figure 1.22: Sound Source in Fricatives

In the production of fricatives, the source is *turbulent noise* which occurs when air flows rapidly through a constriction. The turbulent noise is a result of irregular particle movement, which happens when particles leave a relatively narrow channel and hit inert outside air. The physical factors determining turbulence are the *size of the channel* and *volume velocity* of the airflow (volume of air going past a fixed point per unit of time).

A turbulent noise can be produced at any point along the vocal tract where a constriction can be formed, including the glottis. Noise produced at a glottal constriction is known as *aspiration noise*. [h]and whispered speech have aspiration noise source. *Frication noise* is produced through a supra-glottal constriction that can occur at many different points along the vocal tract.

Frication/aspiration noise source can be combined with glottal periodic source as in production of voiced fricatives. Voiced fricatives are not very common cross-linguistically because they are relatively harder to produce. Turbulent noise requires a high velocity of airflow but during voicing, airflow from the lungs is slowed by the constant closing (and opening) of the glottis. As a result voiced fricatives sometimes lose frication and become glides.

There are two major mechanisms for generating turbulence noise at a given constriction and these two mechanisms produce noise sources that are quite different in amplitude/intensity.

i. **Channel turbulence** - air stream is directed in such a way that it does not impinge on any surface; flow of air from the constriction forms a jet which is distributed over the area in front of the constriction. The resulting noise is of low intensity. This source is involved in production of [ɸ], [β] and [h].

ii. **Obstacle turbulence** - air stream at a constriction is directed against a surface or obstruction

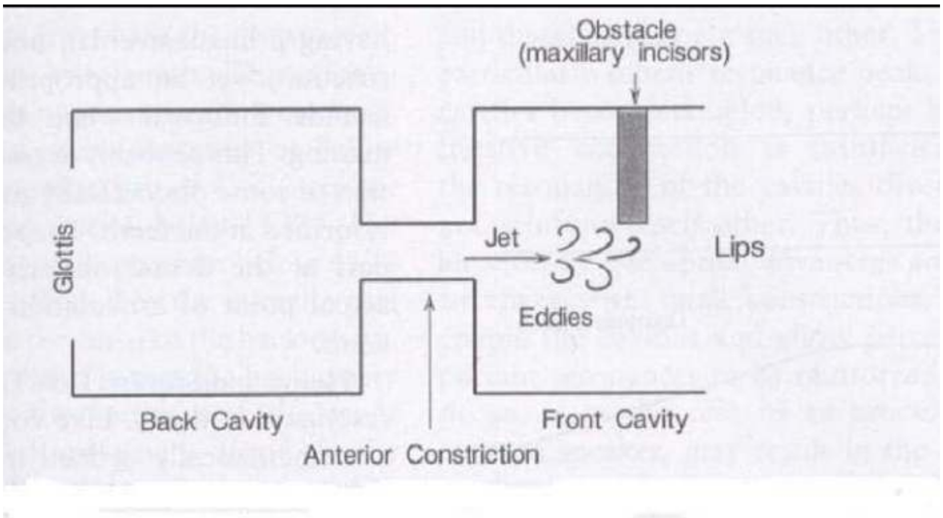


Figure 1.23: Tube Model of the Vocal tract for Obstacle Fricatives (Lass,1996)

Aperiodic turbulent noise is idealised as having a straight line spectrum, i.e. consisting of all frequencies at equal amplitude. However, in speech fricative source spectra vary in shape depending on how the fricative is produced. Shadle (1990) demonstrated that there are at least two types of fricative sources that have distinctive spectra.

a) **Wall** or *surface source* results when the airflow impinges on the surface that is nearly parallel to its direction, as in the production of the sounds [x] and [χ]. Spectrum of a wall source has a broad peak around 4-9 kHz and is relatively low in amplitude

b) **Obstacle source** occurs when the airflow is directed against an obstacle which is roughly at a right angle to its direction. For example, the tongue may be grooved to direct the air stream against upper or lower teeth. This source is involved in production of *strident* fricatives, e.g.[s], [ʃ], [z] and [ʒ].

Spectrum of an obstacle source is gradually falling in higher frequency region but overall source of amplitude is relatively high.

1.1.1.8.5.1.1. Filtering Effects in Fricatives

The main difference between filter function of vowels and fricatives is the presence of *antiformants*, or zeros. Zeros are frequencies that are not transmitted by the vocal tract but are instead 'trapped' in the back cavity. They arise because in production of fricatives the vocal tract is radically constricted; the front cavity behaves like a half-open tube and acts as a main filter. Thus filtering function and poles of high energy are determined by the size and length of this cavity. Longer front cavity gives rise to lower resonant frequencies conversely, constrictions occurring more forward in the mouth result in shorter front cavity that has high resonances.

Thus *alveolars* have an energy peak at around 7500 Hz although there may also be a minor peak at 4500 Hz. *Post-alveolars* have a peak at around 3500-4000 Hz. The further back in the vocal tract the constriction occurs, the lower the peak of spectral energy is. Therefore *pharyngeal* fricatives have peaks at a lower frequency then *velar* ones; velars have peaks at a lower frequency than *palatals* etc…The production of *glottal* fricative does not involve the constriction in the vocal tract; therefore the filtering function is very similar to vowel resonant structure.

Sounds which are produced with constriction in the very front of the mouth, such as dentals, labiodentals and bilabials have flatter spectra because there is almost no vocal tract in front of the constriction to filter the sound. Generally, *labiodentals and dentals* both exhibit a slight emphasis in high frequency energy while *bilabials* have relatively more energy in lower parts of spectrum.

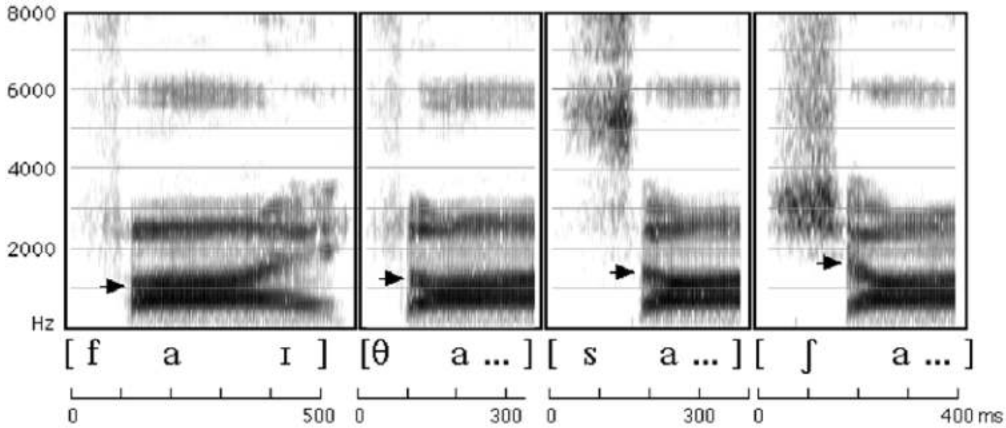


Figure 1.24: Spectrograms of Four Fricatives in Context (Ladefoged, 2001)

1.1.1.8.5.2. Acoustic Properties of Stops and Affricates

The production of stop consonants involves a sequence of articulatory and acoustic events; therefore they have more complicated acoustic properties. When talking about production of stops it is useful to distinguish several stages/phases. Note that these stages are not always discrete; they often overlap temporally and not every stage is necessary for production of a given stop in a given context.

i) if the stop is preceded by a vowel, a closing phase, or transition between the vowel and the stop;
ii) an interval of complete closure;
iii) release of the closure;
iv) in some stops, an interval of aspiration;
v) an interval of transitions from the stop into the following vowel.

As their production requires a complete closure, all stops have a period of silence; except for fully voiced (e.g. some intervocalic voiced stops in English) that have a low level of periodic energy during the closure.

1.1.1.8.5.2.1. Sound Sources in Stop Consonants

Because of the complex nature of stop production there are several sound sources involved.

i) during the closing phase the most common source is *voicing* but in pre-aspirated stops there may also be *aspiration noise*;
ii) usually no source during closure, except for voiced stops when *periodic glottal* source is possible; release produces aperiodic *transient noise* known as *release burst*; following

release *continuous aperiodic noise* is generated; it can be *aspiration* (aspirated stops) or *frication* (affricates); during transition *voicing* starts.

iii) The release burst is very brief transient noise which has a flat spectrum. Irrespective of other variations, bursts tend to show a strength hierarchy: voiceless aspirated > voiceless unaspirated**>** voiced, i.e. the burst spectrum of voiceless stops has more energy than that of voiced stops. The source spectra for aspiration and frication are nearly the same as for fricatives.

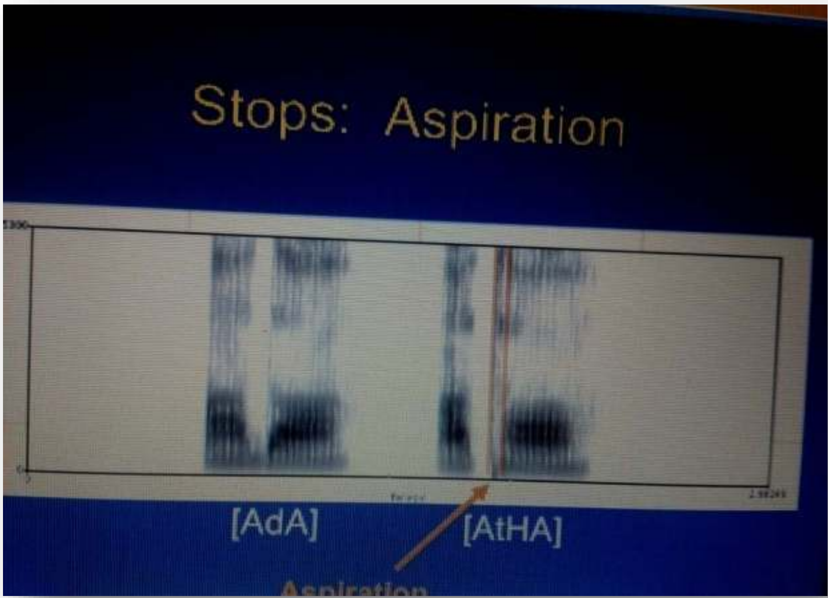


Figure 1.25: Spectrogram of Stops: Aspiration

1.1.1.8.5.2.2. Vocal Tract Filter Function in Stops

Again production of *stops* is determined by several different types of filtering effects.

i) Part of the vocal tract in front of the constriction influences the spectral properties of the release burst, re-shaping its flat spectrum;

ii) Complete closure of the upper vocal tract introduces antiformants into filtering function; it also decreases F1 to a value close to zero.

iii) Before and after release, front and back cavities are coupled; therefore, the filter can be presented as a two-tube model. As the articulators move from the position of the stop to the position of the vowel, the formants move rapidly. These are the *formant transitions*, which are seen in the adjacent vowel.

iv) Formant transitions are important cues for the place of articulation.

1.1.1.8.6. Voicing

i) Voiced Consonants: for voiced consonants the voicing itself starts less than about 30 ms after the release. It can also occur during the closure.

ii) Unvoiced Consonants: the voicing starts more than about 50 ms after the release and it cannot occur during the closure.

1.1.1.8.7. Acoustic Properties of Nasals, Liquids and Glides

Nasals and *liquids* share periodic glottal source with vowels. But their filtering effects are more complicated. Vocal tract filtering effects for vowels were modeled as resonances of coupled tubes, each of which was open at one end. Nasals and laterals are characterized by the presence of a side branch, i.e. a closed tube.

During the production of nasal consonants, the velum is lowered so that the pathway from the pharynx to the nasal passages is open and air flows from the lungs out through the nostrils. In nasal stops, the mouth cavity is closed off by a complete constriction in the vocal tract.

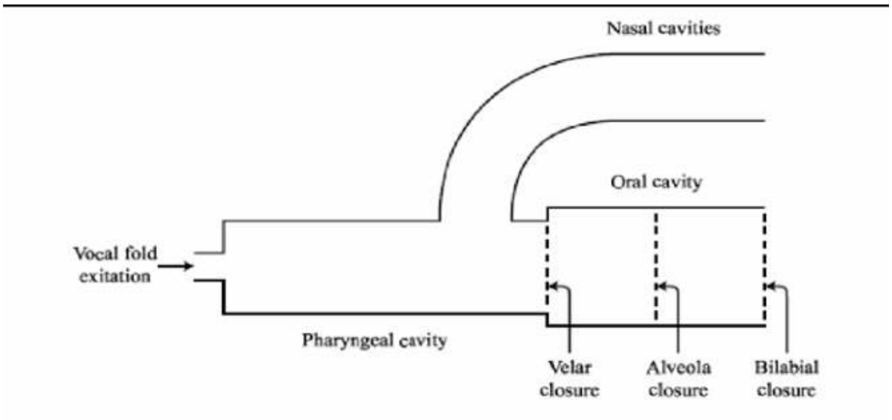


Figure 1.26: Tube Model for the Production of Nasal Consonants

(www.phon.ucl.ac.uk.PLIN2108/PLING216 Intermediate Phonetics)

This means that nasal consonants that are made with an oral constriction add a side cavity onto the long pharyngeal-nasal tube. The side cavity is a tube closed at one end; thus, resonating frequency components of the side branch are not transmitted out of the vocal tract but are absorbed by the side cavity. This gives rise to antiformants. Nasals have several antiformants which are determined by the size of the side cavity. As [m] has the longest side resonator its antiformants are low and narrowly spaced (first antiformant A1 around 800-1000 Hz). Antiformants of [n] are higher and wider spaced (A1 around 2000 Hz) and short oral branch produces the highest and widest antiformants for [ŋ] (A1 around 3000 Hz).

Other characteristic properties of *nasals* include:

(i) Low F1 (sometimes called N1 or *nasal murmur*); typical figures 250-300 Hz, as a result of increase in overall length of the filtering tract;

(ii) Formants are much weaker than in vowel sounds, i.e. the peaks are lower in amplitude due to the fact that soft mucous membranes of nasal cavity absorb sound.

Liquids demonstrate similar features to those of nasals. They typically have a clear formant structure, though with less energy than vowels. F1 is much lower than in vowels and formant transitions are usually very sharp, although less so in trills.

Laterals also have antiformants, due to the presence of a side cavity – a pocket of air on top of the tongue. However their antiformants are usually less strong than those in nasals.

Filter characteristics of the *glides*[j] and [w] are very similar to those of the vowels [i] and [u].The main difference between glides and vowels is that the former do not have a steady-state positions, i.e. they cannot be characterized in terms of stable formant values as they have lower F1 than for the vowels; ultimately they are formant transitions, that is to say they tend to have more formant movement than vowels.

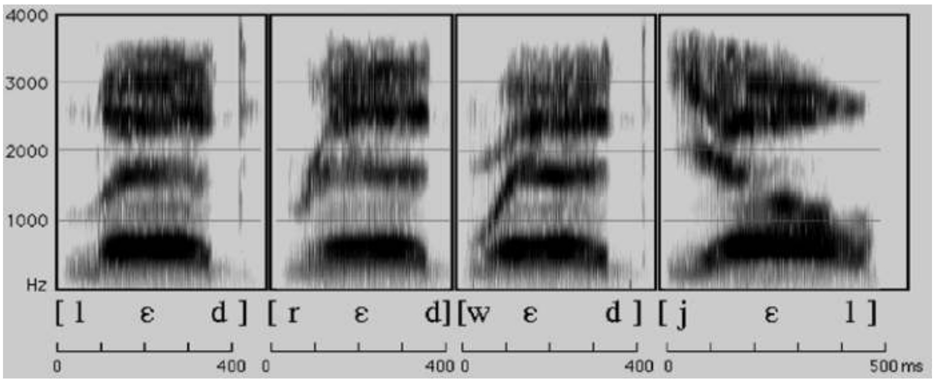


Figure1.27: Spectrograms Illustrating Distinctive Properties of Liquids and Glides in English (Ladefoged,2001)

1.1.1.8.8. Vowels

Since vowels are segments which are produced without any obstruction to the air flow as it passes from the larynx to the lips and no significant constriction of the vocal tract, articulations change resonances of the vocal tract; for instance: moving the tongue, lips and jaw alter the shape of the vocal tract. The resonances of the vocal tract are called *Formants.* On a *spectrogram,* vowels usually have very clearly defined *formant bars*, as in the following:

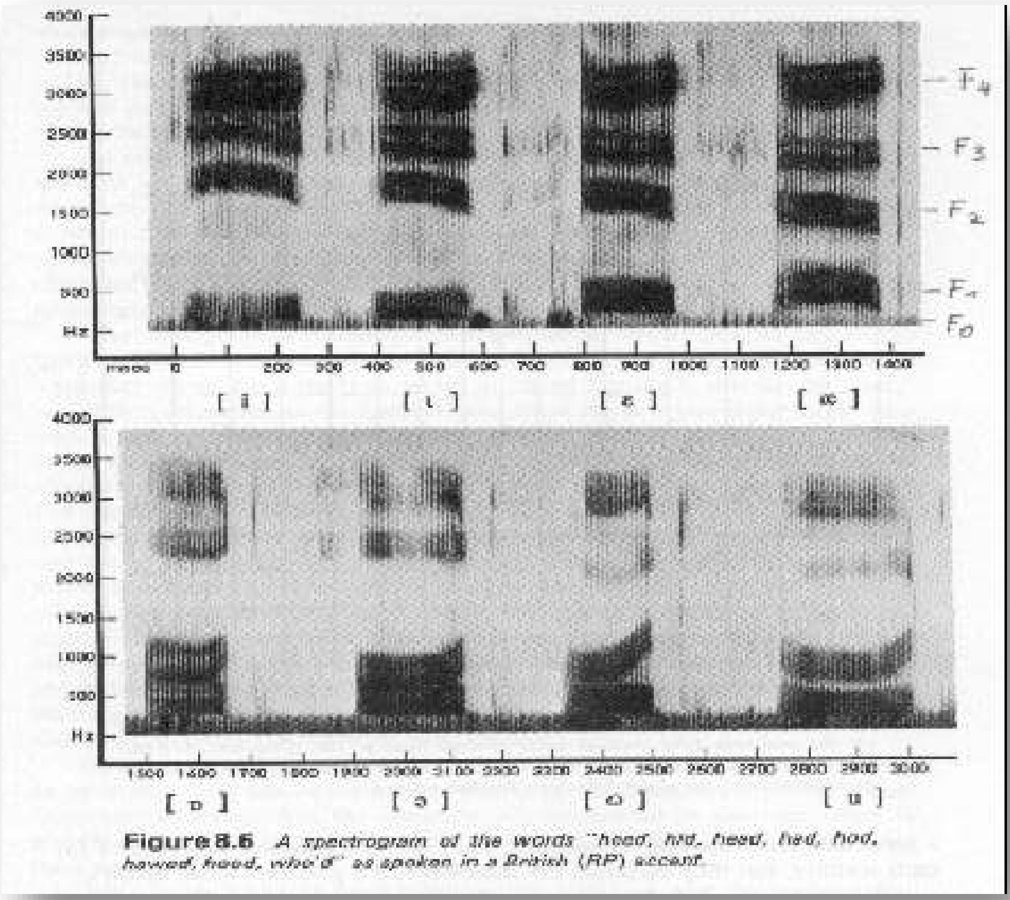


Figure1.28: A spectrogram of the words: heed, hid, head, had, hod, hawed, hood, who'd, (Ladefoged, 2006)

In diphthongs, you can see the formants change frequency as the tongue body moves through the mouth.

You cannot always tell reliably *which* formant you are looking at -- F1, F2, F3, etc…-- unless you already have a good idea of where to expect them. But the existence of formants is usually obvious enough that you can at least be sure you are looking at a vowel.(There are some especially common difficulties in identifying formants. In [ɑ], and sometimes other back vowels, F1 and F2 are often so close together that they appear as a single wide formant band. In [i], F2 and F3 also often appear merged together in a single wide band.)

1.1.1.8.8.1. Monophthong Vowels: these are characterized by strong stable voicing, as represented in Figure 1 below. The formants in the vowel, visible as a grouping of three components: (1) a red band of increasing energy, (2) a maximum in green and yellow, and (3) decreasing energy in blue, are stable in time. Geometrically this means that they are horizontal, showing no motion in the y-axis of frequency. In Figures below, we see F1 and F3, since F2 has been absorbed into F1.

1.1.1.8.8.2. Diphthong Vowels: the diphthongs have strong moving voicing, as represented in the figures below. The formants are not horizontal throughout the life of the vowel as they were in the monophthong vowels, but move from a beginning configuration to a target configuration. The spectrograms below represent some of the diphthongs:

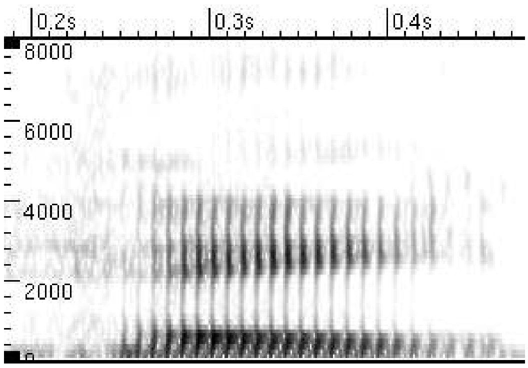
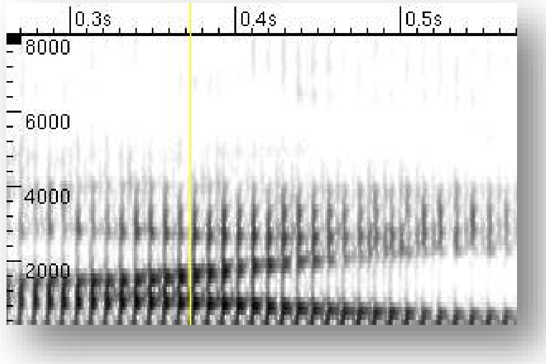


Figure 1.29: Spectrogram of the Diphthong /ei/ in Hate.



.

Figure1.30: Spectrogram of the Diphthong /ai/ in Like

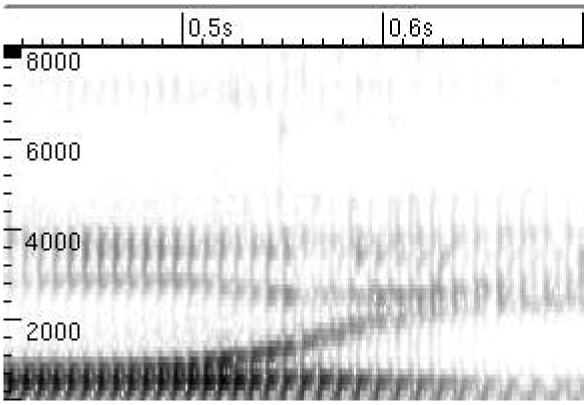


Figure1.31: Spectrogram of the Diphthong /ɔɪ/ in Hoy

1.1.1.8.8.3. Syllabics: When liquids or nasals occur in an unstressed syllable, the vowel is often merged into the liquid or nasal, which becomes syllabic in that it bears the weight of the syllable. The spectral appearance of the syllabics is midway between that of a vowel and that of a liquid or nasal.

Conclusion

To conclude what we have been through, teachers should be aware of the differences between the two tightly linked subfields of linguistics, which are phonetics and phonology, and to have a thorough mastery of both.

Phonetics and phonology are worth studying too for several reasons. One is that as any study of language, the study of phonology gives us insight into how the human mind works. Two more reasons are that the study of the phonetics of a foreign language gives us a much better ability both to hear and to correct mistakes that we make, and also to teach pronunciation of any foreign language. As both phonetics and phonology deal with sounds, and English spelling and English pronunciation are two different things, it is important to mention that we were much more interested in sounds and not in letters.

We have also explained what a spectrogram is and provided some cues to help identifying sounds according to their places and manner of articulation, which makes them different from each other. The spectrogram, which displays immediately our speech in a real time, is very helpful in depicting errors of pronunciation and thus correcting them.

Chapter Two

Chapter Two: Speech Production

Introduction

Knowing the minute details of speech production is of paramount importance for phoneticians and also for speech therapists and teachers of oral language and pronunciation. It is the purpose of this chapter to provide an overall account of the mechanism of this phenomenon which is exclusive to human beings.

2.1. Speech Chain

According to Denes and Pinson (1993), we usually take for granted our ability to produce and understand speech and give little thought to its nature and function, just as we are not particularly aware of the action of our hearts, brains, or other essential organs. It is not surprising, therefore, that many people overlook the great influence of speech on the development and normal functioning of human society.

Wherever human beings live together, they develop a system of talking to each other; even people in the most primitive societies use speech. Speech, in fact, is one of those few, basic abilities that set us apart from animals and are closely connected with our ability to think abstractly.

Why is speech so important? One reason is that the development of human civilization is made possible—to a great extent—by man's ability to share experiences, to exchange ideas and to transmit knowledge from one generation to another; in other words, his ability to communicate with other men. We can communicate with each other in many ways. The smoke signals of the Apache Indian, the starter's Pistol in a 100-yard dash, the finger signing language used by deaf people, the *Morse Code* and various systems of writing are just a few examples of the many different systems of communication developed by man. Undeniably, however, speech is the system that man found to be far more efficient and convenient than any other.

One may think that writing is a more important means of communication. After all, the development of civilization and the output of printing presses seem to parallel each other, and the written word appears to be a more efficient and more durable means of transmitting intelligence. It must be remembered, however, that no matter how many books and newspapers are printed, the amount of intelligence exchanged by speech is still greater. The use of books and printed matter has expanded greatly in societies, but so has the use of telephones, radio and television.

In short, human society relies heavily on the free and easy interchange of ideas among its members and, for many reasons; we have found speech to be our most convenient form of communication.

Through its constant use as a tool essential to daily living, speech has developed into a highly efficient system for the exchange of even our most complex ideas. It is a system particularly suitable for widespread use under the ever changing and varied conditions of life. It is suitable because it remains functionally unaffected by the many different voices, speaking habits, dialects and accents of the millions who use a common language. And it is suitable for widespread use because speech-to a surprising extent-is invulnerable to severe noise, distortion and interference.

Speech is well worth careful study. It is worthwhile because the study of speech provides useful insights into the nature and history of human civilization. It is worthwhile for the communications engineer because a better understanding of the speech mechanism helps in developing better and more efficient communication systems. It is worthwhile for all of us because we depend on speech so heavily for communicating with others.

The study of speech is also important for the development of human communication with *machines***.** We all use automatons, like push-button telephone-answering machines and automatic elevators, which either receive instructions from us or report back to us on their operations. Frequently, they do both, like the *computers* used so extensively in our society; their operation increasingly relies on frequent, fast, and convenient exchanges of information with users. In designing communication systems or *languages* to link users and machine, it should prove worthwhile to have a firm understanding of speech, that system of person-to-person communication whose development is based on the experience of many generations.

When most people consider speech, they think only in terms of moving lips and tongue. A few others, who have found out about sound waves, perhaps in the course of building or using stereo systems, will also associate certain kinds of sound waves with speech. In reality, speech is a far more complex process, involving many more levels of human activity, than such a simple approach would suggest.

A convenient way of examining what happens during speech is to take the simple situation of two people talking to each other. For example, you as the speaker want to transmit information to another person, the listener. The first thing you have to do is arrange your thoughts, decide what you want to say and then put what you want to say into a *linguistic form*. The message is put into a linguistic form by selecting the right words and phrases to express its meaning, and by placing these words in the order required by the grammatical rules of the language. This process is associated with activity in the speaker's brain, and it is from the brain that appropriate instructions, in the form of impulses along the motor nerves, are sent to the muscles that activate the *vocal organs*- the lungs, the vocal cords, the tongue, and the lips. The nerve impulses set the vocal muscles into movement which, in turn, produce minute pressure changes in the surrounding air. We call these pressure changes a *sound wave*. Sound waves are often called *acoustic waves*, because acoustics is the branch of *physics*, which is concerned with sounds.

The movements of the vocal organs generate a speech sound wave that propagates through the air between both the *speaker* and the *listener*. Pressure changes as the ear activate the listener's hearing mechanism and produce nerve impulses that travel along the acoustic nerve to the listener's brain. In the listener's brain, a considerable amount of nerve activity is already taking place, and this activity is modified by the nerve impulses arriving from the ear. This modification of the brain activity, in ways that are not yet fully understood, brings about recognition of the speaker's message. We see, therefore, that speech communication consists of a chain of events linking the *speaker's brain* with the *listener's brain*. We shall call this chain of events the *Speech Chain* (see the forthcoming figure)

It might be worthwhile to mention at this point that the speech chain has an important side link. In the simple speaker-listener situation just described, there are really two listeners, not one, because speakers not only speak, but also listen to their own voice. In listening, they continuously compare the quality of the sounds they produce with the sound qualities they intended to produce and make the adjustments necessary to match the results with their intentions. There are many ways to show that speakers are their own listeners. Perhaps the most amusing is to delay the sound *fed back* to the speaker. This can be done quite simply by recording the speaker's voice on a tape recorder and playing it back a second later. The speaker listens to the delayed version over ear-phones. Under such circumstances, the unexpected delay in the *fed-back* sound makes the speaker stammer and slur. This is the so-called delayed speech *feedback effect*. Another example of the importance of *feedback* is the general deterioration of the speech of people who have suffered prolonged deafness. Deafness, of course, deprives people of the speech chain's feedback link. To a limited extent, we can tell the kind of deafness from the type of speech deterioration it produces. Let us go back now to the main speech chain, the links that connect a speaker with a listener. We have seen that the transmission of a message begins with the selection and ordering of suitable words and sentences. This can be called *the linguistic level* of the speech chain (Denes & Pinson,1993).

We will consider how Stevens (1998) dealt with the speech chain, that is, the different forms in which a spoken message exists in its progress from the speaker's mind to the listener's mind with some illustrations.

2.1.1. Linguistic Level

According to Stevens (1998), words are the fundamental production of speech, which are organized into sentences. Words, in turn, are composed of syllables, which are constructed from phonemes or (what we can refer to as *segments*- which are vowels and consonants), so that a word like *fat* has three such segments. The segments, in turn, can be classified in terms of discrete, binary features. Thus /p, t, k/ for instance, have in common the feature *voiceless*, whereas, /b, d,g/ are *voiced*. Therefore, all of these six consonants have a common attribute which is stop *consonants*.

The speech event continues at *the physiological level*, with neural and muscular activity, and ends, on the speaker's side, with the generation and transmission of a sound wave, *the physical (acoustic) level* of the speech chain. At the listener's end of the chain, the process is reversed. Events start at the physical level, when the incoming sound wave activates the hearing mechanism. They continue at the physiological level with neural activity in the hearing and perceptual mechanisms. The speech chain is completed at the linguistic level when the listener recognizes the words and sentences transmitted by the speaker. The speech chain, therefore, involves activity atleast three levels-linguistic, physiological and physical -first on the speaker's side and then at the listener's end.

2.1.2. Physiological Level

The segments and features that make up words specify instructions indicating how the sounds in the words are produced by the articulators, for instance the tongue, lips, larynx, etc. These instructions state how the different articulators are to be moved and coordinated via the muscles, and, for a given feature, there may be some instructions that depend on the context in which it appears. So, a central problem is understanding the transformation from the discrete linguistic representation to the continuous movements of the articulators.

2.1.3. Acoustic Level

The movements of the articulators, together with the movement of air from the lungs, results in the creation of sound sources which are filtered by acoustic cavities that are shaped by the articulators. The physics of sound production draws on concepts of acoustics, source and filter theory, and signal processing. The sound radiated from a speaker's lips is the most accessible component of the speech chain. The sound has some quantal aspect and some continuous aspects. This sound impinges on the ear of a listener where it is converted to mechanical vibration of the eardrum (the tympanic membrane).

2.1.4. Auditory Level

In a series of stages, the sound is converted to mechanical motion, it undergoes some frequency analysis in the ear, and it is converted into electrical activity in the auditory nerve and in higher centers in the brain. The listener must decode this sound in order to extract the intended sequence of words at the linguistic level. The decoding presumably requires identification or estimation of the segments and features that underlie the sound pattern.

From another angle, Denes and Pinson (1993) went to say that we may also think of the speech chain as a communication system in which ideas to be transmitted are represented by a code that undergoes transformations as speech events proceed from one level to another. We can draw an analogy here between speech and Morse code. In Morse code, certain patterns of dots and dashes stand for different letters of the alphabet; the dots and dashes represent a code for the letters. This code can also be transformed from one form to another. For example, series of dots and dashes on a piece of paper can be converted into an acoustic sequence, like *beep-beep*. In the same way, the words of the English language area code for concepts and material objects. The word *dog* is the code for a four-legged animal that wags its tail, just as *dash-dash-dash* is Morse code for the letter “o”. We learn the code words of a language -and the rules for combining them into sentences- when we learn to speak.

During speech transmission, the speaker’s linguistic code of words and sentences is transformed into physiological and physical codes; in other words, it is transformed into corresponding sets of muscle movements and air vibrations- before being reconverted into a linguistic code at the listener’s end. This is analogous to translating the written *dash-dash-dash* of Morse code into the sounds *beep-beep-beep*.

Although we can regard speech transmission as a chain of events in which a code for certain ideas is transformed from one level or medium to another, it will be a great mistake to think that corresponding events at the different levels are the same. There is some relationship, to be sure, but the events are far from identical. For example, there is no guarantee that people will produce identical sound waves when they pronounce the same word. In fact, they are more likely to produce different sound waves when they pronounce the same word. By the same token, they may very well generate similar sound waves when pronouncing different words.

This state of affairs was demonstrated experimentally by Miller (1991), when a group of people listened to the same sound wave, representing a word, on three occasions when the word was embedded in three different sentences. The listeners agreed that the test word was heard either as *bit* or *bet* or *bat*, depending on which of the three sentences was used.

The experiment clearly showed that the general circumstances or the context under which we listen to speech profoundly affect the specific words we associate with particular sound waves. Put differently, the relationship between a word and a particular sound wave, or between a word and a particular muscle movement or pattern of nerve impulses, is not unique. There is no label on a speech sound wave that invariably associates it with a particular word. Depending on context, we recognize a particular sound wave as one word or another. A good example of this is reported by people who speak several languages fluently. They sometimes recognize indistinctly heard phrases as being spoken in one of their languages, but realize later that the conversation was in another of their languages.

Knowledge of the right context can even make the difference between understanding and not understanding a particular sound wave sequence. We may have listened to announcements made over a loudspeaker in an unfamiliar, noisy place like a bus or subway station. The chances are that many of the words were incomprehensible to some because of noise and distortion. Yet this same speech would be clearly intelligible to regular users of the station, simply because they have more knowledge of the context than others. In this case, the context is provided by their experience in listening under noisy conditions, and by their greater knowledge of the kind of messages to expect.

The strong influence of circumstance on what we recognize is not confined to speech. When we watch television or movies, we probably consider the scenes we see as quite life-like. But pictures on television are much smaller than life-size and those on a movie screen are much larger. Context will make the small television picture, the life-sized original, and the huge movie scene appear to be the same size. Black-and-white television and movies also appear quite life-like, despite their lack of true color. Once again, context makes the multicolored original and the black and white screen seem similar. In speech, as in these examples, we are usually quite unaware of our heavy reliance on context.

We can say, therefore, that speakers will not generally produce identical sound waves when they pronounce the same words on different occasions. Listeners, in recognizing speech, do not rely only on information derived from the speech wave they receive. They also rely on their knowledge of an intricate communication system, subject to the rules of language and speech, and on cues provided by the subject matter and the identity of the speaker.

In speech communication, then, we do not actually rely on precise knowledge of specific cues. Instead, we relate a great variety of ambiguous cues against the background of the complex system we call our common language. When we think about it, there is no other way speech could function efficiently. It does seem unlikely that millions of speakers, with all their different voice qualities, speaking habits and accents, would ever produce anything like identical sound waves when they say the same words. Even though our instruments for measuring the characteristics of sound waves are more accurate and flexible than the human ear, we are still unable to build a machine that can recognize speech nearly as effectively as a human being. We can measure characteristics of speech waves with great accuracy, but we do not know the nature and rules of the contextual system against which the results of our measurements must be related, as they are so successfully related in the listener's brain.

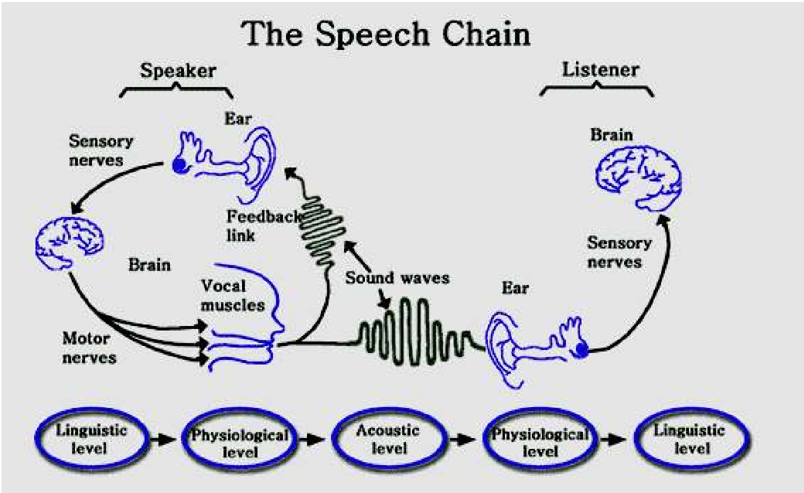


Figure2.1: Speech Chain (Adapted from Denes&Pinson,1993,p.5)

2.1.5. Physiological Level

Any manifestation of language by means of speech is the result of a highly complicated series of events. The communication in sound of such a simple concept as '*it's snowing*' involves a number of activities on the part of the speaker. In the first place, the formulation of the concept will take place at a linguistic level, i.e.: in the brain; the first stage may, therefore, be said to be *psychological*. The nervous system transmits this message to the *organs of speech* and these in turn behave in a conventional manner, which, as we have learned by experience, will have the effect of producing a particular pattern of sound; the second important stage for our purpose may thus be said to be *articulatory* or *physiological.* The movement of our organs of speech will create disturbances in the air, or whatever the medium may be through which we are talking, these varying air pressure may be investigated and they constitute the third stage in our chain, the physical or acoustic. Since communication generally requires a listener as a speaker, these stages will be reversed at the listening end: the reception of the sound waves by the hearing apparatus (*physiological*) and the transmission of the information along the nervous system to the brain, where the linguistic interpretation of the message takes place (*psychological*). Phonetic analysis has often ignored the role of the listener. But any investigation of speech as communication must ultimately be concerned with both the production and the reception ends. A number of phonetic features, e.g. stress, must be defined in different terms according to whether the emphasis is laid on the speaker's feeling for his own speech activity or on the listener's appreciation of significant features.

Our concern however is with the speaker's behaviour and more especially, on the concrete speech level, with the activity involved in the production of sounds. To discover how the various organs behave in order to produce the sounds of speech, we must examine the speaker's articulatory stage (his *speech mechanism*).

2.2. Speech Mechanism

Humans possess, in common with many other animals, the ability to produce sounds by using certain of their body's mechanisms .Humans differ from other animals in being able to organize the range of sounds they can emit into a highly efficient system of communication. Non-human animals rarely progress beyond the stage of using the sounds they produce as a reflex of certain basic stimuli to signal fear, hunger, sexual excitement, and the like. Nevertheless, like other animals, man when he speaks makes use of organs whose primary physiological function is unconnected with vocal communication, namely, those situated in the respiratory tract.

2.2.1. Sources of Energy: The lungs

The most usual source of energy for our vocal activity is provided by an airstream expelled from the lungs. There are languages which possess sounds not requiring lungs air for the articulation (*pulmonic sounds*);in English almost all the essential sounds use lung air for their production. The utterances we make, are, therefore, largely shaped by the physiological limitations imposed by the capacity of our lungs and by the muscles which control their action. Hence, we are obliged to pause in articulation in order to refill our lungs with air and this will to some extent condition the division of speech into intonational phrases. In those cases where the airstream is not available for the upper organs of speech (for instance, after the removal of the larynx, lung air does not reach the mouth but escapes from an artificial opening in the neck), a new source of energy, such as stomach air, has to be employed in this case.

A number of techniques are available for the investigation of the cavity during speech of the lungs and their controlling muscles. At one time air pressure within the lungs was observed by the reaction of an air-filled balloon in the stomach. Based on such evidence from a gastric balloon, it was at one time claimed that syllables were formed by chest pulses (Stetson, 1951). Such a primitive procedure was replaced by the technique of electromyography, which demonstrated the electrical activity of those respiratory muscles most concerned in speech, notably the internal intercostals; this technique disproved the relationship between chest pulses and syllables (Ladefoged, 1967).

Air coming out of the lungs is the source of power in nearly all speech sounds. When lung air is pushed out, we say that there is a *pulmonic airstream mechanism*. The lungs are sponge-like tissues within a cavity formed by the rib cage and the diaphragm (a dome-shaped muscle).When the diaphragm contracts, it enlarges the lung cavity so that air flows into the lungs. The lung cavity can also be enlarged by raising the rib cage which is the normal way of taking a deep breath in. Air can be pushed out of the lungs by pulling the rib cage down, or by pushing the diaphragm upward by contracting the abdominal muscles. X-ray photography and CT scans can reveal the big movements of the ribs and hence influence the surrounding muscles, although the technique of Magnetic Resonance Imaging (MRI) is not preferred on medical grounds.

In the description of most sounds, we take it for granted that the pulmonic airstream mechanism is the source of power. But in the case of obstruent consonants (stops and fricatives), other airstream mechanisms may be involved. Stops that use only an egressive or outward-moving, pulmonic airstream are called plosives. Obstruents made with other airstream mechanisms will be specified by other terms.

2.2.2. The Larynx and the Vocal Folds

The airstream provided by the lungs undergoes important modifications in the upper parts of the respiratory tract before it acquires the quality of a speech sound. First of all, at the top end of the **Trachea** or windpipe, it passes through the *larynx*, containing the *vocal folds*, less correctly called the *vocal cords* (see the figure below).

The larynx is a casing, formed of cartilage and muscles, situated in the upper part of the trachea. Its forward portion is prominent in the neck below the chin and is commonly called *Adam's apple***.** Housed within this structure from back to front are the vocal cords, two folds of ligament and elastic tissue which may be brought together or parted by the rotation of the arytenoids cartilages (attached at the posterior end of the fold) through muscular action. The inner edge of these folds is typically about 17-22 mm long in males and about 11-16 mm in females (Clark and Yallop, 1995 p. 181). The opening between the folds is known as the *glottis*. Biologically, the vocal folds act as a valve which is able to prevent the entry into the trachea and lungs of any foreign body or which may have the effect of enclosing the air within the lungs to assist in muscular effort on the part of the arms or the abdomen. In using the vocal folds for speech, the human being has adapted and elaborated upon this original open-or-shut function in the following ways.

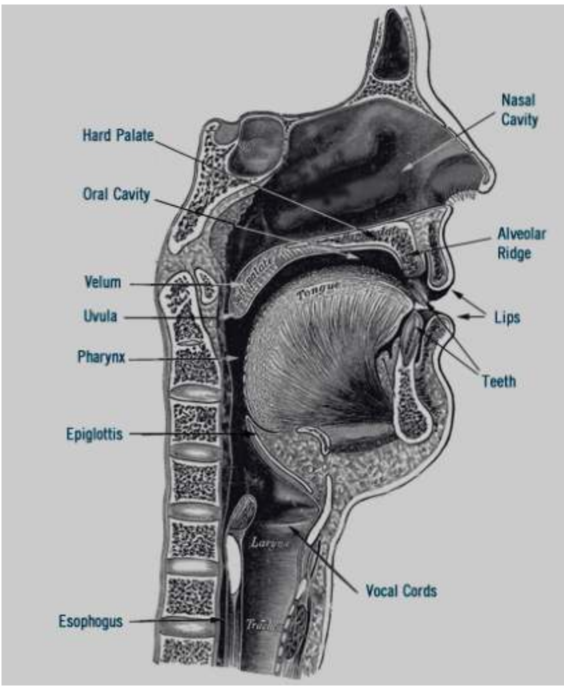


Figure2.2.: Organs of speech.[6]

a. The glottis may be held tightly closed, within the lung air pent up below it. This *glottal stop*/ʔ/ frequently occurs in spoken English, e.g. when it precedes the energetic articulation of a vowel as in *apple***[ʔǽpl]** or when it reinforces /p, t, k/ as in *clock* [klɒʔk] or even replaces them, as in *cotton* [kɒ**ʔn̩].**It may also be heard in defective speech, such as that arising from soft palate, when [**ʔ]**may be substituted for the stop consonants, which, because of the nasal air escape, cannot be articulated with proper compression in the oral cavity.

b. The glottis may be held open as for normal breathing and for voiceless sounds like [s] in *sip* and [p] in *peak*.

[6]From http://projects.iq.harvard.edu/cb45/pages/diagram-organs-speech-replacement

c. The most common action of the vocal folds in speech is as a vibrator set in motion by lung air, which produces *voice*, or *phonation*; this vocal fold vibration is a normal feature of all vowels or of a consonant such as [z] compared with voiceless [s]. In order to achieve the effect of voice, the vocal folds are brought sufficiently close together that they vibrate when subjected to air pressure from the lungs. This vibration, of a somewhat undulatory character, is caused by compressed air forcing the opening of the glottis and the resultant reduced air pressure permitting the elastic folds to come together once more; the vibratory effect may easily be left by touching the neck in the region of the larynx or by putting a finger over each ear flap when pronouncing a vowel or [z] for example. In typical speaking, a man's voice, this opening and closing action is likely to be repeated between *100* and *150* times in a second, i.e. there are that number of *cycles of vibration* (*referred to as Hertz, which is abbreviated to Hz*); in the case of a woman's voice, this frequency of vibration might well be between *200* and *325Hz*. We are able, within limits, to consciously vary the speed of vibration of our vocal folds in order to change the pitch of the voice; the more rapid the rate of vibration is, the higher is the pitch (an extremely low rate of vibration being partly responsible for what is usually called *creaky* voice). Normally the vocal folds come together rapidly and part more slowly, the opening phase of each cycle hence being longer than the closing phase (see the figure below).This gives rise to *modal*(or normal) voice which is used for most English speech. Other modes of vibration result in other *voice qualities*, most notably breathy and creaky voice which are used contrastively in a number of languages and may be used in English by some individuals and in some styles.

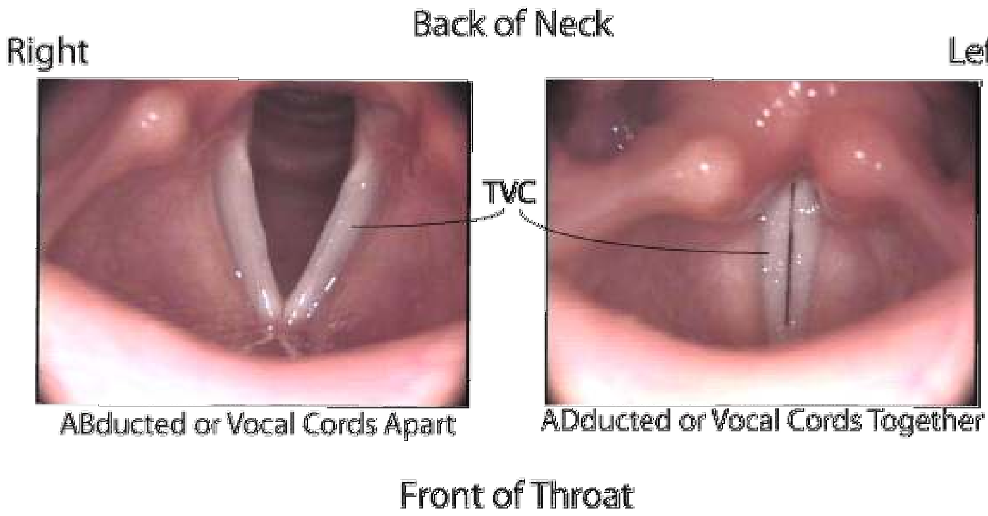


Figure2.3.: Vocal Folds Image Abducted (Apart) and Adducted (Together).[7]

Moreover, we are able, by means of variations in pressure from the lungs, to modify the size of the puff of air which escapes at each vibration of the vocal folds; in other words, we can change the *amplitude* of the vibration, with a corresponding change of loudness of the sound heard by a listener. The normal human being soon learns to manipulate his glottal mechanism so that most delicate changes of pitch and loudness are achieved. Control of this mechanism is, however, very largely exercised by the ear, so that such variations are exceedingly difficult to teach to those who are born deaf and a derangement of pitch and loudness control is liable to occur among those who become totally deaf later in life.

d. One other action of the larynx should be mentioned. A quiet whisper may result merely from holding the glottis in the voiceless position throughout speech. But the more normal whisper, by means of which we are able to communicate with some ease, involves energetic articulation and considerable stricture in the glottis region. Such a whisper may in fact be

---

[7]Adapted from http://fauquierent.blogspot.ca/

uttered with an almost total closure of the glottis and an escape of air in the region of the arytenoids cartilages.

The simplest way of observing the behaviour of the vocal folds is by the use of a laryngoscope, which gives a stationary mirrored image of the glottis. Using stroboscopic techniques, it is possible to obtain a moving record and high-speed films have been made of the vocal folds, showing their action in ordinary breathing, producing voice and whisper, and closed as for a glottal stop in the Laboratory of phonetics and phonology in Paris. The modern technique of observation is to use fibro-optic endoscopy coupled if required with a tiny video camera.

2.2.3. Resonating Cavities

The air stream, having passed through the larynx, is now subject to further and many modifications according to the shape assumed by the upper cavities of the pharynx and mouth, and according to whether the nasal cavities are brought into use or not. These cavities function as the principal resonators of the voice produced in the larynx.

2.2.3.1. Pharynx

Cruttenden (2008), stated that the pharyngeal cavity extends from the top of the trachea and oesophagus, past the epiglottis and the root of the tongue, to the region at the rear of the soft palate. It is convenient to identify these sections of the pharynx by naming them: Laryngopharynx, Oropharynx, and Nasopharynx. The shape and the volume of this long chamber may be considerably modified by the constrictive action of the muscles enclosing the pharynx, by the movement of the back of the tongue, by the position of the soft palate which may, when raised, exclude the nasopharynx and by raising the larynx itself. The position of the tongue in the mouth, whether it is advanced or retracted, will affect the size of the

Oropharyngeal cavity; the modifications in the shape of this cavity should, therefore, be included in the description of any vowel. It is a characteristic of some kinds of English pronunciation that certain vowels, e.g. the [ǽ] vowel in ‘sad’, are articulated with a strong pharyngeal contraction. Additionally, a constriction may be made between the lower raised part of the tongue and the wall of the pharynx so that the friction, with or without voice, is produced, such as fricatives being a feature of a number of languages, for instance Arabic.

The pharynx may be observed by means of a *laryngoscope* or the *fibre-optic nasendoscopy* and its constrictive actions are revealed by lateral *X-ray* photography or nowadays preferably the ***RMI.***

The escape of the air from the pharynx may be affected in one of the three ways:

- The soft palate may be lowered, as in normal breathing, in which case the air may escape through the nose and the mouth. This is the position taken up by the soft palate in articulation of the French nasalized vowels in such a phrase as ‘*un bon vin blanc*’, the particular quality of such vowels being achieved through the resonance of the nasopharyngeal cavities.
- The soft palate may be lowered so that a nasal outlet is afforded to the airstream, but a complete obstruction is made at some point in the mouth, with the result that, although air enters all or part of the oral cavity, no oral escape is possible. A purely nasal escape of this sort occurs in such nasal consonants as [m,n,η] for instance in the English words *ram*, *sun*, *sum* and *sung*. The nasal escape may be accompanied by a friction between a rear side of the soft palate ‘*velum*’ and the pharyngeal wall in case of snoring or defective speech.
- The soft palate may be held in its raised position, eliminating the action of the *nasopharynx*, so that the air escapes solely through the mouth. All the normal English sounds, with the exception of the nasal consonants mentioned above, have this oral escape.

Moreover, if for any reason the lowering of the soft palate cannot be effected, or if there is an enlargement of the organs enclosing the nasopharynx or a blockage brought about the mucus, it is often difficult to articulate either nasalized vowels or nasal consonants. In such speech, typical of adenoidal enlargement or obstruction caused by a cold. The French sentence previously mentioned above would have its nasalized vowels turned into their oral equivalents, and the English word '*morning'* would have its nasal consonants replaced by [b, d, g] becoming [bɔːdig]. On the Other hand, inability to make an effective closure by means of the raising of the soft palate, either because the soft palate itself is defective or because an abnormal opening in the roof of the mouth gives access to the nasal cavities, will result in general nasalization of vowels and failure to articulate oral stop consonants such as [b, d, g](Cruttenden, 2008. pp. 12.13).

It is evident that the action of the soft palate is accessible to observation by direct means, as well as by lateral *X-ray* photography and *Magnetic Resonance Imaging*; the pressure of the air passing through the nasal cavities may be measured at the nostrils or within the cavities themselves.

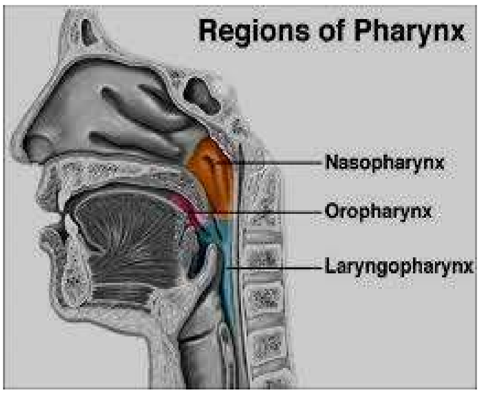


Figure2.4.: Regions of the pharynx (adapted from Saladin,2011. Chapter 22)

2.2.3.2. Mouth

Although all the cavities so far mentioned play an essential and vital part in the production of speech sounds, most attention has traditionally been paid to the behaviour of the cavity formed by the mouth. Indeed, in many languages the word for '*tongue*' is used to refer to speech and language activity. Such as a preoccupation with the oral cavity is due to the fact that it is the most readily accessible and easily observed section of the vocal tract; but there is in such an attitude a danger of oversimplification. Nevertheless, it is true that the shape of the mouth determines finally the quality of the majority of our speech sounds. Far more finely controlled variations of shape are possible in the mouth than in any other part of the speech mechanism.

The only boundaries of this oral chamber which may be regarded as relatively fixed are, in the front, the teeth, in the upper part, the hard palate, and, in the rear, the pharyngeal wall. The remaining organs are movable: the lips, the various part 'subdivisions'of the tongue, and the soft palate with its pendant the uvula. The lower jaw, too, is capable of very considerable movements; its movement will control the gap between the upper and lower teeth and also to a large extent the disposition of the lips. Thus, the space between the upper and lower teeth will often enter into our description of the articulation of sounds; in all such cases, it is clear that the movement of the lower jaw is ultimately responsible for the variation previously described. Movement of the lower jaw is also one way of altering the distance between the tongue and the roof of the mouth.

For our descriptive purposes, it is convenient to divide the roof of the mouth into three parts. Moving backwards from the upper teeth, first, there is the teeth ridge (for which the adjective is *alveolar*) which can be clearly felt behind the upper front teeth; second, there is a bony arch which forms the hard palate which is generally referred to as the roof of the mouth and for which the adjective is *palatal* and which varies in size and arching from one individual to another; and finally, there is the soft palate for which the adjective is *velar* (in reference to the velum) which, as we have mentioned, can be raised or lowered, and at the extremity of which is the uvular (the adjective is *uvular*). All these parts can be readily observed by using a mirror with the mouth open.

i. **Lips**: (adjective *Bilabial* or *Labial*) they are among the movable parts and constitute the final obstruction to the airstream when the nasal passage is shut off. The shape, which may affect considerably the shape of the total cavity. They may be shut or held apart in several ways. When they are held tightly shut, they form a complete obstruction or occlusion to the airstream, which may either be momentarily prevented from escaping at all, as in the initial sounds of both '*path*' and '*bath*', or may be directed through the nose by the lowering of the soft palate, as in the initial sound of '*mad*'. If the lips are held apart, the positions they assume may be summarized under five headings:

a. Held sufficiently close together over all their length that friction occurs between them. *Fricatives* of this sort, with or without voice, occur in many languages, for instance /s/, /z/.
b. Held sufficiently far apart for no friction to be heard, yet remaining fairly close together and energetically spread. This shape is taken up for both vowels like that in '*see*'and is known as the *spread lip position* in addition to the voiceless fricatives;
c. Held in a relaxed position by lowering the lower jaw. This is the position taken up for the vowel of '*sat*'and is known as the *neutral position*;
d. Tightly pursed, so that the aperture is small and rounded, as in the vowel of '*do'*, or more precisely, in the French vowel '*doux*'. This is the *close-rounded position*;
e. Held wide apart, but with slight projection and rounding, as in the vowel of '*got'*. This is the *open –rounded position***.**

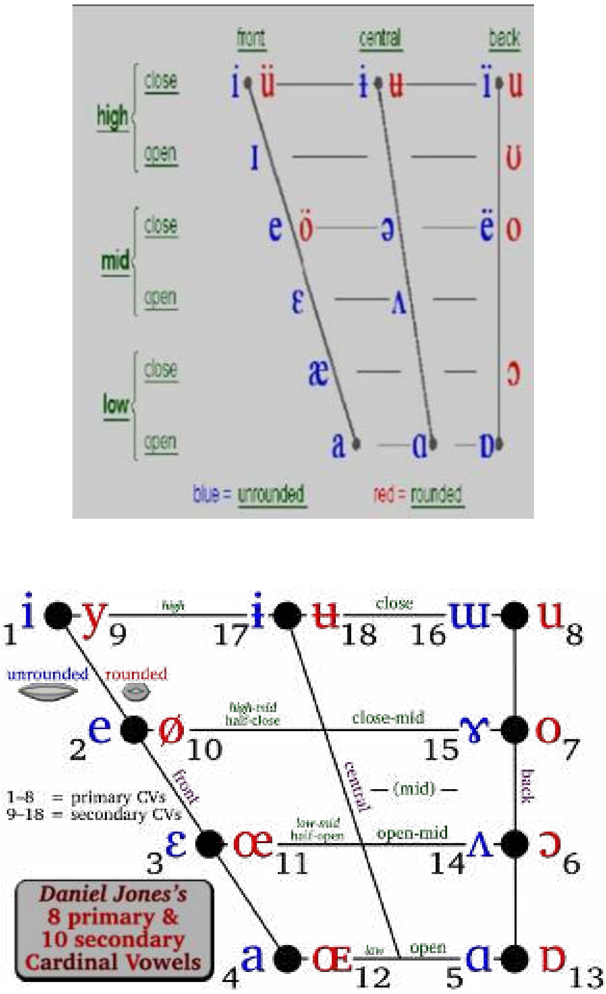


Figure2.5.: Cardinal Vowels (Adapted from Jones,1972)

Variations of these five positions may be encountered, e.g.in the vowel in '*saw'* for which a type of lip rounding between open rounded and close rounded is commonly used. It will be seen from the examples provided that lip position is particularly significant in the formation of vowel quality. On the other hand, English consonants, even including [p,b,m,w]whose primary articulation involves lip action '*bilabial'*, will tend to share the lip position of the adjacent vowel. In addition, the lower lip is an active articulator in the pronunciation of [f,v], a slight contact being made between the lower lip and the upper teeth '*labiodental'***.**

Of all the movable organs within the mouth, *the tongue* is by far the most flexible and is capable of assuming a great variety of positions in the articulation of both vowels and consonants. The tongue is a complex muscular structure which does not actually show obvious sections; yet, since its position must often be described in considerable detail, certain arbitrary division are made. When the tongue is at rest, with its tip lying behind the lower teeth, that part which lies opposite the hard palate is called the *back***,** with the region where the front and back meet known as the *centre* **(**for which the adjective is *central***)** **.**These areas together with the *root* **(**which forms the front side of the pharynx**)** are sometimes referred to as the body of the tongue. The tapering or the narrow section in front of the body and facing the teeth ridge is called the *blade* **(**the adjective is *laminal***)** *and its extremity is the tip* **(**the adjective is *apical***).**The edges of the tongue are known as the *rims***.**

Generally, in the articulation of vowels, the tongue tip remains low behind the *lower teeth*. The body of the tongue may, however, '*be bunched up*' in different ways, e.g. The front may be the highest part as when we say the vowel of /i:/ in '*he***';** or the back may be most prominent as in the case of the vowel in '*who*'**;** or the whole surface may be relatively low and flat as in the case of the vowel in '*ah*'**.** Such changes of shape can be felt if the above words are said in succession. These changes, together with the variations in lip position, have the effect of modifying very considerably the size of the mouth cavity and of dividing this chamber into two parts: that part of the cavity which is in the forward part of the mouth behind the lips and that which is in the rear in the region of the pharynx.

The various parts of the tongue may also come into contact with the roof of the mouth. Thus, the tip, blade and rims may articulate with the teeth as for the '/Z/ and /T/' sounds in English, or with the upper alveolar ridge as in the case of /t,d,s,z,n/ or the apical contact may be only partial as in the case of /l/ (where the tip or the blade makes a firm contact with the alveolar ridge while the rims make none; they are just curved allowing the air to escape laterally);that is why the /l/ is called *lateral***.**

The front of the tongue may articulate against or near to the hard palate, a region that is situated between the alveolar ridge and the hard palate and is called *palato-alveolar*. Such a raising of the front of the tongue towards the palate is an essential part of the /S,Z/ consonants in English words such as '*she'* and '*pleasure'*, being additional to an articulation made between the blade of the tongue and the post-alveolar region referred to as *palato-alveolar* too , or, again, it is a main feature whenever the center of the tongue is against the hard palate to produce a /j/ sound initially as in '*yield'* , which is called *palatal***.**

The back of the tongue can also form a total obstruction by its contact with the front of the soft palate, the back side is being raised in the case of /k,g/ and lowered for /N/as in '*sing*'; or, again, there may merely be a narrowing *constriction* between the soft palate and the back of the tongue allowing a friction. And finally, the uvula may vibrate against the back of the tongue, or there may be a narrowing in this region which causes uvular friction, as at the beginning of the French word '*rouge*'.

Therefore, it will be seen from these few examples that whereas for vowels the tongue is generally held in a position which is out curved or we would rather say *convex* in relation to the roof of the mouth, some articulations, such as the southern British English /r/ in '*red'* and the /l/ in '*table'*, will involve the hollowing of the body of the tongue so that it has, at least partially, a concave contact with the roof of the mouth.

Moreover, the surface of the tongue, viewed from the front, may take on different and various forms, for instance: there may be a narrow groove form running from back to front down the midline as for the fricative voiceless alveolar /s/ in *see*, or the grooving may be very much more diffuse as in the case of /S/ in *ship*. This we will see in the coming chapters on spectrograms.

The speech mechanism is almost accessible to direct observation as far as the lip movements are concerned as are many of the tongue movements which take place in the forward part of the mouth. A lateral view of the shape of the tongue over all its length and its relationship with the palate and the velum can be obtained by means of still and moving X-ray photography and by *Magnetic Resonance Imaging*. It is not, however, to be expected that pictures of the articulation of the vowel /&/in '*cat'* will show an identical tongue position for the pronunciation of a number of individuals. Not only is the sound itself likely to be slightly different from one individual to another, but, even if the sound is for all practical purposes '*the same'* the tongue positions may be different, since the shape of the mouth cavity is not identical between two speakers; and, in any case, two sounds judged to be the same may be produced by the same speaker with different articulations. Therefore, when, we describe an articulation in detail, it should be taken into consideration and understood that such an articulation is typical for the sound in question, but that variations are to be expected.

There is another practical and informative way of not only showing the extent of the area of contact between the tongue and the roof of the mouth, but also of recording tongue movements. At one time the palate was coated with a powdery substance, the articulation was made, and the wipe-off subsequently photographed. But the modern method uses *electropalatography*, whereby electrodes on a false palate respond to any tongue contact, the contact points being simultaneously registered on a visual display. This has the advantage of showing a series of representations of the changing contact between the tongue and the top of the mouth during speech. *Electropalatograms* of this sort are used to illustrate the articulations found in the index.

Conclusion

In this chapter, we have provided a thorough description of speech production, which is the process by which spoken words are selected by the speaker to be produced, have their phonetics formulated and then finally are articulated involving a set of different organs 'motor system' in the vocal apparatus and after that they are to be perceived by the listener. Exchanging the roles between both the speaker and the listener is referred to as the *speech chain*. The process of human communication in which a thought begins in an individual's mind, is articulated, emerges and is transmitted as an acoustic signal, is heard, listened and perceived by the listener, processed by the listener and interpreted in the mind.

CHAPTER THREE

Chapter Three: Pronunciation Teaching

Introduction

Teaching pronunciation involves a variety of challenges. To begin with, teachers often find that they do not have enough time in class to give proper attention to this aspect of English instruction. When they do find the time to address pronunciation, the instruction often amounts to the presentation and practice of a series of tedious and seemingly unrelated topics. Drilling sounds over and over again (e.g., minimal pair work) often leads to discouraging results, and discouraged students and teachers end up wanting to avoid pronunciation altogether.

There are also psychological factors that affect the learning of pronunciation in ways that are not so true of studying grammar or vocabulary. For one thing, the most basic elements of speaking are deeply personal. Our sense of self and community are bound up in the speech-rhythms of our first language (L1). These rhythms were learned in the first year of life and are deeply rooted in the minds of students. Therefore, it is common for students to feel uneasy when they hear themselves speak with the rhythm of either a second or foreign language.

They find that they *sound foreign* to themselves, and this is troubling for them. Although the uneasiness is usually unconscious, it can be a major barrier to improved intelligibility in the target language.

A teacher can help overcome this psychological barrier and other challenges by thinking of the goal of pronunciation instruction not as helping students to sound like native speakers but as helping them to learn the core elements of spoken English so that they can be easily understood by others. In other words, teachers and students can overcome the frustrations, difficulties, and boredom often associated with pronunciation by focusing their attention on the development of pronunciation that is *listener friendly*. After all, English pronunciation does not amount to mastery of a list of sounds or isolated words. Instead, it amounts to learning and practicing the specifically English way of making a speaker's thoughts easy to follow.

Strategic learning and pronunciation learning are both areas of study that have recently received wide-spread attention in second language research. We can refer to different researchers who explored this virgin area, for instance, (Brown, 2001; Bruen, 2001 Celce-Murcia, Brinton, & Goodwin, 1996; Dornyei&Skehan, 2003; Fan, 2003; Norton &Toohey, 2001). Strategic learning research has sought to advance the understanding of how students tackle difficult language learning tasks using learning strategies (Chamot, 2004; Chamot& El-Dinary, 1999; El-Dib, 2004). On the other hand, the field of pronunciation learning research has attempted to discover which areas of pronunciation are most beneficial for instructors to teach (Celce-Murcia, Brinton, & Goodwin, 1996; Derwing, Munro &Carbonaro, 2000; Florez, 1998; Riney&Flege, 1998; Riney, Takada, & Ota, 2000). Little crossover between these two fields has taken place, so second language researchers have yet to discover how second language learners tackle difficult pronunciation learning tasks through the use of learning strategies.

This study is meant to help bridge the existing gap between strategic learning research and pronunciation learning in a Communicative Language Teaching (CLT) context. The gap is currently very large for several reasons. First, language-learning strategies have only been examined in light of pronunciation learning and much research is still required to determine what pronunciation learning strategies exist (Peterson, 2000). Second, no categorization scheme exists whereby pronunciation strategies can be organized into pedagogically founded groupings. Third, no overall theoretical construct currently ties pronunciation-learning strategies to pronunciation acquisition theory. Finally, no study has examined to what extent students' ability to develop and use pronunciation-learning strategies correlates with actual pronunciation skill.

With so many unexplored factors in the new domain of pronunciation learning strategies, the field is both overwhelming and ripe for research. One way to investigate the effect of strategic learning on pronunciation learning is to examine the use of pronunciation learning strategies used intuitively by language learners. Learning strategies have had their greatest impact in language learning on the four major skill areas of second language learning: speaking, listening, reading, and writing. Only a limited number of studies have requested students to reflect on their pronunciation learning techniques and report the strategies they use (Peterson, 2000). This is extremely unfortunate since second language research has begun to show that pronunciation learning is an essential element of second language acquisition (SLA) (e.g. Celce-Murcia, Brinton, & Goodwin, 1996; Hinofotis& Bailey, 1980; Derwing&Rossiter, 2001; Vitanova& Miller, 2002).

Another way to investigate the effect of strategic learning on pronunciation learning is to ascertain the degree to which pronunciation students use personal pronunciation learning strategies in the major areas of pronunciation learning: *input*, *noticing*, *practice*, and *feedback*. It is, after all, not the number of learning strategies a student uses; it is rather the ability students have in developing a set of personal learning strategies, which determines good language learners (Chamot& Rubin, 1994).

Therefore, a student who reports the ability to successfully develop and implement theoretically informed pronunciation learning strategies will likely be a good pronunciation learner.

We aim at investigating techniques to look at what strategies are better used to help students improve their pronunciation.

Pronunciation involves far more than individual sounds. Word stress, sentence stress, intonation, and word linking all influence the sound of spoken English, not to mention the way we often smear words and phrases together in casual speech. English pronunciation involves too many complexities for learners to strive for a complete elimination of accent, but improving pronunciation will boost self-esteem, facilitate communication, and possibly lead to a better job or at least more respect in the workplace. Effective communication is of greatest importance, so we choose first to work on problems that significantly hinder communication and let the rest go. Remember that our students also need to learn strategies for dealing with misunderstandings, since native pronunciation is, for most, an unrealistic goal.

A student's first foreign language often interferes with English pronunciation. For example, /p/ is aspirated in English but not in French, and does not exist in Arabic, so, when a French speaker pronounces 'pig' without a puff of air on the /p/, an American may hear 'big' instead. Sometimes the students will be able to identify specific problem sounds and sometimes they will not. We can ask them for suggestions, but we will also need to observe them over time and make note of problem sounds. For instance, we encountered a student from Khenchela (a city in the east of Algeria), who was unable to pronounce /p/, at the beginning; we thought she did not hear us, but after several times, we figured out that she had a problem in pronouncing the voiceless bilabial occlusive sound... Another challenge resulting from differences in the first language is the inability to hear certain English sounds that the native language does not contain. Often these are vowels, as in *ship* and *sheep*, which many learners cannot distinguish. The Japanese are known for confusing /r/ and /l/, as their language contains neither of these but instead has one sound somewhere between the two. For problems such as these, listening is crucial because students cannot produce a sound they cannot hear. Descriptions of the sound and mouth position can help students increase their awareness of subtle sound differences. That is why we highly recommend the use of Praat software, as it contains even different tests of perception; the *identification* and the *discrimination* tests which when used correctly, afford satisfactory results about how well students perceive the sounds in particular and the speech in general. But as far as pronunciation is the core of our work, many inquiring questions should be asked and dealt with.

3.1. The Need to Teach Pronunciation

Most ESL or EFL teachers now agree that explicit pronunciation teaching is an essential part of language courses. On the one hand, confidence with pronunciation allows students in particular and learners in general the interaction with native speakers that is so essential for all aspects of their linguistic development. On the other hand, poor pronunciation can mask otherwise good language skills, condemning learners to less than their deserved social, academic and work advancement.

While there is little doubt about the teachers' appreciation of the importance of pronunciation instruction, there is even less doubt about learners' own demand for effective pronunciation teaching: almost all learners rate this as a priority and an area in which they need more guidance. For any remaining sceptics, it may be worth briefly rehearsing the following responses to reasons sometimes given for not teaching pronunciation explicitly in an EFL program.

a. It is true that students are very unlikely to attain a native-like accent - but their intelligibility can be greatly improved by effective pronunciation teaching;

b.It is true that pronunciation improves most through the gradual intuitive changes brought about by real interaction with native speakers – for instance chat rooms, but for a large proportion of EFL students, the skills that enable this type of interaction do not come naturally; most need a *leg-up* from explicit pronunciation teaching.

c.It is true that it is offensive to prescribe an *accent norm* to which students must assimilate, and it is true that people should be free to express themselves in whatever accent they choose- but it is not true that this freedom is given by withholding pronunciation teaching. On the contrary, it is effective pronunciation teaching that offers learners a genuine choice in how they express themselves.

3.2. To Teach Pronunciation and How to teach it

Despite widespread agreement about the importance of pronunciation teaching, in EFL courses around the world, pronunciation is the aspect of language that receives least attention or simply none. The reason is not unwillingness to teach pronunciation, but uncertainty as to how best to help learners. This is of course not to discount the contribution of significant numbers of gifted pronunciation teachers all around the world. The problem is that the effectiveness of these teachers is based largely on individual experience and insights. No training in pronunciation teaching is given in TEFOL programs at least in our department of English. For those already teaching, who seek to improve their skills, the literature on phonological theory and English phonetics is often (with notable exceptions) found to be opaque, and of little applicability in the classroom.

The question, then, is not whether to teach pronunciation, but how to teach pronunciation. We have been reading much about the theoretical approach to phonology that we think is useful in this regard. It sees phonology not as computational processing, but in a *communicative* and *meaning-based framework*. Without going into technical detail here, we would like to consider some implications for pronunciation teaching.

There are many ways of teaching pronunciation, and many different opinions as to which ways are the best or most effective. However there has been to date relatively little serious comparative research on what really works in helping learners of a second language with pronunciation. This is an area which needs considerable improvement. Nevertheless, there are a few things which are becoming well established as key factors in effective pronunciation tuition.

In this section, we outline some of the pronunciation-teaching practices that have been shown to be effective, and then set out some concepts that are necessary in understanding why these particular practices are effective.

3.3. Efficient Means to teach Pronunciation

It is important to emphasize that pronunciation teaching is currently undergoing a revival after several decades of neglect. There are many questions requiring detailed research and empirical investigation. The account presented here represents a *best guess* for which there is considerable evidence but which is most certainly not the last word on the subject.

Here are some of the factors that have been shown to be most relevant in creating good outcomes in pronunciation teaching. The first three are becoming more widely known and accepted. The last, though, is less well understood. It will be given more extensive discussion below.

- Pronunciation teaching works better if the focus is on larger chunks of speech, such as words, phrases and sentences, than if the focus is on individual sounds and syllables. This does not mean that individual sounds and syllables should never be referred to; it simply means that the general focus should be on the larger units. It is the reason why we selected a sentence test rather than single words.
- Pronunciation lessons work best if they involve the students in actually speaking, rather than in just learning facts or rules of pronunciation. Many students of course feel more comfortable learning the rules of the language, because it is less threatening than actually speaking. However, the transfer of explicit knowledge of rules into pronunciation practice is very limited. Teachers need to devise activities which require learners to actually speak in their pronunciation classes.
- Learning pronunciation requires an enormous amount of practice, especially at early stages. It is not unreasonable for learners to repeat a particular phrase or sentence twenty or fifty times before being really comfortable with it. Unfortunately, *drilling* has been out of favour in language classes for some time, due to association with several bad aspects of the behaviourist method of teaching. Indeed some forms of drilling are at best a waste of time, and can even be a hindrance to learning. However, drilling of real, useful phrases which can actually be used outside the classroom is highly advantageous to learners.

- Pronunciation teaching requires thorough preparation through work on the perception of English sounds and contrasts, and the formation of concepts of English phonology. The latter has never been taught in our department.

According to Kolb (1984), there are different strategies of teaching pronunciation. The following chart presents a synthesized taxonomy of pronunciation strategies using a theoretical framework which accords with pronunciation acquisition theory. Some strategies may be listed in two different categories. The reason for this overlap is that some categories compassed several steps of the pronunciation acquisition construct. For example, intent listening was listed in both the input and noticing sections because a learner may encounter sounds when listening intently (input) and can be able to notice important pronunciation distinctions at the same time (noticing).

Kolb's (1984) Learning Cycle Construct	Pronunciation Acquisition Construct	Pronunciation Learning Strategies
Concrete Experience	Input / Practice	Input •Intent listening •Focusing on articulatory gestures of others •Active listening •Eagerly listening to new sounds •Putting self in proximal points for hearing L2 pronunciation: TV, Movies, Radio, etc. •Representing sounds in memory •Focusing on individual syllables of words Practice •Reading aloud •Practicing new sounds •Imitating and/or mimicry of native speakers •Practicing 'mock talk' or imitating L2 prosody using L1 words •Talking aloud/role-play •Memorizing the pronunciation of words •Helping facial muscles become accustomed to accommodating L2 pronunciation •Practicing different sounds, first in isolation and then in the context of words •Repeating after tapes in a language laboratory

Kolb's (1984) Learning Cycle Construct	Pronunciation Acquisition Construct	Pronunciation Learning Strategies
Reflection on Observation	Noticing /Feedback	Noticing • Noticing the intricate differences between L1 and L2 pronunciation • Focusing on suprasegmentals of language • Intent listening • Distinguishing errors among other speakers • Focusing on articulatory gestures of others • Listening carefully to errors made by native speakers to infer key sounds or structures • Acquiring a general knowledge of phonetics Feedback • Self-monitoring • Focusing on suprasegmentals of own speech. • Using phonetic symbols and transcriptions • Monitoring and eliminating negative interference • Active listening • Asking for help • Cooperating with peers
Abstract Conceptualization	Hypothesis Forming	Hypothesis Forming • Monitoring and eliminating negative interference • Self-correcting • Acquiring a general knowledge of phonetics • Doing special exercises for sounds not existing in the learner's native language • Finding out about the target language pronunciation
Action Based on New Conceptualization	**Hypothesis Testing**	**Hypothesis Testing** **• Repeating new words according to new hypotheses** **• Skipping difficult words** **• Rehearsing sounds** **• Using proximal articulations** **• Increasing or decreasing volume of speech** **• Using a slower rate of speech** **• Using clear speech** **• Lowering anxiety**

Table: 3.1. Connection between Kolb's (1984) Construct, SLA, and Pronunciation Learning Strategies

3.4. What to Avoid in Teaching Pronunciation

The demonstrated ineffectiveness of drilling phonemes, minimal pairs and formal or stilted dialogues was a large part of the reason that pronunciation was all but dropped from language classes with the introduction of the communicative methods. Basing lessons around detailed descriptions of the articulation of sounds, intensive IPA transcription, or lectures on

English phonology are also minimal in their effectiveness, while at the same time placing high knowledge demands not just on learners but also on teachers. Many teachers consider their own expertise in these areas to be inadequate, preferring to delegate pronunciation classes to those with more knowledge. The point we wish to drive attention to is that while knowledge of English phonetics and phonology is certainly useful, it is not, in itself, what is needed for EFL pronunciation classes. Much more importance is for teachers to have insight into the kinds of problems learners face in pronouncing English, and tools to provide for their needs at different stages.

What if we change the focus of pronunciation for a while, from teachers to learners and try to investigate students' perception and responses about pronunciation.

3.5. Difficulty of Learning Pronunciation

Pronunciation is one of the most difficult areas for learners as well as for teachers. In quest of effective teaching, it is worth diagnosing carefully the nature of the difficulties that they face.

a. There is a significant skill component for students. Pronunciation is not just a cognitive *knowing-that*, it is also a physical *knowing-how*, similar to playing a sport or musical instrument. Students need *motivation* and *time* to really practice pronunciation. It is very much worth spending class time discussing with students their own ideas about what is involved in learning pronunciation. They are often surprised by the suggestion that they should practice speaking. Students need help in overcoming both their wishful idea that pronunciation is a subject like, say, history, which can be learned merely by listening to a teacher, but what if a teacher does not pronounce English correctly? And the psychological and social barriers that make it difficult for them to practice effectively. Wehave never heard our students practicing their English outside classrooms...Taking into account that even in classrooms, their participation remains insufficient to allow them refining their pronunciation.

b. This skill, however, is only a component. There is also a significant cognitive component in pronunciation learning, which is much less often acknowledged. Indeed, we believe it is useful to think of learning to pronounce a new language as involving a kind of concept formation, about which there is a large literature in psychology and education, rather than as a purely physical skill. Both teachers and students themselves generally assume it is the latter and can benefit greatly from a better understanding of the conceptual aspects of pronouncing a new language. For instance, very often a sound that causes great difficulty to a student is one that they can produce quite acceptably in a different context. As just one example: a French student who mispronounces /T/ in *thing*, thumb etc., as /s/ might use this very/T/ sound as a mispronunciation of the /ʒ / in or *father as/z/*. Helping them to realize this and to reallocate the sounds to more appropriate phonemes might well be more useful than instruction in the articulation of the dental fricatives.

The nature of phonological systems means that there is very often a radical difference between:

a. What students think they are saying (their description of their own speech),
b. A phonetic description of the sounds they are actually producing, and
c. How someone from a different language background (e.g. a teacher) describes their speech.

This creates scope for significant miscommunication between teachers and students regarding the nature of any errors they are producing. We think many teachers are only partially aware of how unlike their students' perception of speech is from their own. An over-emphasis on the notion of *transfer* in the literature on second language acquisition can increase the teachers' sense that what learners are doing is *mixing up (say) /r/ and /l/*. What is actually needed in my view is for teachers to change this interpretation of what is going on. It suggests that students are using the same descriptions of sounds as teachers are, whereas in fact the problem is that the students are unaware that there is any distinction between the sounds. To them, they are alike.

Another area in which miscommunication between teachers and learners is evident is in discussion of syllables and word stress. The number of syllables people think they are producing can be quite different from the phonetic reality. For example, the same sequence of sounds, say *sport*, which to English speakers has one syllable, when borrowed into Japanese for instance, is interpreted as having three syllables. (Note that the pronunciation in connected speech is phonetically very similar in both languages - it is the phonological interpretation that differs.) This can seem quite bizarre and even strange to English speakers until we realize that we are producing two syllables in *support* is often equally incorrect from a phonetic point of view (English words like *support*, *police*, etc., are often pronounced with the first vowel phonetically non-existent, that is the elision of the weak vowel schwa). The difference between what people think they are saying and what they are actually saying means that the level of insight people have into the phonetics of their pronunciation is generally very poor. Giving learners a phonetic description of the target sound is often not helpful, even if it is done accurately. In our opinion, discussion of the articulation of sounds should be restricted to those which can actually be seen and felt by the learner: lip rounding and spreading, tongue between teeth, etc. In other cases, especially for vowels, it is much more useful to concentrate on training the students' perception, and helping them develop better ways of thinking about the sounds of the new language - rather than in giving anatomy lessons. The interesting thing is that people act, in pronunciation as in many things, not on the basis of the phonetic reality of sounds they hear, but on the basis of their description or concept of the sounds. A nice example of this comes from the case of word boundaries. Many learners speak in a jerky disconnected way, because they separate the words, rather than stringing them together… (What we refer to as a connected speech). This is hardly surprising - most speakers of whatever language believe that their words are separated by short pauses which equal the punctuation in writing. In fact of course, all languages run words together in connected speech (though the exact way in which the words are run together varies greatly from language to language). When people learn a new language, they will naturally try to keep the words separate. Overcoming their misconception that separating the words makes their speech clearer, and encouraging them to really listen to how words are said together in the new language, can be more effective than instruction in the rules of English co articulation - which may make little sense to them if they still subconsciously believe that clear speech has separate words. A large part of a teacher's skill, then, is in giving the learners new ways of thinking about or conceptualizing words and sentences in the new language. This skill can include knowledge of articulation and abstract rules of phonology but it is not dependent upon it. It is a skill which many teachers have or develop but which needs greater acknowledgement and explicit attention to start by the suprasegmental features which unconsciously will encompass the segmental ones.

3.6. How to teach Pronunciation

Given these considerations about the nature of pronunciation, we think we should first of all set some objectives and goals about how and what to teach in pronunciation, in other words, better sow to reap well. Some recommendations about pronunciation teaching are afforded.

3.6.1. Pronunciation in a Communicative Context

Learners benefit greatly from explicit explanation of how pronunciation fits into the overall process of communication. A very simple model of communication, showing a listener trying to interpret a message on the basis of cues in the speakers' speech, is sufficient. This gives learners a framework within which to understand what goes wrong when they are not understood or are misunderstood, and to gain a clear, practical idea of the nature of linguistic contrast - not just a classroom drill with *thigh* and *thy*, but the living basis of our ability to communicate in real life contexts.

3.6.2. Advantages of the Framework

These advantages include:

a.It takes learners' focus away from their own *performance* and places it more clearly on the listener's experience of their speech. This can be very helpful in reducing anxiety and the expectation of failure. Discussion of learners' experience of listening to foreign speakers of their own language can help them see how tolerant listeners are (in terms of understanding accented speech) and give them a sense that accents are nice - it is incomprehensibility that is bad, not the accent as such. These considerations can help give learners enough confidence in their segmental production to allow them the fluency and rhythm so important to intonation.

b.It changes the goal of pronunciation from one of mimicking a native accent (which is a bit difficult to achieve), to one of creating intelligible messages (perfectly possible). Errors can be defined in terms of intelligibility rather than in terms of non-attainment of a perfect model, which allows much more scope for teachers to encourage successful communication rather than constantly focusing on deviations from native-like production. The rules of English can be defined in terms of what listeners need in order to understand a message correctly and easily, which makes them more meaningful and easier to relate to real speech.

c.It allows a blurring of boundaries between segmental and suprasegmental aspects of speech, and an easy way in to teaching learners about the information structure of speech, which is highly useful in teaching English prosody. Following from the emphasis on the listener's experience of their speech, they can learn that in English we use stress to highlight the information the speaker considers will be unpredictable to the listener. Many other languages of course do not use stress for this function – the fact that English does needs to be explicitly taught and demonstrated. But spending time on this can give learners a *handle* on understanding stress and intonation in terms of the meaning of the message, rather than as a set of classroom rules.

3.7. Student- Centred Teaching

To enhance students' pronunciation learning, a wide range of techniques may be introduced, which will definitely have a positive effect on their performance, facilitating a strong social learning experience to emphasize the transition from the traditional practice, where the teacher is the sole source of knowledge, towards a Learner Centred Approach.

3.7.1. A Learner-Centred Approach

This type of teaching naturally encourages the use of naturalistic exercises and practice of real communicative situations. Classes must be learner-centred in the sense that learners should be able to practice speech that will be directly useful to them in their real lives. We think it is essential that the basis of all classroom exercises should be phrases and sentences. Not that individual sounds and words should ever be discussed; indeed they should, but always in the context of the larger structures of communicative language use. Students should be encouraged to bring examples of communication failure to class for work shopping – much can be gained from discussion of why a learner got a cup of tea instead of the cappuccino they ordered after lunch. They should also be encouraged to anticipate situations they will encounter after the class, and practice speech that will be useful in those situations. In these ways, the material learned in class will have maximum transferability to the learner's real world. But there are more profound respects in which pronunciation teaching must be learner-centred. We have seen already that teachers need to be aware that students actually hear speech very differently from the teachers themselves do. This seems to be a natural and essential extension of efforts in all other areas of language teaching to be culturally sensitive in communication with students. Teaching can be learner-centred by developing skills in communicating with learners about speech and pronunciation in ways that make sense to the students - as opposed to giving them phonetically accurate descriptions. This ability to start from where students are in order to lead them to new understanding is the basis of all effective teaching. Even more important is for students to be *learner-centred*: to develop their own skills in what is called *critical listening*- the ability to notice, diagnose and repair their own errors, and those of their classmates, rather than always relying on the teacher's feedback. It is through critical listening that perceptual discrimination, and appropriate conceptual analysis of English words and sentences into sounds and letters, can best develop. Many indications suggest that it is these that are the foundation of improvements in whether second or foreign language pronunciation (Fraser 1999).

3.7.2. A Need based Curriculum

When we teach anything, we start by helping students acquire some basic concepts on which they can build more complex understanding. For example, if we teach science, we first make sure students have studied in a scientific stream and have a basic understanding of solids and liquids before we teach them about molecules, atoms and subatomic particles. Sometimes teaching the elementary concepts involves letting students believe some things that are not strictly accurate but which help their understanding. Later on, they can go on to refine their concepts.

It is exactly the same with teaching pronunciation. Although in many cases we do not have the opportunity to establish and follow a full curriculum course on pronunciation with students, it is always important to offer students help at a level appropriate to their needs. This means having a rough curriculum for pronunciation teaching in our minds so that we can access material relevant to particular situations.

There are many ways of developing a pronunciation curriculum. For example, some people like to work through the various classes of phonemes or contrasts in order; others like to tackle *common problems*, such as /r/ and /l/ or *vowel length*, one at a time; others like to have lessons on topics such as *questions vs. statements* or *contrastive stress*.

In the communicative approach, the order in which pronunciation needs are addressed is based on the needs of the people who will be listening to the learners (i.e.ordinary native speakers of English), and the curriculum involves helping learners acquire the concepts most relevant to making themselves understood in English. In other words, the *curriculum* for pronunciation is based on the relative importance of different aspects of pronunciation in terms of how they affect listeners' comprehension.

If the stress pattern of a phrase is correct the phrase can be comprehended in context even though some other aspects are incorrect. However, even if the consonant pronunciation is perfect, the overall meaning of the message will be missed if the stress pattern and vowel characteristics are not given correctly. Since our goal is to help students to acquire functional oral communication, we start with aspects of pronunciation that most affect listener's comprehension. Once they can manage functional oral communication, they can certainly go on to improve the details of their pronunciation.

In general, it is not a good idea to communicate about pronunciation solely in words – at least until we have built up a deep understanding with students about the metalinguistic vocabulary we use. It is very important to use audio and visual aids to help them understand what is meant. Simple visual representations of the words in ordinary spelling with a few well-chosen annotations are usually the best for students.

Another important concept for teachers is to test students often using either formal or even informal tests to be sure they really understand the terms and concepts which are used. We may ask them to tell us about the stress pattern of words. We ask them to tell us whether the pronunciation of a word is right or wrong. If they have a problem with these tasks, we can take a step back and go over some key concepts.

3.7.3. Best Knowledge in Teaching

In relation to pronunciation, many students believe that they need information, e.g. about articulation or grammar, in order to overcome their perceived inability to pronounce the sounds in question. Many students believe they need to master the *phonetic alphabet* (really the phonemic alphabet) in order to learn pronunciation. Many students believe they have no right to speak unless they can sound like a native speaker. Many students believe that learning the rules of English phonology is the same thing as learning pronunciation.

As we have seen, all of these beliefs are at least partially false. Without going into a detailed explanation to learners, it is important not to just *give them what they want*. We have often heard justification of the use of vocal tract animations with *but that's what the students want*. This may be so, but the students' greater want is to learn pronunciation.

The same goes for practicing real speech in class, or for getting learners to record their voices and listen to them critically. While you cannot force adults to do things your way, you can certainly encourage them and give them confidence that your way will work. Indeed it is your responsibility to do this.

3.7.4. Self-reliant Learners

Many students have quite incorrect ideas about what is involved in learning pronunciation – or in learning a language in general, for that matter. For example, many students believe that learning vocabulary involves writing words on cards and storing them in a card file! Certainly doing this is useful but the learning only happens when the cards are actually used. In regard to pronunciation it is useful to tell learners that pronunciation is a skill that involves both thinking and doing – just like learning a sport or a musical instrument.

It is also very useful to give learners a framework within which they can think about pronunciation, can understand and extend the information we give them, and even, as they become more experienced, we may from learn their own mistakes.

Give learners themselves a simplified version of the idea that in communication it is not what we say that matters but what the listener understands. Help learners understand the importance of helping their listener by speaking loudly enough and slowly enough that the listener can process their speech, not just rushing to finish with their ordeal! You might even like, with some learners, to discuss their own experience of listening to a foreign learner speaking their own native language.

Sometimes they find this interesting and are encouraged to realise that learners can be quite comprehensible even if they have an accent.

3.7.5. Good Teaching and Giving Opportunities to Practice

Although we have emphasized the cognitive aspects of pronunciation, this has been purely to redress the balance in favour of an often neglected aspect. In reality, pronunciation is a skill, and practice is just as important as cognitive understanding.

One of the main values of a classroom situation is that it gives learners a safe place to try out and rehearse the speech they will need to use in *the real world*. We do not let them wriggle out of practicing by saying they are embarrassed! We can just encourage them by saying: “If *you are going to make a mistake, it is better to make it here with me than out there where it really matters*.”

A good method to use with a large group is to let the students practice in chorus for several repetitions, then choose one student for individual rehearsal, go back to chorus rehearsal, then choose another student, go back to chorus, and so on. You might think they would get bored with this, but as long as the material is useful and challenging, students generally love this kind of work.

3.7.6. Motivating and Encouraging

Sometimes learners feel that learning pronunciation is a hopeless task because there is so much to learn, or because of previous teachers' avoidance of the issue, or inability to teach pronunciation effectively. It is important to give students a feeling of confidence and optimism. This of course depends upon the teacher really believing that it is possible to learn pronunciation, and having confidence in the approach they are learning. Ultimately this comes only through experiencing success, but it is also a frame of mind.

We try to notice and point out to students the positive aspects of their pronunciation, and praise any improvement, even if we feel there is still a long way to go. We try to show confidence in the process that learners are going through, and to build up a sense of what works and what does not.

On the other hand, we do not tell students they are perfect if they are not! We give praise in relation to any improvement we do notice, and encouragement that we are sure they will get it eventually.

3.8. Pronunciation and Intelligibility

Most teachers would claim that a native-like pronunciation is liked and preferred for their students (Stern, 1983). However, it is accepted that intelligibility is the most sensible goal of teaching pronunciation (Dalton and Seidlhofer, 1994).The question here would be "what is meant by intelligibility"?.

Intelligibility is *being understood by a listener at a given time in a given situation* (Kenworthy, 1987, p. 13).It seems that understandability seems to replace intelligibility. Hence intelligibility "refers to *the degree to which a listener understands a speaker* and comprehensibility is a judgment of how easy or difficult an individual's pronunciation is to understand"(Derwing, 2010, p.29). Consequently, many researchers now highlight the importance of intelligibility and comprehensibility over a native-like accent. It is argued that "*………with an increasing focus on communication,…it has become of critical importance to provide instruction that enables students to become, not 'perfect pronouncers' of English , ………but intelligible, communicative, confident users of spoken English for whatever purposes they need"* (Jenkins,1998, p. 489).

Dalton and Seidlhofer (1994, p. 9) tried to show understandability in the following phrases:"*making yourself understood is not just a matter of accurate and clear articulation, and pronunciation cannot be considered or taught in isolation, dissociated from questions such as "why do we talk"?, what determines how we talk to whom?, and getting our meaning across?*"

According to Dalton and Seidlhofer (1994), although the foreign learner cannot produce sound features accurately, it is quite possible for the listener to approximate or match the sound feature heard with the sound used by native speakers.

Hence, intelligibility, as a goal means that we are aiming at and for "*something close enough to native like performance*" (Kenworthy, 1987, p.13)

3.9. Intelligibility or Accuracy

Trying to imitate native pronunciation is actually a challenging task for any learner of English as a foreign language. Some students are really better than others at imitating not only sounds but speech patterns too. . Therefore, we should not expect almost all our students to pick up sounds and intonation patterns instantly and theoretically unless we put them into practice using software to enable students to immediately visualize their speech and the pitch variations.

Intelligible pronunciation is an essential component of communicative competence (Morley, 1991, p.488). Therefore, intelligibility as a goal means that we are aiming for "*something close enough to native like performance*."

It has been mentioned that intelligibility is related and dependent on 'counts of sameness', but there are some other aspects of speech affecting a person's intelligibility. For instance, if a person's speech is full of hesitations, self-corrections, and grammatical restructuring, then a listener will be misled , and will tend to find the speaker difficult to be understood, because it has been proved by Kenworthy (1987, P. 13) that speakers who hesitate a lot also tend to exhibit many pronunciation problems.

Another factor is associated to the speaker's speed. When we feel that someone is speaking fast, we face difficulty in grasping what is being said; more often , it is not the speed that causes the misunderstanding, but rather the fact that we do not seem to pick out the most important bits – the crucial words from the less important bits. Features like stress, rhythm and intonation are essential in showing the important bits of a message, and then, there would be less intelligibility problems in a rapid speech. That is why, familiarity and exposure to the target language are of paramount importance and thus, considered as intelligibility factors. A person who is exposed to different accents of English will find it easier to understand those speakers with different accents for the simplest reason is that he/she is familiar with those features of pronunciation. That is to say, the more opportunity an EFL learner has to listen to different varieties of English, the more intelligible those accents are to him/her.

3.10. Components of Pronunciation

We have mentioned so far intelligibility and the factors that enhance pronunciation learning. To master the previously mentioned factors of intelligibility, one should be aware of the various aspects of pronunciation, which are classified into three major components: segmentals, suprasegmentals and aspects of connected speech.

3.10.1. Segmental vs.Suprasegmental

Speech parameters are the segmental physiological articulatory elements upon which the *suprasegmental* or *prosodic* elements (*duration*, *intensity*, *melody* and *pitch*) are superimposed.

Segmental: Vowels, consonants, tone and phonological duration**.**

Suprasegmentals: "*Suprasegmental features are established by a comparison of items (segments) in a sequence."(*Lehiste, 1970)

3.10.1.1. Segmental

Segmental aspects of accent and pronunciation are individual sounds, also known as phonemes, meaning consonants and vowels. Syllables and one-syllable words are considered as segmental aspects of accent and pronunciation, as well. Segmental aspects of accent and pronunciation are not heard as being affected over more than one sound segment. For instance, the first /s/ in business always sounds like /z/.

3.10.1.2. Suprasegmentals

Suprasgemental refers to how speech sounds function, and are affected at the sentence and discourse level, which is to say over multiple sound segments that combine to make phrases, clauses, and sentences. Improvements in suprasegmental aspects of English accent and pronunciation play an important role in the production of comprehensible speech patterns in English and are key factors in accent reduction and speaking ability in general.

Suprasegmental aspects of accent and pronunciation are intonation, added stress, time-stress, tone, linking and connecting, thought groups, volume, speed, reduction, assimilation, deletion, weak forms of function words, conversational contractions, and standard contractions which are also used in writing.

3.10.2. Visual Feedback in Suprasegmental Pronunciation Training

Prosody plays a critical role in pronunciation training because it contributes the most to comprehensibility of speech. Pronunciation training with use of computers has become an imminent part of Computer Assisted Language Learning CALL; at first, the systems were able to provide audio feedback only (recording and playback of student's speech); later, with the use of speech analysis systems the feedback became more and more useful. Teacher's research of these displays showed that for pronunciation training,visual feedback combined with the auditory feedback, is more effective than auditory feedback alone.'

However, there is still a long way to go to achieve human teachers' level of ability to provide students with meaningful, corrective procedures and feedback.

The biggest challenge for a pronunciation training system is to show users relevant features of speech without overwhelming them with information. There exists a compromise between how rudimentary is the analysis of a user's speech and how easy it is for an average user to understand the resulting visual feedback. For example, speech spectrogram is easy to calculate, but teaching students and teachers, '*who are not used to see these spectrograms and what these displays mean might take longer than the pedagogical potential their use might warrant.*'We aim at "showing displays that are useful, making them easy to interpret, and that assist in pronunciation learning"(Price, 1998).

The focus here is the use of computers for suprasegmental pronunciation training. We will speak primarily about pronunciation training for our students of English in the coming chapters in the field work.

3.10.3. Prosody in Pronunciation Training

Acquisition of prosody, in addition to grammar and vocabulary, is essential for learners of any foreign language. Several methods can be used to train these learners in order to learn and acquire prosody,

Recently, there has been more interest in exploring automated methods to measure prosodic features of pronunciation. For example, Levow et al (2009) used a SVM (Support Vector Machine) based classifier for pitch accent recognition. Support Vector Machines provided a fast, easily trainable classification framework that had been proven effective in a wide range of application tasks. Honig et al (2009) used a large feature set based on duration, energy, pitch and pauses to detect word accents.

3.10.3.1. Importance of Prosody

The following Prominent language and speech specialists agree that prosody plays a crucial role in comprehensibility of speech and that prosodic training is imperative in pronunciation training:

- *Ordinary people who know nothing of phonetics or elocution have difficulties in understanding slow speech composed of perfect sounds, while they have no difficulty in comprehending an imperfect gabble if only the accent and rhythm are natural.*(Bell, 1916)

*Learners who use incorrect rhythm patterns or who do not connect words together are at best frustrating to the native-speaking listener; more seriously, if these learners use improper intonation contours, they can be perceived as abrupt, or even rude; and if the stress and rhythm patterns are too non-native like, the speakers who produce them may not be understood at all.(*Celce- Murcia, et al, 1996)

- *When students start to learn a new language, some time is usually devoted to learning to pronounce phones that are not present in their native language. Yet experience shows that a person with good segmental phonology who lacks correct timing and pitch will be hard to understand. Intonation is the glue that holds a message together. It indicates which words are important, disambiguates parts of sentences, and enhances the meaning with style and emotion. It follows that prosody should be taught from the beginning.(*Eskinazi,1999)

3.10.3.1.2. Some Examples of Suprasegmental Pronunciation Problems

Suprasegmental pronunciation problems are omnipresent. Different acoustic or even grammatical means are necessary to produce the same prosodic results in different languages, thus creating communication problems for non-native speakers of a language. For example, Japanese speakers of English put equal stress on each syllable, have troubles with schwa, insert vowels, have difficulty in understanding the link between stress placement and meaning, and have low flat pitch (males). Italian speakers of English have troubles with the stress-timed nature of English, since all syllables in Italian have full vowels. French speakers of English elongate the last syllable in a phrase and drop pitch on it; they also do not use reduced vowels. Spanish speakers of English have narrow pitch range, use word order for contrastive and sentence stress, and have difficulty in distinguishing between stressed-unstressed and unstressed-stressed words like ”PROject” and “proJECT”. Chinese (Mandarin and Cantonese) speakers of English have problems with all aspects of prosody (Komissarchiket al, 2000).Let us deal with these prosodic aspects and explain with illustrations how they are used and should be used to reduce the foreign accent and reach the intelligibility.

3.10.3.2. Intonation

Intonation is a variation of spoken pitch that is not used to distinguish words; instead it is used for a range of functions such as indicating the attitudes and emotions of the speaker, showing the difference between *statement* and *question*, and between different types of question, focusing attention on important elements of the spoken message and also helping to regulate conversational interaction. It contrasts with *tone*, in which pitch variation in some languages *does* distinguish words, either lexically or grammatically. (The term *tone* is used by some British writers in their descriptions of intonation, but this is to refer to the pitch movement found on the *nucleus* or tonic syllable in an intonation unit. The intonation of the language thus refers to the rise and fall of the pitch of the voice when we speak. Therefore, intonation is very important for intelligibility, because as mentioned earlier, it is used to express intentions. This is because an inappropriate intonation pattern or a mispronounced word can lead to misunderstandings.

The acoustic manifestation of intonation is fundamental frequency **F0**, which is perceived by listeners as pitch**.**

3.10.3.2.1. Intonation Patterns

Intonation is expressed using different patterns used for different types of sentences. Yet, there are three basic patterns in English, which are: normal, rising and falling. Intonation patterns and how they are expressed are illustrated below:

i) Falling tone **yes** ↘

ii) Rising tone **yes** ↗

iii)Fall-rise tone ˅ **yes**

iv)Rise-fall tone ˄ **yes**

v) Level tone — **yes**

- **How to Express a Certain Attitude**

a. Different voice qualities for different attitudes

b. One may use different facial expressions, gestures and body movements.

3.10.3.2.1.1. Falling Intonation

- This is the tone that is usually regarded as more or less neutral. If someone is asked a question and the reply is *Yes* or *No*, it will be understood that the question is now answered and that there is nothing more to be said. Thus the falling tone gives an impression of finality.
- At the end of the sentence tone is falling. It is the most common type of intonation in English.
- It is used for asking and giving information in normal, quiet and certain style.
- It sounds more categorical, confident and convincing than rising intonation.

3.10.3.2.1.1.1. Use of Falling Intonation

Falling intonation is used on the last stressed syllable of the sentence in:

- **Statements** (declarative sentences): We live in \MOScow. .
- **Special Questions**: Where do you \Live?
- **Commands** (imperative sentences): \STOP it! Sit \DOwn.
- **Exclamatory Sentences**: What a wonderful SUR\prize!
- **The last part of Alternative Questions** (after “or”): Do you want /TEA or \COFfee?
- **Tag Questions** (when the speaker is *sure* that the answer will be *yes*): You \LIVE here, \DON’T you? (The speaker is sure and expects the answer to be *yes*.)

3.10.3.2.1.1.2. Examples of Falling Intonation

a) |This is the end of the *news*|

b) |I am absolutely *certain*|

c) |Stop *playing*|

d) |I have finished *working*|

e) |Stop *talking*|

3.10.3.2.2. Rising Intonation

Rising intonation in English is a very complicated phenomenon. It can express a number of various emotions, such as non-finality, surprise, doubt, interest, politeness, lack of confidence, etc. It is used in:

- **General Questions**: Was she glad to /SEE him?
- **Dependent or Introductory Parts of Sentences**: If he /CALLS, ask him to /COME.

- **The First Part of Alternative Questions** (before 'or'): Would you like an /APple or a /PEAR?
- **Direct Address**: /SIR, you dropped your note/BOOK
- **Enumerating Items in a List**: She bought/bread,/cheese and to/MATOES.
- **Tag Questions** (when the speaker is *not sure* that the answer will be *yes* or wants your opinion): It's a beautiful \TOWN, /ISN'T it? (The speaker thinks that the town is beautiful but asks for your opinion and confirmation.)

This tone conveys an impression that something more is to follow:

a. |I phoned **them**| (but they were not home)

b. |You must write it **again**| (and this time gets right)

c. |I have to leave **now**| (because I am getting late)

It can be used while making general questions:

d. |Can you **help**?|

e. |Is it **over**|

f. |Can I go **now?**|

This tone can be used while listing things**:**

g. |**Red**, **brown**, **yellow**, **green**, and **blue**|

h. |**Peter**, **Jack**, **Roger**, and **Sam**|

i. |**Oranges**, **bananas**, **mangoes** and **apples**|

It may be used while encouraging someone:

j. |It won't **hurt**|

k. |You will get it **right**|

l. |There is always next **time**|

3.10.3.2.3. Fall-Rise Tone

The fall-rise tone is a combination of a fall from high or mid to low followed by a low rise, that is to say, a rise from low to mid. It is considered to be one of the tones that make English speech sound too theatrical. A fall-rise tone may be used for several purposes; on the one hand, it may indicate that the speaker is not telling everything, but a part of the message is only implied, thus, the listener has to grasp it from the context.

This tone shows limited agreement, response with reservation, uncertainty, or doubt:

a. |You may be **right**|
b. |Its **possible**|
c. |If I am not **mistaken**|
d. |He may be **honest**|
e. |It can be **true**|

3.10.3.2.4. Rise-Fall Tone

According to Crystal (1997), this tone is used to convey strong emotional involvement feelings of approval, disapproval or surprise; depending on the face, the attitude might be delighted, challenging, or complacent.

a. |It’s **impossible**|
b. |You were **first**|
c. |All of **them**|

d. |He is **honest**|

e. |Its **true**|

3.10.3.2.5. Level Tone

- This tone is neutral and uninterested.
- It is used in English language in a restricted context. It almost always conveys a feeling of routine, sarcastic, uninteresting or boring. For example: a teacher calling names of the pupils from a register.

3.10.3.2.6. Structure of Intonation

The foundation of intonational structure includes the units of Intonational Phrase (IP) and intermediate phrase (IP). Every IP consists of at least one IP and a boundary tone (%), and every IP consists of at least one pitch accent (*) (where the intonation peak occurs). Accents in a phrase are determined by looking at the pitch track of the speech file, to determine where the pitch levels are high or low (Gussenhoven, 2004).

3.10.3.3. Boundary Tones

Intonation contours are produced by an interpolation of the pitch level between tonal accents, with boundary tones at the end of an intonational phrase that determine rising or falling interpolation. In figure below, adapted from Gussenhoven's sentence "I've *invited John Peck as well*" (1999, p.286), we can see that it ends with the symbol **H%**, which represents a high boundary tone or a rising target at the end of the phrase. Though tonal accents vary, they correspond to the lexical accents present in words. In this way, the structure of intonation is crucially based on both lexical stress (often termed *accent* as in French), and boundary tones. So both stress and tone play a key role in the formation of intonation contours. Stress provides the high and low points of the contour, and intonation itself is realized as the interpolation of pitch between the accents and the boundary tones.

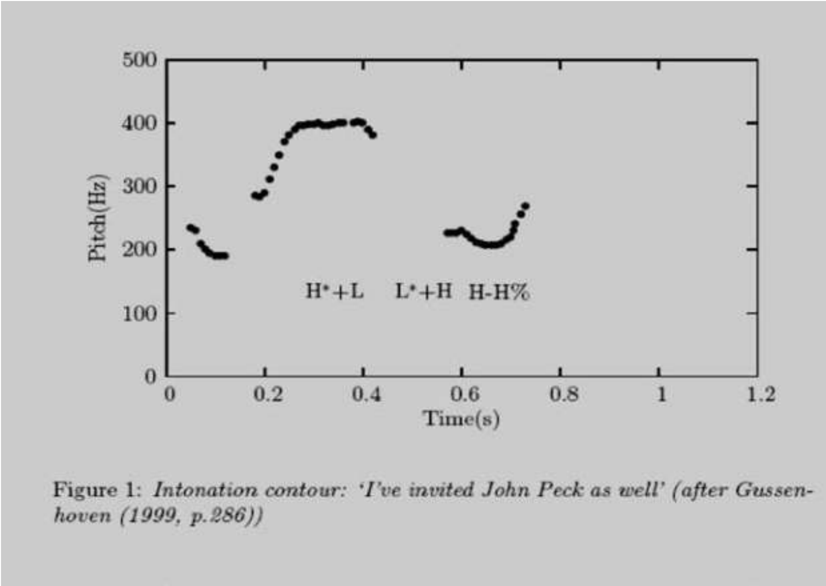


Figure3.1: Intonation contour of "*I've invited John Peck as well*" adapted from Gussenhoven(1999, p. 286)

3.10.3.4. Functions of Intonation

If every syllable is said on the same level pitch, no pauses, no changes in speed and loudness and without intonation, speech will be as though produced by a *mechanical speech device*.

Intonation makes it easier for the listener to understand the meaning a speaker is trying to convey. According to Crystal (1997), intonation has six important communicative functions: emotional, grammatical, information structure, textual, psychological and indexical. The following are the different functions of intonation, the way Crystal described them:

3.10.3.4.1. Emotional: The most obvious function is to express a wide range of attitudinal meanings, for instance: excitement, boredom, surprise, friendliness, reserve, etc.

3.10.3.4.2. Grammatical: Intonation plays an important role in the marking of grammatical contrasts. The identification of such major units as clauses and sentences depends on the way pitch contours break up an utterance; and several specific contrasts, such as question and statement, or positive and negative, may rely on intonation.

3.10.3.4.3. Information Structure: Intonation conveys a great deal about what is new and what is already known in the meaning of an utterance, what is referred to as the informal structure of the utterance. If someone says I saw a BLUE car, with maximum intonational prominence on blue, this pronunciation presupposes that someone has previously mentioned the colour, whereas if he emphasizes on I, it presupposes a previous question about which person is involved.

3.10.3.4.4. Textual: Intonation is not only used to mark the structure of the sentences, it is also an important element in the construction of larger stretches of discourse. A well-illustrated example is in the paragraphs of information, which are given a distinctive melodic shape in radio news-reading. As the news-reader moves from one item to the next, the pitch level jumps up, then gradually descends, until by the end of the item, the voice reaches a relatively low level.

3.10.3.4.5. Psychological: Intonation can help to organize language into units that are more easily perceived and memorized. Learning a long sequence of numbers, for instance, would be easier if the sequence is divided into rhythmical chunks.

3.10.3.4.6. Indexical: An indexical function refers to language features, which also have a significant function as markers of personal identity. In particular, they help to identify people, their occupations and which social groups they belong to (such as teachers, lawyers, street vendors, etc. (Crystal, p.249).

Therefore, intonation can be communicated in different ways to convey different meanings. We will review some of these functions mainly the attitudinal, the accentual, the grammatical and the discourse one, with some illustrations.

3.10.3.4.7.Attitudinal Function

- Intonation enables the speaker to express emotions and attitudes which adds a special meaning to spoken language as a difference from its written counterpart.
- It allows us to express finality, confidence, interest, surprise, doubt, joy, pain, irony, anger, boredom, gratefulness and so on.

3.10.3.4.8.Accentual Functions

The term accentual refers to *accent*. Some writers attach the word accent to *stress* like in the *French language*. When it is said that intonation has accentual function, it implies that the placement of stress is somewhat determined by intonation.

The most common position for the placement of tonic syllable is the last lexical word (nouns, adjectives, words, adverbs) and not the functional words. For contrastive purpose, however any word may become the bearer of *tonic syllable.*

3.10.3.5.Tonic Syllable: The tonic syllable is an obligatory component of tone unit. The centre of the unit, around which everything else is constructed, is the *tonic syllable*, or *nucleus*.

3.10.3.5.1. Location of the Tonic Syllable

- It is of great linguistic importance. The most common position for the placement of tonic syllable is the *last lexical word* (nouns, adjectives, verbs, adverbs) and not the functional words.
- But, for contrastive purposes any word can become the bearer of the tonic syllable.
- Thus, the placement of the tonic syllable represents the focus of the information.

Examples

a.|She was wearing a red *dress*| (Normal placement)

b.|She was not wearing a **red** dress| She was wearing a *green* dress| (Contrastive purpose placement)

a. |I want to know where he is *travelling* to|
b. |I don't want to know where he is travelling *to*| I want to know where he is travelling *from*|(Roach, p. 173)

3.10.3.5.2. Functions of Accentual Intonation

Similarly for the purpose of emphasis the tonic stress can be placed in other positions.

In the following examples, a is emphatic and b is non-emphatic

a. |The movie was very *boring*|

b. |The movie was *very* boring|

a. |You shouldn't talk so *loudly*|.|You *shouldn't* talk so loudly| (Roach, p.173)

3.10.3.5.3.Grammatical Functions

The listener is better able to recognize the grammar and syntax structure of what is being said by using the information contained in the intonation. For example such things as:

a. The placement of boundaries between phrases, clauses and sentences

b. The difference between questions and statements

Roach (1991) stated that grammatical intonation is used in those sentences which when written are ambiguous, and whose ambiguities can only be removed by using differences of intonation. In the following example the difference caused by the placement of tone-unit boundaries causes two different interpretations of sentence.

a. |Those who sold *quickly*| made a *profit*| (a profit was made by those who sold quickly)

b.|Those who *sold*| quickly made a *profit*| (a profit was quickly made by those who sold) (P.174)

Intonation is used to distinguish sentence types. Compare:

- **She was˥ not there. ‖ (a declarative)**
- **She was˥ not /there? ‖ (a Yes/No interrogative)**
- **Shut the door. ‖ (an imperative)**
- **| Shut the /door. ‖ (a request)**
- **Hasn't she been /clever? Question**
- **Hasn't she been \clever! Exclamation**

Intonation is, then, based on several key components, such as pitch, sentence stress and rhythm. *Pitch* is the degree of height of our voice in speech. Normal speaking pitch is at mid-level. Intonation is formed by certain pitch changes, characteristic of a given language; for example, falling intonation is formed by pitch changes from high to low, and rising intonation is formed by pitch changes from low to high.

Sentence stress makes the utterance understandable to the listener by making the important words in the sentence stressed, clear and higher in pitch and by shortening and obscuring the unstressed words. It provides rhythm in connected speech. All words have their own stress in isolation, but when they are connected into a sentence, important changes take place: content words are stressed and function words are not; thought groups (i.e. logically connected groups of words) are singled out by pauses and intonation; the stressed syllables occur at regular intervals and are usually higher in pitch than the unstressed syllables; the unstressed syllables are blended into a stream of sounds between the stressed syllables; emphatic stress may be used in the sentence to single out the most important word; the last stressed word in the sentence gets the strongest stress with the help of falling or rising intonation. Developing the ability to hear, understand and reproduce sentence stress is the main prerequisite to mastering English intonation.

3.10.3.6. Intonation Analysis

The primary method of intonation analysis involves the use of the *tones* and *Breaks Indices* (*ToBI*) system of transcription, developed by Beckman and Elam (1997). The ToBI transcription system, based on the work of Pierrehumbert (1980), provides a model for transcribing the intonation contour by breaking up the intonational phrase into discrete units of High and Low accents, which are hierarchically structured. This system utilizes both the physical speech signal analyzed by a computer, and the transcribers' perception of the contours. The physical speech signal for intonation is the pitch track or the measure of the fundamental frequency over the course of the utterance.

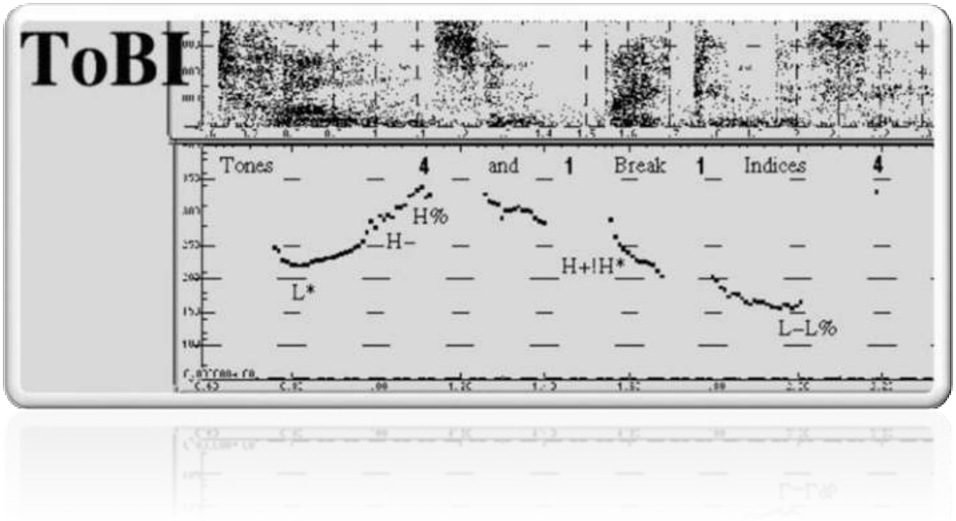


Figure3.2: Transcription with ToBI's system (adapted from http://www.ling.ohio-state.edu//~tobi

3.10.3.6.1. ToBI's System

In ToBI system, target tones **H*** and **L*** (called **H** star and **L** star) are typically written on a line (called tier) above the segmental symbols that represent stressed syllables. A high tone, **H***, can be preceded by a closely attached low pitch, written **L+H*,** so that the listener hears a sharply rising pitch. Similarly, **L*** can be followed by a closely attached high pitch**,** **L*+H,** so that the listener hears a scoop upward in pitch after the low pitch at the beginning of stressed syllable. Sometimes, a stressed syllable can be high but nevertheless can contain a small step-down of the pitch. This is known as high plus downstepped high, and is written **H+!H*,** with the exclamation mark indicating the small downstep in pitch. In special circumstances, a downstepped high syllable **!H*** can itself be a pitch accent.

According to Ladefoged & Johnson (2006), there are six possibilities that can be regarded as the possible pitch accents that occur in English.

The last pitch accent in a phrase is called the *nuclear pitch accent.* The ***ToBI*** system allows the phrase to be marked by an additional tone after the nuclear pitch accent. This tone, called the *phrase accent*, is written **H–** (**H** minus) or **L–** (**L** minus). Finally, there is a boundary tone, which is marked **H%** or **L%,** depending on whether the phrase ends on a rising or a falling pitch.

3.10.4.Tone

Tone can be defined as the phonologically contrastive use of pitch in the domain of the segment or syllable (Yip, 2002), and when this is present it is termed a tonal language. A tonal language can have a variety of different tonal systems; some simply have a distinction between High tone and Non High (Low) tone, while others can have many more levels. Some languages also have contour tones, in which the pitch can rise or fall over the course of the segment or syllable. An example of a tonal language with both level and contour tones is Cantonese, which is outlined in the following:

[yau]

a. high level ‘worry’

b. high rising ‘paint’

c. mid level ‘thin’

d. low level ‘again’

e. very low level 'oil'

f. low rising 'have' (Yip, 2002, p.2)

Then, *tone* refers to significant (i.e. meaningful, contrastive, phonemic) contrasts between words signalled by pitch differences. Tone may be *lexical*, as in Mandarin Chinese:

Tone Number	Description		IPA Transcription	Meaning
1	high level	1.	[má]	mother
2	high rising	2.	[māá]	hemp
3	low (falling+)rising	3.	[mǎ] or [màá]	horse
4	high fall	4.	[mâ] or [máà]	scold
no tone/neutral tone	(depends on preceding syllable)	5.	[ma]	(question marker)

Table3.2: Pitch levels in Mandarin Chinese

Or *grammatical* tone, as in many African languages, e.g. Edo:

Tense	**Monosyllabic Verbs**	**Disyllabic Verbs**
Timeless	[ì mà] `I show'	[ì hrùl<u>è</u>] `I run'
Continuous	[í mà] `I am showing'	[í hrùl<u>é</u>] `I am running'
Past	[ì má] `I showed'	[ì hrúl<u>è</u>] `I ran'

Table 3.3: Grammatical Pitch Levels in Edo

However, there may also be non-pitch aspects of tone. Lexical tones are often related to durational, phonatory and vowel quality distinctions as well as frequency distinctions. For example, Mandarin Chinese tone **3** (low rise) is long with creaky voice, Hunanese tone **2** has breathy or chesty voice. Tibetan tone 1 words have voiceless initial consonants whereas tone **2** words have voiced beginnings. Long vowels in tone **4** or **5** open syllables in Thai are checked by a final glottal stop.

Tone can be measured phonetically by looking at the pitch track of an utterance, which is a measure of the fundamental frequency of the speech sound. Though pitch is the primary cue for phonological tone, scholars have noted that high tones also have the correlates of high intensity and longer duration. Intensity, however, has been found to be less perceptible to hearers, and as was discussed above, there are anatomical reasons behind why pitch and intensity may often co-occur. Tonal languages also have a tendency to assign stress to higher toned syllables, as is discussed in De Lacy (2002), which also makes sense because higher tones are likely to have longer durations, and stress is often correlated with heavy syllables. Tone is the manifestation of pitch in the domain of the segment or syllable, and also plays an important role in intonation.

3.10.5.Pitch

Cruttenden (1986) describes pitch as the "*perceptual correlate of fundamental frequency*" (p. 1), which, in essence, is the continuous variation in the sounds we perceive as a result of the vibration of the vocal cords. As such, intonation can be described as the movements or variations in pitch to which we attach familiar labels describing levels (e.g. high / low) and tones (e.g. falling /rising), etc. (Ranalli, 2002).

The phonetic phenomenon of pitch, or fundamental frequency, is created by the vibration of the vocal cords during speech. Pitch and intensity are related notions, and higher pitch tends to be accompanied by higher intensity. This is because in order to raise pitch, a higher sub-glottal pressure must occur, which causes the vocal cords to vibrate more rapidly.

This higher sub-glottal pressure also causes an utterance to be louder (have a higher intensity level).

3.10.6. Stress

The term stress is used in the literature with a bewildering variety of definitions. Jones (1957) describes stress as *the degree of force with which a sound or syllable is uttered* (p.245). Abercrombie (1967) talks in terms of *force of breath impulse* while Crystal (1969) uses *loudness* as the main indicator of stress. Others regard stress as being controlled by the relative durations of syllables. Lehiste and Peterson (1959) state that the perception of stress is produced by variations in intensity, **F0** phonetic quality and duration. Often the terms *stress* and *accent* are used analogously, which adds to the confusion.

According to Ladefoged & Johnson (2006), stress is a suprasegmental feature of utterances. It applies not to individual vowels and consonants, but to whole syllables. A stressed syllable is then pronounced with a greater amount of energy than an unstressed syllable and is more prominent in the flow of speech. Thus, stressed syllables are most easily identified in citation forms. In conversational speech, words can be unemphesized, and when this happen, some of the properties of stressed syllables may not be realized. In citation forms, a stressed syllable is usually produced by pushing more air out of the lungs in one syllable relative to others. A stressed syllable thus has greater respiratory energy than neighbouring unstressed syllables. It may also have an increase in laryngeal activity.

English and other Germanic languages make great use of differences in stress than do most of the languages of the world, having a kind of variable word stress, in which the placement of stress is not always predictable from the segmental structure of the word, for example, the case of word class pairs to distinguish between a verb and a noun *to insult* and *an insult*. In many other languages, the position of the stress is fixed in relation to the word. For instance, in Czech, the words nearly always have the stress on the first syllable, no matter the number of syllables in the word. While, in Polish and Swahili, the stress is usually on the penultimate syllable.

Variations in the use of stress cause different languages to have different rhythms, but stress is only one factor in causing rhythmic differences. In English, the rhythm of a sentence depends on several interacting factors, not just stress. Perhaps a better way to describe stress differences among languages would be to divide languages into those which have variable word stress (such as the previously mentioned examples of English and German), those which have fixed word stress (such as Czech, Polish, and Swahili), and those that have fixed phrase stress (such as French, which belongs to syllable –timed languages, in which syllables tend to recur at regular intervals of time).

3.10.6.1. Levels of Stress

We can have three levels of stress, which are as follows:

- **Primary Stress**: is considered to be the strongest type of stress in which the syllables are characterized to be longer, stronger, higher in pitch and different in quality than the neighbouring sounds. It is a small vertical line put high up the stressed syllable. (‘)
- **Secondary Stress**: is weaker than the primary stress but stronger than the unstressed. Syllables that are not reduced, but not the most prominent in the word are called *secondary stressed* syllables. A small vertical line put below the stressed syllable(ˌ)
- **Tertiary:** generally syllables which contain a weak vowel schwa are never stressed, and said to be weak syllables. Unstressed syllables undergo vowel reduction.

3.10.6.2. Types of Stress

There are two types of stress: Sentence stress and word stress.

3.10.6.2.1. Sentence Stress

Sentence stress needs not be referred to as the particular anxiety an EFL student experiences when attempting to pronounce a particularly wordy sentence in English. Sentence stress is actually the music of English; the thing that gives the language is particular *beat* or *rhythm*. In general, in any given English utterance there will be particular words that carry more *weight* or *volume* than others. From a speaking perspective, sentence stress will affect the degree to which an EFL student sounds *natural*. In terms of listening, it affects how well a student can understand the utterances they hear.

3.10.6.2.2.Word Stress

Whereas sentence stress refers to the process whereby particular words are stressed within an overall sentence, word stress refers to the process whereby particular syllables or *parts of words* are stressed within an overall word. In general, sentence stress is more of a consideration for overall fluency. Words stress tends to have more of a phonological and morphemic importance.

3.10.6.3. Mechanism and Function of Stress

In any given sentence in English, there will be words that carry stress and others that do not. This is not a random pattern. Stressed words carry the meaning or the sense behind the sentence, and for this reason they are called *content or lexical words*; they carry the content *meaning* of the sentence. Unstressed words tend to be smaller words that have more of a grammatical significance- they help the sentence *function* syntactically and for this reason they are called *function words* and sometimes, they are referred to as *structure words*.

3.10.6.4.Rules of Sentence Stress in English

The basic rules of sentence stress are:

i. Content words are stressed
ii. Structure words are unstressed
iii. The time between stressed words is always the same.

3.10.6.5. Sentence Stress and Intonation

Sentence stress is the key component of English intonation. Intonation organizes words into sentences, distinguishes between different types of sentences and adds emotional colouring to utterances. Let us sum up the functions of sentence stress.

a. Sentence stress organizes separate words into sentences by making content words stressed and function words unstressed.
b. Sentence stress makes the utterance understandable to the listener by making the important words in the sentence stressed, clear and higher in pitch and by shortening and obscuring the unstressed words.
c. Sentence stress organizes the words in the sentence rhythmically, making the stressed syllables occur at regular intervals and jamming together the unstressed syllables between the stressed syllables.
d. Sentence stress organizes the words in the sentence into logically connected thought groups by joining the unstressed syllables to the main stressed syllable in the group and marking the end of the thought group with a slight pause, if necessary.

e. If necessary, sentence stress singles out the most important word in the sentence by giving it emphatic stress.

f. Sentence stress marks the end of the sentence by giving the strongest stress to the last stressed syllable with the help of falling or rising intonation.

It is not possible, of course, to learn sentence stress and rhythm just by talking about them. Listening and repeating should become the most important part of our work on pronunciation. That is why; we should always choose the textbooks that come with corresponding listening materials. When we practice repeating sentences after the recorded speaker, we always mark sentence stress and reduced unstressed words.

3.11. Connected Speech

When people communicate, they use language in either its written or spoken form, in a general sense, we bear in mind the use of other forms of language, for instance sign-language Braille, Morse, semaphore, etc.

Speech is then a continuous stream of sounds, without clear –cut borderlines between them, and the different aspects of connected speech help to explain why written English is so different from spoken English. So, what is it that native speakers do when stringing words together that cause so many problems for students?

Speed is a factor in fluency. When we speak quickly, we speak in groups of words which are continuous and may not have pauses between them. This causes changes to the *shape* of words. Unstressed words always sound different when used in a sentence as opposed to being said in isolation. It is an expression used to refer to spoken language when analyzed as a continuous sequence, as in utterances and conversations spoken at natural speed in everyday situations of life.

The most common features of connected speech are the weak forms of grammatical and some lexical words (*and, to, of, have, was, were*) and contractions, (*can't, won't, didn't, I'll, he'd, they've, should've*). However, we often ignore other features which preserve rhythm and make the language sound natural.

3.11.1. Aspects of Connected Speech

Another feature of pronunciation in which sounds change when they come in contact with others is connected speech. Words in every utterance are a continuous, changing, pattern of sound quality with associated (prosodic) features of quantity, accent, and pitch. The word is, like the phoneme, an abstraction from this continuum and must be expected to be realized in phonetically different ways according to the context. If the word is admitted as an abstracted linguistic unit, it is important to note the differences which may exist between its concrete realization when said (often artificially) in isolation and those which it haswhen in connected speech. In the latter case, it is subject to the pressures of its sound environment or of the accentual or rhythmic group of which it forms part. The variations involved may affect the word as a whole, e.g. weak forms in an unaccented situation or word accentual patterns within the larger rhythmic pattern of the complete utterance, or may affect more particularly the sounds used at word boundaries, such changes involving a consideration of the features of morpheme and word junctures, in particular, *assimilations*, *elisions*, and *liaisons*, in addition to other aspects that we will explain eventually.

3.11.1.1. Assimilation

The term assimilation describes how sounds modify each other when they meet, usually across word boundaries, but within words too. In other words, assimilation is the influence exercised by one sound upon the articulation of another sound, or it is the process by which a sound changes to more closely another sound.

Assimilation is something that varies in extent according to speaking rate and style; it is more likely to be found in rapid, casual speech and less likely in slow, careful one. Sometimes the differences caused by assimilation are very noticeable, and sometimes it is very slight.

According to Roach (1983), assimilation is indicated in three kinds which are:

- Regressive Assimilation: when the first sound of the following word influences the last sound of the previous word.
- Progressive Assimilation: when the last sound in the first word influences the first sound in the following word.
- Coalescent Assimilation: when there is a fusion of sounds in connected speech and hence there is a mutual influence.

3.11.1.1.1. Practicing Assimilation

This makes articulation easier. But notice that the change from one consonant sound to another should not interfere seriously with comprehension because the resulting sounds are quite similar to the original ones.

The alveolar consonants /n/ /t/ /d/ /s/ and /z/ can change to become more like the following sound. It is a question of making things easier for the speaker. For instance, if you are going to close your lips for /p/, then it is easier to close them for the preceding nasal /n/, so /n/ assimilates into /m/. Examples are summed up within the following tables:

Assimilation of		Becomes bilabial	For instance:
/n/	Before bilabial stops /p/ or /b/	/m/	In bed
/t/		/p/	Fat boy
/d/		/b/	Good boy

Table3.4: Assimilation before Bilabial Stops

Assimilation of		Becomes velar	For instance:
/n/	Before velar stops /k/ or /g/	/ŋ /	Ten girls
/t/		/k/	That girl
/d/		/g/	Good girl

Table3.5: Assimilation before Velar Stops

Assimilation of	Before palato-lveolar	Becomes palato-lveolar	Examples
/s/	/ʃ/ /tʃ/ /dʒ/	/ ʃ/	This shop, this chapter
/z/		/ʒ/	Cheese shop, those churches.

Table3.6: Assimilation before Palato- Alveolar

Coalescent Assimilation	**Examples**
If a word ends in /t/and the following word begins with/j/both sounds may coalesce to become **/tʃ /**	Can't you
/d/+/j/=/ dʒ/	Could you
/s/+/j/=/ ʃ /	Is this yours?
/z/+/j/=/ʒ/	He's your brother

Table 3.7: Coalescent Assimilation

The most common form involves the movement of place of articulation of the alveolar stops /t/, /d/ and /n/ to a position closer to that of the following sound. For instance, in the phrase *ten cars*, the /n/ will usually be articulated in a velar position, /ˈteŋ ˈkɑ:z/ so that the organs of speech are ready to produce the following velar sound /k/. Similarly, in *ten boys* the /n/ will be produced in a bilabial position, /ˈtem ˈbɔɪz/ to prepare for the articulation of the bilabial /b/.

Phoneme	Realised as	Example
/n/	**/ŋ/**	bank /bæŋk/
/d/	**/g/**	good girl /gʊg ˈgɜ:l/
/t/	**/k/**	that kid /ðæk ˈkɪd/

Phoneme	Realised as	Example
/n/	**/m/**	ten men /tem 'men/
/d/	**/b/**	bad boys /bæb ˈbɔɪz/
/t/	**/p/**	Hot mushrooms /ˈhɒp ˈmʌʃru:mz/

3.11.1.2. Rhythm

It has often been claimed that English speech is rhythmical, and that the rhythm is detectable in the regular occurrence of stressed syllables; of course, it is not suggested that the timing is as regular as a clock—the regularity of occurrence is only relative. The theory that English has *stress timed rhythm* implies that stressed syllables will tend to occur at relatively regular intervals whether they are separated by unstressed syllables or not; this would not be the case in *mechanical speech*.

3.11.1.3. Elision

The term *elision* describes the disappearance or the omission of a sound in a word or phrase. Native speakers unconsciously practice elision to make their pronunciation easier; that is why many of the English words have a sound or more dropped. For example, in the utterance *he leaves next week* speakers would generally elide (leave out) the /t/ in next saying

/nekswi:k/. Again here, the reason is an economy of effort, and in some instances, the difficulty of putting certain consonant sounds together while maintaining a regular speech rhythm and speed.

3.11.1.3.1. Rules of Elision

The most common consonants to find involved in elision are /t/ and /d/.

Elision of /t/ and /d/

They are elided when they are at the end of a word (in the last syllable) and between two other consonants.

Elision of /t/

/fɜst' θri/	[ˈfɜˈsθri]	First three
/ˈlɑstˈjɪə/	[ˈlɑsˈjɪə]	Last year
/məʊst' risənt/	[məʊsˈrisənt]	Most recent

b. Elision of /d/

/d/ elides even more readily than /t/ and in more environments:

/ˈwɜld' wɑɪld' laɪf' fʌnd/	[ˈwɜl' wɑɪ' laɪ' fʌnd]	World Wild Life Fund
/hɜld ˈtwentɪ/	[hɜl ˈtwentɪ]	Hurled twenty (yards)
/ ˈræpɪdlɪ/	[ˈræpɪlɪ]	Rapidly

3.11.1.3.2. Elision of Identical Sounds

They are elided when a word ending in a consonant sound is followed by another word starting with that sound.

E.g. *lamp post* *six students* *lettuce salad*

3.11.1.3.3. Elision of Initial Sounds in Pronouns

-Weak Pronouns

e.g. I saw him half an hour /aisɔ: imha:f@n hau@/

The most important occurrences of this phenomenon is when the alveolar consonants /t/ and /d/ are *sandwiched* between two consonants, e.g.

The ne**xt d**ay….	/ðə ˈneks ˈdeɪ/
The la**st c**ar…	/ðə ˈlɑ:s ˈkɑ:/
Ho**ld th**e dog!	/ˈhəʊlðə ˈdɒg/
Se**nd F**rank a card.	/sen ˈfræŋk ə ˈkɑ:d/

This can also take place when the affricates /tʃ/ and /dʒ/ are preceded by a consonant, **e.g.**

Lunchtime	/ˈlʌntʃtaɪm/	become	/ˈlʌnʃtaɪm/
strange days	/ˈstreɪndʒˈdeɪz/		/ˈstreɪnʒˈdeɪz/

The phoneme /t/ is a fundamental part of the negative particle *not*, the possibility of it being elided makes the foreign students life more difficult. Consider the negative of *can* – if followed by a consonant the /t/ may easily disappear and the only difference between the positive and the negative is a different, longer vowel sound in the second:

I can speak….	/aɪkən ˈspi:k/
I can't speak…	/aɪ ˈkɑ:n(t) ˈspi:k/

3.11.1.4. Linking

It is adding or joining sounds between words. We tend to link final consonants and initial vowels across word boundaries. In what follows are some examples:

The phoneme **/r/** cannot occur in syllable-final position in RP, but when a word's spelling *suggests* a final **/r/**, and a word beginning with a vowel follows, the usual pronunciation for RP speakers is to pronounce with letter **r**. For example:

'Here' /hɪə/*but* 'here are' /hɪər ə/

'Four' /fɔ:/ *but* 'four eggs' /fɔ:regz/

Many RP speakers use **r** in a similar way to link words ending with a vowel even when there is no *justification* from spelling, as in:

'Formula A' / fɔ:mjələreɪ/

'Australia all out' /ɒstrelɪərɔ:laʊt/

'Media event' /mi:dɪəɪvent/

This has been called *intrusive***/r/**; some English speakers and teachers still regard this as incorrect or sub-standard pronunciation, but it is undoubtedly widespread. *Linking* and *intrusive **r*** are special cases of *juncture*; this name refers to the relationship between one sound and the sounds that immediately precede and follow it, and has been given some importance in phonological theory. If we take the words 'my turn' /mɑɪtɜ:n/, the relationship between **/m/** and **/ai**/, between **/t/** and **/ɜ**:/ and between **/ɜ**:/ and **/n/** is said to be one of *close juncture*. **/m/** is preceded by silence and **/n/** is followed by silence, and so **/m/** and **/n/** are said to be in a position of external *open juncture*. The problem lies in deciding what the relationship is between **/a**ɪ/ and **/t/**; since we do not usually pause between words, there is no silence (or *external open juncture*) to indicate word division and to justify the space left in the transcription. But if English speakers hear /m**a**ɪtɜ:n/ they can usually recognize this as *my turn* and not *might earn*. This is where the problem of internal *open internal juncture* (usually called *juncture* for short) becomes apparent. What is it that makes perceptible the difference between /m**a**ɪtɜ:n/ and /m**a**ɪtɜ:n/? The answer is that in the case the /**t/** is aspirated (initial in *turn*), and in the other case it is not (being final in *might*). In addition to this, the**/ai**/ diphthong is shorter in *might*, but we will ignore this for the sake of a simple argument. If a difference in meaning is caused by the difference between aspirated and unaspirated **/t/**, how can we avoid the conclusion that English has a phonemic contrast between aspirated and unaspirated **/t/**? The answer is, of course, that the position of a word boundary has some effect on the realization of the **/t/** phoneme; this is one of the many cases in which the occurrence of different allophones can be properly explained by making reference to units of grammar (something which was disapproved of by many phonologists).(Roach 1991 p. 128)

3.11.1.5. Intrusion and Linking

Intrusion is the insertion of an extra sound. If the words 'go' and 'up' are said together, there is a new **/w/** sound between the two words. When two vowel sounds meet, we tend to insert an extra sound which resembles either a **/ j /, /w/ or /r/,** to mark the transition sound between the two vowels. For example:

i. Intruding / r/

The media / ***r*** */are to blame.*

Law(r) and order.

ii. Intruding / j /

I / ***j*** */ agree.*

They / ***j*** */are here!*

iii. Intruding / w/

I want to/ ***w*** */eat.*

Please do/ ***w*** */it.*

You are /w/ will sound as/juːˈwɑː/ rather than /juːˈɑː/.

Word boundaries involving a consonant and a vowel are also linked, as we tend to drag final consonants to initial vowels or vice versa. For example:

- *Get on*. (**geton**)
- *Not at all.* (notatall)
- *It´s no joke*. (snow **joke**)

3.11.1.6. Liaison

It is the insertion of an extra phoneme in order to facilitate articulation. In other terms, it is adding or joining sounds between words. We tend to link final consonants and initial vowels across word boundaries.

3.11.1.6.1. Linking /r/

The /r/ sound is heard connecting two words when there is an R in the spelling and there is a following vowel sound.

Examples: *Peter and Tom; far away, more ice*

3.11.1.6.2. Intrusive /r/

In many words ending with the written consonant R, the final vowel sound is one of the following:

/@/ teacher, harbour, actor /ɔ:/ four, door . /a:/ car, far

No doubt, as a result of this, there is a tendency to insert an intrusive /r/ when a word ends in one of these vowels, even when no written **R** exists. Many people consider that intrusive /r/ is substandard, and certainly not to be imitated. For instance:

America *and Asia;*

Asia and America;

Law and order;

Vanilla ice cream;

I saw it

3.12. Weak Forms

Weak forms, in the phonetics of English, are a series of words which have one pronunciation (strong), when isolated, and another (weak), when not stressed within a phrase. The following is an example.

a car	/ˈeɪ ˈkɑ:/
I bought a car	/aɪ ˈbɔ:t ə kɑ:/

In connected speech, many words are pronounced in a weak form. Weak forms are usually distinguished by a change in vowel quality from a border position on the vowel quadrilateral to a central position. The vowel in a weak form is usually the schwa (ə). Weak forms are pronounced more quickly and at lower volume in comparison to the stressed syllables. They are also not central to changes in intonation.

There is a logical explanation behind the occurrence of weak forms: they are present in words which are necessary to construct a phrase; yet, at the same time, they do not communicate a large quantity of information; in other words, they are not content words.

For example in the following phrase:

I went to the hotel and booked a room for two nights for my father and his best friend.

- The most important words, those that are central to the message, can be emphasised: *I went to the hotel and booked a room for two nights for my father and his best friend.*

If we eliminate the words that are not emphasised, can we still understand the message?

went hotel booked room two nights father best friend.

Perhaps it is difficult to be certain but it is possible to predict what the missing words might be. The words which we emphasised would bear the stress, while many of those which we eliminated would become weak forms, simply because they are less important in the conveyance of the message. Look at the sentence in transcription:

/aɪ ˈwent tə ðə həʊ ˈtel ən ˈbʊkt ə ˈru:m fə ˈtu: ˈnaɪts fə maɪ ˈfɑ:ðər ən hɪz ˈbest ˈfrend/

We will notice that most of the unstressed words are pronounced with the sound **/ə/:** prepositions such as ***to*** and ***for*,** articles ***a, an*** and ***the***, and the conjunction ***and***. Auxiliary verbs frequently have weak forms.

Some of the most common examples of weak forms are:

Auxiliary Verbs: Do, are was were, would

Prepositions: To, for, from, into

Others: And, but, then, that (as a relative), you (as object pronoun)

3.12.1. Verbal Fillers in Speech

These are words, phrases, and sometimes just noises like *er* which do not contribute much, if anything, to the new information content of an utterance but perform several valuable functions in speech. The exemplification will be drawn from the speech of public speakers, informal conversations and interviews. Sometimes people do make phonetic errors which they correct on mid-sentence. So we hear them say:

I saw three [brig] + I mean big + dogs having a big fight out there. More other aspects which are often used by native speakers are the repetition and the fillers. The repetitive use of words in a sentence doesn't add anything to the meaning of the sentence, it rather gives the speaker time to work out what he/ she is going to say next. Fillers are commonly used too, when people are asked direct questions and expected to immediately respond. The most common filler used is *well* closely followed by a non-verbal noise like *mm* or *er*

Conclusion

In this chapter, we have attempted to pull some principles out of the fundamentals, and discussed in general terms how best to help students with pronunciation. We have been through various features, taking into account the basic ones from segmental to suprasegmental, to a speech which is never said in isolation but always connected, which leads to altering the pronunciation for the sake of making it easier and with less efforts. Most of these principles are based on familiar ideas about good pronunciation teaching practice, which we hope teachers undoubtedly use in other aspects of their teaching. The focus would be on:

- Having a suitable student-centred curriculum
- Helping learners become self-reliant when working with pronunciation software
- Giving opportunities to practice
- Knowing what is best for their students
- Knowing how to apply these familiar ideas to pronunciation which requires a fair bit of the mastering of not only phonetics and phonology but psycholinguistics as well.

Intonation is then, based on several key components, such as pitch, sentence stress and rhythm.

- Pitch is the degree of height of our voice in speech. Normal speaking pitch is at mid-level.
- Intonation is formed by certain pitch changes, characteristic of a given language, for example, falling intonation is formed by pitch changes from high to low, and rising intonation is formed by pitch changes from low to high.
- Sentence stress makes the utterance understandable to the listener by making the important words in the sentence stressed, clear and higher in pitch and by shortening and obscuring the unstressed words.
- Sentence stress provides rhythm in connected speech. All words have their own stress in isolation, but when they are connected into a sentence, important changes take place: content words are stressed and function words are not; thought groups (i.e. logically connected groups of words) are singled out by pauses and intonation; the stressed syllables

occur at regular intervals and are usually higher in pitch than the unstressed syllables; the unstressed syllables are blended into a stream of sounds between the stressed syllables; emphatic stress may be used in the sentence to single out the most important word; the last stressed word in the sentence gets the strongest stress with the help of falling or rising intonation. Developing the ability to hear, understand and reproduce sentence stress is the main prerequisite to mastering English intonation.

Chapter Four

Chapter Four: History of Pronunciation Teaching

Introduction

In his very comprehensive history of language teaching, Kelly (1969) calls pronunciation the *Cinderella* area of foreign language teaching. He shows that Western philologists and linguists had studied grammar and vocabulary much longer than pronunciation. For this reason, grammar and vocabulary had been much better understood by most language teachers than pronunciation, which began to be studied systematically shortly before the beginning of the twentieth century.

4.1. Movements, Approaches and Reforms of Pronunciation Teaching

The field of modern language pronunciation teaching has witnessed many movements, approaches and reforms. In what follows are the major ones.

4.1.1. Intuitive Imitative Approach

Before the late nineteenth century only this approach was used, occasionally supplemented by the teacher‘s or textbook writer's impressionistic (and often phonetically inaccurate) observations about sounds based on orthography (Kelly 1969). The *intuitive–imitative* approach depended on the learner's ability to listen to and imitate the rhythms and sounds of the target language without the intervention of any explicit information; it also presupposed the availability of good models to listen to, a possibility that was enhanced by the availability first of phonograph records, then of tape recorders and language labs in the mid-twentieth century, and more recently of audio and videocassettes and compact discs.

4.1.2. Analytic-linguistic approach

It used information and tools such as the phonetic alphabet, articulatory descriptions, charts of the vocal apparatus, contrastive information, and other aids to supplement listening, imitation and production. It explicitly teaches the learner the sounds and rhythms of the target language and focuses attention on them. This approach was developed to complement rather than to replace the intuitive-imitative approach, which was typically retained as the practice phase used in tandem with the phonetic information.

When we look at the various language-teaching methods that had some currency throughout the twentieth century, we must acknowledge that there were methods, such as the Grammar translation and the reading based approaches, in which the teaching of pronunciation was largely irrelevant. In such methods grammar or text comprehension was taught through the medium of the learner's native language, and oral communication in the target language, was not a primary instructional objective. In the following overview of methods we focus on those methods and approaches for which the teaching and learning of pronunciation is a genuine concern.

4.1.3. Direct Method and More Recent Naturalistic Approaches

In the Direct Method, which first gained popularity in the late 1800s and early 1900s, pronunciation was taught through intuition and imitation; students imitated a model-the teacher or a recording-and did their best to approximate the model through imitation and repetition. This instructional method was grounded on observations of children learning their first language and of children and adults learning foreign languages in non-instructional settings. Successors to this approach were the so called naturalistic methods, including comprehension methods that devoted a period of learning solely to listening before any speaking was allowed. Examples included Asher’s (1977) Total Physical Response and Krashen and Terrell’s (1983) Natural Approach. Proponents maintained that the initial focus on listening without pressure to speak gives the learners the opportunity to internalize that target sound system. When learners do speak later on, their pronunciation is supposedly quite good despite their never having received explicit pronunciation instruction.

4.1.4. Reform Movement

The first linguistic or analytic contribution to the teaching of pronunciation emerged in the 1890s as part of the Reform Movement in language teaching. This movement was influenced greatly by phoneticians such as Sweet, Viëtor, and Passy, who founded the International Phonetic Association in 1886 and developed the International Phonetic Alphabet (IPA). This alphabet resulted from the establishment of phonetics as a science dedicated to describing and analysing the sound systems of languages. A phonetic alphabet made it possible to accurately represent the sounds of any language because, for the first time, there was a consistent one to one relationship between a written symbol and the sound it represented.

The phoneticians involved in this international organization, many of whom had also had experience teaching foreign languages, did much to influence modern language teaching by specifically advocating the following notions and practice:

- The spoken form of a language is primary and should be taught first.
- The findings of phonetics should be applied to language teaching.
- Teachers must have solid training in phonetics.
- Learners should be given phonetic training to establish good speech habits.

- The basic principle of the IPA is : 1 sound = 1 symbol

4.1.5. The 1940s and 1950s

Many historians of language teaching (e.g. Howatt, 1984) believe that the Reform Movement played a role in the development of Audio-Lingualism in the United States and of the Oral Approach in Britain during the 1940s and 1950s. In both the Audio-lingual and Oral Approach classrooms, pronunciation was very important and was taught explicitly from the start. As in the Direct Method classroom, the teacher (or a recording) modelled a sound, a word, or an utterance and the students imitated or repeated it. However, the teacher also typically made use of information from phonetics such as a visual transcription system (modified IPA or some other system) or charts that demonstrated the articulation of sounds.

Furthermore, the teacher often used a technique derived from the notion of contrast in structural linguistics: the minimal pair drill –drills that used words that differ by a single sound in the same position. This technique, based on the concept of the phoneme as a minimally distinctive sound (Bloomfield, 1933), was used for both listening practice and guided oral production as follows:

Sample of Minimal Pair Teaching Materials

i. Word Drills

A /i:/	**B/I/**
Sheep	Ship
Green	Grin
Least	List
Meet	Mitt
Deed	Did

Figure 4.1. Word Drills

ii. Sentence Drills

a. Syntagmatic Drills (Contrast Within a Sentence)

- Don't **sit** in the **seat.**
- Did you at **least** get the **list**?

b. Paradigmatic Drills (Contrast across Two Sentences)

- Don't **slip** on the floor.
- Don't **sleep** on the floor**.**
- Is that a black **sheep**?
- Is that a black **ship**?

Using such minimal pairs, the teacher first had the students practice listening skills.[1]The teacher used to say two words (e.g., *sheep*, *sheep* or *sheep*, *ship*) and to ask the students to decide if they were the same or different. Alternatively, the teacher could read a word or words from either list A or list B and asked the student to identify which sound (A or B) was being produced.

Listening (Perception Tests)

i) **Same** or **Different**? (Sheep, sheep; ship, sheep) *Discrimination Test*

ii) **A** or **B**? (Ship; ship, sheep) *Identification Test*

Such listening discrimination practice was followed by guided oral production practice. Following a teacher model, students practiced lists A and B first in isolation (i.e., reading list A and then list B), then in contrast (i.e., reading across columns A and B).

Guided Oral Production

- Read down column **A**, then column **B** (sheep, green, etc.)
- Read across the columns (sheep, ship, etc.)

Finally, the teacher asked individual students to read the lists without a model.

4.1.6. The 1960s

In the 1960s, the *Cognitive Approach*, influenced by Transformational Generative Grammar (Chomsky 1957, 1965) and cognitive psychology (Neisser, 1967), viewed language

[1] This technique can be used to all minimal pair contrast involving vowel or consonant discrimination.

as a rule-governed behaviour rather than habit formation. It deemphasized pronunciation in favour of grammar and vocabulary because, its advocates argued:

(a) Native pronunciation was an unrealistic objective and could not be achieved (Scovel, 1969); and

(b) Time would be better spent on teaching more learnable items, such as grammatical structures and words.

4.1.7. The 1970s

As can be seen from the preceding overview, the language teaching profession changed positions many times with respect to the teaching of pronunciation. Various methods and approaches placed this skill either at the forefront of instruction, as was the case with the Reform Movement practices and the Audio-lingual/Oral Method, or in the back wings, as with the Direct Method and naturalistic comprehension-based approaches, which operated under the assumption that errors in pronunciation (and other errors, for that matter) were part of the natural acquisition process and would disappear as students gained in communicative proficiency. Other methods and approaches either ignored pronunciation (e.g., Grammar Translation, Reading-Based Approaches, and the Cognitive Approach) or taught pronunciation through imitation and repetition (Direct Method), or through imitation supported by analysis and linguistic information (Audiolingualism).

The methods that came to attention during 1970s, such as the Silent Way and Community Language Learning, continued to exhibit interesting differences in the way they dealt with pronunciation. This aspect of both methods is described in the following two sections (Stevick, 1980; Blair, 1991; and Larsen –Freeman, 1986).

4.1.8. The Silent Way

Like Audiolingualism, the Silent Way (Gattegno 1972, 1976) can be characterized by the attention paid to accuracy of production of both the sounds and structures of the target language from the initial stage of instruction. Not only are individual sounds stressed from the very first day of a Silent Way class, but learners' attention is focused on how words combine in phrases –on how blending, stress, and intonation all shape the production of an utterance. Proponents claim that this enables Silent Way learners to sharpen their inner criteria for accurate production. The difference between Audiolingualism and the Silent Way, however, is that in the Silent way learner's attention is focused on the sound system without having to learn a phonetic alphabet or a body of explicit linguistic information.

4.1.8.1. The Silent Way and Pronunciation Teaching

The teacher, according to the method's name, spoke as little as possible, indicating through gestures what students should do. This included an elaborate system in which teachers tapped out rhythmic patterns with a pointer, held up their fingers to indicate the number of syllables in a word or to indicate stressed elements, or modelled proper positioning of the articulators by pointing to their own lips, teeth, or jaw. The Silent Way teachers also used several indispensable tools of the trade such as a sound-colour chart, Fidel charts, word charts, and coloured rods3.

[3] All tools are available from Educational Solutions in New York, the late Caleb Gattegno's company.

The Sound-Colour Chart was created by Gattegno to bypass the ear (Gattegno 1985). It was a large rectangular wall chart which contained all the vowel and consonants of a target language in small coloured rectangular shapes. The upper half of the chart represented the vowels. The primary vowels were represented by one colour each and the diphthongs by two colours. The consonants were located in the bottom half of the chart and were separated from the vowels by a solid line. Colours for consonants were assigned randomly although there was consistency in colour from language to language when sounds overlapped.

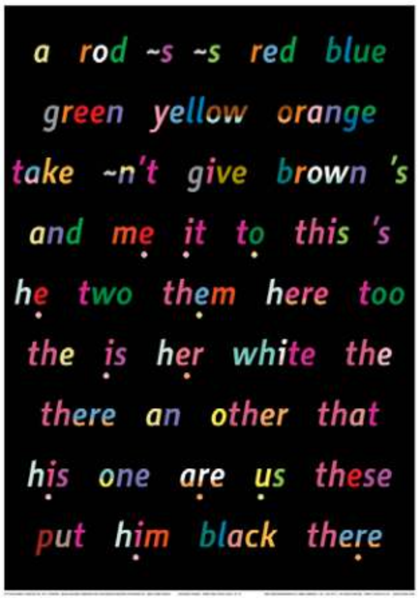


Figure4.2: Word Chart 1 for English as a Second Language (Adapted from Gattegno, 1972)

The set of Fidel *wall charts* contained all the possible spelling patterns for each sound in the language. Each letter or combination of letters was colour-coded: sounds that were pronounced alike were coloured alike. Because of the complex nature of English spelling eight charts in the set represented sound-spelling correspondence.

The large coloured word chart (1 to 12, *see Figure4.2.*) were similar in size to the sound-colour chart; they reflected and reinforced the system used in the sound–colour chart4.The wall charts contained common words of the target language, along with some words useful to the method (e.g., rod). These were grouped semantically in a way that allowed the leader (teacher or proficient student) to *silently dictate* or tap out phrases, which were then practiced orally and/or written down as a dictation. For example, the class could take several steps to progress from *Take a blue rod* to *Take a blue rod* and *a red rod*. *Give the blue one to him* and *the red to her*.

The final tool was a set of small coloured blocks of wood or plastic of varying lengths, with all red rods being the same size, all white the same size, and so on. The rods were used for many purposes, but when the focus was pronunciation, the rods could be used to build and visually demonstrate intonation patterns and to indicate the differing pronunciation of morphological endings (past tense, plural marker, etc.).

In one Silent Way lesson that we observed, the students were foreign–born professionals, advanced in English but with heavy accents. As an accent reduction exercise, the instructor was helping students to introduce themselves in a way that would be intelligible and acceptable to native English speakers. They first practiced giving their names (e.g., 'my name is 'Alexander Eliopoulos') by placing the coloured rods on the table in front of them in a configuration that approximated the stress, intonation, and blending of the phrase. This visual configuration was adjusted as students discovered ways in which they could produce a more intelligible form of the utterance. For example, the contraction of *name* into *name*'s was

[4] Since the Fidel wall chart is not reproduced in colour, some of the important clues to pronunciation are lost here.

achieved by moving the small white rod representing *is* directly next to the red rod representing *name.* The teacher remained very much in the background, and there was intense peer assistance both in monitoring the utterances and suggesting alternatives. Once a high level of intelligibility had been attained for the first phrase, students then practiced their professions (e.g. 'I'm a commercial real estate broker') in the same manner. Finally, they combined the two phrases and introduced themselves to their peers.5

The Silent Way is better understood if experienced rather than read about, since any description fails to capture the actual learner engagement. The method appears to have a special focus on teaching pronunciation, and many language educators agree that the principal of sound-colour correspondence, which the Silent Way invokes, provides learners with an "*inner resource to be used*" (Stevick,1980:46), which helps establish a true feel for the language, "*its diction, rhythm, and melody*" (Blair, 1991:32).

4.1.9. Community Language Learning

Rooted in the humanistic client-centred learning exemplified by Rogers (1951), Community Language Learning (CLL) is a method developed by Curran (1976) for teaching second and foreign languages. A typical lesson in a CLL classroom proceeded as follows: Students sat around a table with a tape recorder–a key tool of the method. The counsellor (i.e., the teacher) stood behind one of the students with hands on the student's shoulders. After speaking reassuringly, the counsellor asked the student to say an utterance in the target language. The utterance was then provided by the teacher in the target language, who took care to phrase it idiomatically. The counsellor provided the utterance (broken into chunks for

[5] Judith Weidman's class at the American Language Center of UCLA Extension

ease of repetition), the student repeated, and once the student was able to produce the whole utterance fluently, it was recorded on tape.

In the next phase of the lesson, the utterances were played back and students matched the new target language with the word-for-word translation provided by the counsellor. Next, the teacher asked if the students wished to further practice the pronunciation of any of the new utterances they had learned. If they did, the counsellor again stood behind the student who requested further practice and engaged in a technique known as human computer**.** The counsellor/computer could be turned on or off at will by the student, who could request the correct pronunciation of a given phrase or piece of a phrase from the computer. This provided the raw data for the student to mimic and repeat until he or she was satisfied with the pronunciation.

Several tools and techniques were critical to the treatment of pronunciation in CLL. First, the audiotape recorder not only captured what was said in the student-generated utterances but also provided a way for students to distance themselves from what was said so they could focus on how it was said and compared their pronunciation with that of the counsellor. Second, the human computer technique, which gave no overt correction of pronunciation, allowed the student to initiate pronunciation practice by selecting the item (s) to practice and deciding the amount of repetition needed. In this way, students were able to approximate the target pronunciation to the extent that they desired. Thus the teaching approach was intuitive and imitative as in the Direct Method, but its exact content and the extent to which practice took place were controlled by the learner/client rather than the teacher or textbook.

4.1.10. Pronunciation Teaching in the 80s

The Communicative Approach, which took hold in the 1980s and is still dominant in language teaching, held that since the primary purpose of language is communication, using language to communicate should be central in all classroom language instruction. This focus on language as communication brought renewed urgency to the teaching of pronunciation, since both empirical and anecdotal evidence indicated that there was a threshold level of pronunciation for non-native speakers of English; if they fell below this threshold level, they would have oral communication problems no matter how excellent and extensive their control of English Grammar and vocabulary could have been (this claim was better supported in Hinofotis and Bailey's research 1980).

Morley (1987:2) suggests that there are at least four groups of English language learners whose oral communication needs mandate a high level of intelligibility and therefore require special assistance with pronunciation.

i) Foreign teaching assistants and -sometimes foreign faculty- in colleges and universities in English-speaking countries.
ii) Foreign-born technical, business, and professional employees in business and industry in English speaking countries
iii) International business people and diplomats who need to use English as their working lingua franca.
iv) Refugees (adults and adolescents) in resettlement and vocational training programs wishing to relocate in English-speaking countries.

To Morley's four categories, we should add at least two more groups:

v) Teachers of English as a foreign language, who are not native speakers of English and who expect to serve as the major model and source of input in English for their students.

vi) People in non-English speaking countries working as tour guides, waiters, hotel personnel, customs agents and the like, who use English for dealing with visitors who do not speak their language.

The goal of teaching pronunciation to such learners is not to make them sound like native speakers of English. With the exception of a few highly gifted and motivated individuals, such a goal is unrealistic. A more modest and realistic goal is to enable learners to surpass the threshold level so that their pronunciation will not detract from their ability to communicate.

Having established that intelligible pronunciation is one of the necessary components of oral communication, the next issue is methodological: How can teachers improve the pronunciation of unintelligible speakers of English so that they become intelligible? This is a problem for Communicative Language Teaching, since proponents of this approach have not dealt adequately with the role of pronunciation in language teaching, nor have they developed an agreed-upon set of strategies for teaching pronunciation communicatively (Celce-Murcia, 1983 and Pica, 1984)

We can begin to answer the question of how to teach pronunciation as part of the Communicative Approach by reviewing the kinds of techniques and practice materials that have traditionally been used -and are still being used-to teach pronunciation. The following is a fairly comprehensive list:

a. **Listen and Imitate**: It is a technique used in the Direct Method in which students listen to a teacher-provided model and repeat or imitate it. This technique has been enhanced by the use of tape recorders, language labs, and video recorders.

b. **Phonetic Training**: It is the use of articulatory description, articulatory diagrams, and a phonetic alphabet (a technique from the Reform Movement, which may involve doing phonetic transcription as well as reading phonetically transcribed texts).

c. **Minimal Pair Drills**: It is a technique introduced during the Audio-Lingual era to help students distinguish between similar and problematic sounds in the target language through listening discrimination and spoken practice. Minimal pair drills typically begin with word-level drills and then move on to sentence –level drills (both paradigmatic and syntagmatic).

d. **Contextualized Minimal Pairs*:*** It is Bowen's (1972, 1975b) attempt to make minimal pair drills responsive to Cognitive criticisms of meaninglessness and lack of context. In the technique, the teacher establishes the setting (e.g., a blacksmith shoeing a horse) and presents key vocabulary; students are then trained to respond to a sentence stem with the appropriate meaningful response (**a** or **b**).

- **Sentence Stem**

  The blacksmith (**a**. hits/**b**. heats) the horseshoe.

- **Cued Student Response**

  a. with the hammer **/** **b**. in the fire

e. **Visual Aids**: It is enhancement of the teacher's description of how sounds are produced by audio-visual aids such as sound-colour charts. Fidel wall charts, rods, pictures, mirrors, props, realia, etc. These devices are also used to cue production of the target sounds.

f. **Tongue Twisters**: It is a technique from speech correction strategies for native speakers (e.g**.,** '*She sells seashells by the seashore*')

g. **Developmental Approximation Drills**: It is a technique suggested by first-language acquisition studies in which second language speakers are taught to retrace the steps that many English-speaking children follow as they acquire certain sounds in their first language. Thus, just as children learning English often acquire **/w/** before **/r/** or **/j/** before **/l/**, adults who have difficulty producing **/l/** or **/r/** can be encouraged to begin by pronouncing words with initial **/w**/or **/j/**, and then shift to **/r**/or **/l/** respectively:

/w/ →	**/r/**	**/j/** →	**/l/**
Wed	red	yet	let
Wag	rag	yes	less
Witch	rich	you	lou
Wipe	ripe	young	lung

Figure4.3: Developmental Approximation Drills

h. **Practice of Vowel Shifts and Stress Shifts Related by Affixation**: It is a technique based on the rules of Generative Phonology (Chomsky and Halle, 1968) used with intermediate or advanced learners. The teacher points out the rule-based nature of vowel and stress shifts in etymologically related words to raise awareness; sentences and short texts that contain both members of a pair may be provided as oral material:

- **Vowel Shift***:* mime (long /i:/) mimic (short/I/)
- **Sentence Context**: street *mimes* often *mimic* the gestures of passers-by.
- **Stress Shift**: PHOtograph/ phoTOGraphy

- **Sentence Context**: It can be said from these *photographs* that you are very good at *photography.*

i.**Reading Aloud/Recitation***:* These are passages or scripts for learners to practice and then read aloud, focusing on stress, timing, and intonation. This technique may or may not involvememorization of the text, and it usually occurs with genres that are intended to be spoken, such as speeches, poems, plays, and dialogues.

j. **Recording of Learners' Production***:* These are audio-and videotapes of rehearsed and spontaneous speeches, free conversations, and role-plays. Subsequent playback offers opportunities for feedback from teachers and peers as well as for teacher, peer, and self-evaluation.

With the exception of the last two techniques listed, we can see that the emphasis in pronunciation instruction has been largely on getting the sounds right at the word level dealing with words in isolation or with words in very controlled and contrived sentence level environments (Nilsen and Nilsen, 1973, as an instance).

Although the last two techniques allow for practice at the discourse level, the practice material is often fully scripted and sometimes highly, contrived. There is thus some doubt about whether such reading-aloud exercises can actually improve a learner's pronunciation in spontaneous conversation.

When the Communicative Approach to language teaching began to take over in the mid-to late 1970s,(Brumfit and Johnson1979; Widdowson1978) most of the afore mentioned techniques and materials for teaching pronunciation at the segmental level were rejected on theoretical and practical grounds as being incompatible with teaching language as communication. Influenced by the discourse-based approaches and materials being used to teach language communicatively, materials developers and teachers began to search for more appropriate and adequate ways to teach pronunciation. They decided that directing most of their energy to teaching suprasegmental features of language or what we refer to as prosody (i.e., rhythm, stress and intonation) in a discourse context, was the optimal way to organize a short-term pronunciation course for non-native speakers. McNerney and Mendelsohn (1992:186) express this position very clearly:

> *...a short term pronunciation course should focus first and for most on suprasegmentals as they have the greatest impact on the comprehensibility of the learner's English. We have found that giving priority to the suprasegmental aspects of English not only improves learner's comprehensibility but is also less frustration for students because greater change can be effected in a short time*8.

Nowadays, we see signs that pronunciation instruction is moving away from the segmental/suprasegmental debate and towards a more balanced view. This view recognizes that both an inability to distinguish sounds that carry a high functional load (such as **/I/** in list and /i:/ in least) and an inability to distinguish suprasegmental features (such as intonation and stress differences in *yes/no* and alternative questions) can have a negative impact on the oral communication and the listening comprehension abilities of non-native speakers of English. Today's pronunciation curriculum, thus, seeks to identify the most important aspects of both the segmentals and suprasegmentals, and integrate them appropriately in courses that meet the needs of any given group of learners(Morely, 1994 a). In addition to segmental and

[8] One of the best known textbooks that follows this line of thinking is Gilbert (1993)

suprasegmental features of English, there is also the issue of voice quality setting; that is, each language has certain stereotypical features such as *pitch level*, *vowel space*, *neutral tongue position*, and *degree of muscular activity* that contribute to the overall sound quality or *accent* associated with the language.

The good example is a representation of the best effort to compile a comprehensive volume on the teaching of North American English (NAE) pronunciation to non-native speakers of English. As such, the text provides a detailed treatment of both the NAE sound system and grounding in classroom methods and techniques for the teaching of pronunciation. The desired end result is to equip teachers who use the text (either as a course text or reference) with the background and skills to address the pronunciation needs of their students. The North American variety of English has been selected for rather obvious reasons: mainly as it representsa variety that has gained a strong foothold in much of the world, where English is taught as a foreign or additional language is the target variety of many ESL students, who are living, studying or even working in North America (Celce-Murcia,1983).

The underlying philosophy of the text mentioned above is simple: only through a thorough knowledge of the English sound system and through familiarity with a variety of pedagogical techniques, many of which should be communicatively oriented, can teachers effectively address the pronunciation needs of their students. It is our aim to provide a substantial knowledge base and to assist teachers in teaching and assessing their learners' pronunciation in a correct way. Given this knowledge base, we believe that teachers can continue to improve their understanding of the English sound system and expand upon the instructional tools that are now–and will become-available to assist learners in this skill.

4.2. Research on the Teaching and Acquisition of Pronunciation Skills

As outlined previously, the teaching of pronunciation has experienced the same methodological 'swings of the Pendulum' (Prater, 1991) over the years that have characterized the evolution of ESL teaching. With the wide range of curriculum options available, and the lack of a clear consensus regarding any one best way to teach pronunciation (Macdonald, Yule and Powers 1994), informed decisions on the part of the teacher or curriculum designer are of paramount importance.

Several factors underlie the effective teaching of pronunciation. These factors focus on the learner and involve the effects of age, exposure to the target language, amount and type of prior second language instruction, aptitude, attitude and motivation, and the role of the learner's first language on the phonological acquisition of a second language in addition to the spelling interferences, bearing in mind that English is not a language of spelling (Leather and James, 1991 and Pennington, 1994). These factors will help identify what is taking place in the complex world of pronunciation research, to judge the current status research, and to notice the relevance of this research to teaching.

4.2.1. Learner

The first issue encountered in designing the pronunciation curriculum is perhaps the one most immediately evident-the learners themselves. As Wong (1987 b: 17) aptly points out, the teaching of pronunciation '*is not exclusively a linguistic matter*', and we need to take into consideration such factors as our learners' ages, exposure to the target language, amount and type of prior pronunciation instruction, and perhaps most importantly their attitude towards the target language and their motivation to achieve intelligible speech patterns in the foreign language.

As teachers in the pronunciation classroom, we clearly have little control over certain of these factors, such as our students' ages and their amount and type of prior language instruction. However, we need to be aware of how these factors figure in determining performance in speaking English (or alternatively in colouring attitudes towards such performance). For those factors that we can influence (i.e., attitude and motivation), we need to be aware to what degree they determine the acquisition of the target language phonology.

Let us briefly examine each of these learner-based factors, which influence the learning of pronunciation.

4.2.2. Age

Given the ability of many adult second or foreign language learners to attain target-like proficiency in morphology and syntax, their apparent inability to attain native-like proficiency in pronunciation has often intrigued linguists and non-linguists alike. Scovel (1969,1988) terms this lack of adult facility in acquiring second language pronunciation the 'Joseph Conrad phenomenon' after the famous Polish –born author who, despite the brilliant control of the lexis, syntax, and morphology of English displayed in his literary works, was unable to reach anywhere near the same levels of perfection in his acquisition of English phonology. (Conrad's speech, in fact, remained partly unintelligible to English speakers throughout his life.).Subscribing to the idea that "*you can't teach an old dog new tricks*", many would claim along with Scovel that adults are unable to achieve perfect or target-like pronunciation in a second or foreign language. This view goes hand in hand with the generally held notion that prepubescent children with adequate exposure to a second language can achieve perfect or near perfect pronunciation with relative ease.

One line of research that supports these claims was originally formulated by Penfield & Roberts (1959) and Lennberg (1967). This research considers a period (occurring around puberty) after which brain *lateralization*, or the assigning of certain functions to the different hemispheres of the brain, is completed. The period prior to the completion of lateralization, called the *critical period*, represents the biologically determined period of life during which maximal conditions for language acquisition exist. The implications of this theory, as it relates to second language acquisition, are quite clear. Scovel (1969) and later Krashen (1973) claimed that along with lateralization (*which according to Krashen occurs as early as age 5*) comes an increasing loss of brain plasticity, which renders an individual incapable of achieving native -like pronunciation in a second or foreign language after puberty.

Not all second language researchers, however, advocate to the critical period hypothesis. Flege (1981) cites the lack of empirical evidence to support this claim, contending that, *"Neither physiological maturation nor neurological reorganization renders an adult incapable of speaking a foreign language without an accent*" (p.445). Others (e.g., Brown 1994) would argue that psychomotor considerations figure into the picture as well. In other words, while native like command of morphology and syntax in a second language may be the result of plasticity in the central nervous system, the command of second language phonology also involves the neuromuscular realm, which may play an even more crucial role in the overall picture (Jacobs,1988). Finally, many would argue that the critical period hypothesis overlooks such differences between child and adult second language acquisition as exposure to the target language, linguistic expectations of interlocutors, ego permeability, attitude towards the second language and type of motivation.

Overall, then, the importance of the critical period is somewhat downplayed today, and the claim that adults cannot achieve native like pronunciation in a second language is not infrequently countered with anecdotes about successful adult second language learners who have *beaten the odds*. Scovel (1988) suggests that when native speakers compliment a foreigner on *perfect* pronunciation there is usually some exaggeration involved- for example, when people say: *I'm amazed that you sound just like a native speaker!* They really mean, "*One speaks their language brilliantly-especially for a foreigner!*"(Celce-Murcia, 1996,p.16-17)

More recently, cognitive scientists have concerned themselves with the issue of aging as it relates to brain plasticity and the creation of perceptual networks. Rather than posting one critical period for language acquisition, these researchers propose that there are a number of ***sensitive periods*** during which different aspects of language acquisition occur.

Research in the field further indicates that children and adults perceive sounds in a very similar manner (Lieberman and Blumstein, 1988), and that differences between the two age groups may be related more to the information available (including external circumstances) than to any innate differences in ability (Massaro, 1987). In fact, according to cognitive scientists, the idea of the adult brain *atrophying* or in some way becoming incapable of producing new sounds is an erroneous one, since the brain retains a measure of flexibility or *plasticity* throughout its life (Diamond, 1988).

However, it is undoubtedly the case that adults will acquire the phonological system of a second language in a manner different from that of their first language, given that the acquisition of the new sounds in the second language must be integrated into already existing neural networks. Jacobs (1988:327) went to say that: "*Biological factors impose limitations much the same as psychological and socio-cultural factors..., but none of these variables in isolation imposes an absolute upper bound on" [second language acquisition].* Adults are then capable of rising to the challenge of performing competently in a new sound system.

Of course, factors other than the brain's ability to create new neural networks for the processing and production of the target-language sounds also play a role.

Scott (1989), for instance, demonstrates that auditory perception diminishes with age, especially for those over 60, a factor that would definitely hinder older adults in their attempts to acquire target-like pronunciation in a second or foreign language. A similar caution is sounded by Jacobs (1988), who notes that the environment in which adults typically learn a second language (i.e., the classroom) may not be as rich as that experienced by children acquiring a second language in a more natural, input-rich environment. Thus, when we discuss child/adult differences in phonological acquisition we may be comparing the proverbial apples and oranges.

Ausubel (1964), Guiora (1972), Schumann (1975), and others noted that the disparity between child and adult performance may be explained through a complex interplay of social and psychological factors. In sum, we should be cognizant or aware of such external factors when drawing any conclusion about age.

The implications of the foregoing theories with regard to the teaching of pronunciation deserve reflection. For example, if (as some research indicates) adults are capable of acquiring a high degree of pronunciation accuracy in a second language but are more impeded in their acquisition of target language phonology by non-linguistic factors than are children, then we need to build into courses for adults more fluency and confidence building activities(Wong, 1987 a; and Bowen and Marks, 1992); we should also have our adult learners seriously examine their personal goals in the pronunciation class. Likewise, if Scovel's (1988) claims concerning the inability of most adults to achieve target-like pronunciation are valid, then teachers need to redefine the goal of the pronunciation class as comfortable intelligibility rather than accuracy, and ensure that this goal is reflected in the methods, activities, and materials of the ESL or EFL class.

4.2.3. Exposure to the Target Language

According to the language learning theories of Postovsky (1974), Asher (1977), and Krashen (1982), among others, learners acquire language primarily from the input they receive, and they must receive large amounts of comprehensible input before they are required to speak. If true, learners' exposure to the target language will be a critical factor in determining their success. In EFL settings, especially those where students have little opportunity to surround themselves with native input in the target language, and to ascertain that students have opportunities outside of class (e.g. in language–laboratory or learning-centre environments) to experience samples of the authentic oral discourse of the native speakers, similarly, it will fall to teachers to encourage out-of- class conversational use of the target language. However, even in ESL settings, where the learners are surrounded by the English–speaking world, many speakers live in linguistic *ghettos* with relatively little or non-exposure to native speakers of the target language in their homes and even at work. Again, in such cases, the teacher should try to maximize students' exposure to the target language and to encourage them to expand their own domains of linguistic competence, stressing the importance of language exposure in the process of acquiring all aspects of language: pronunciation, grammar and vocabulary.

4.2.4. Amount and Type of Prior Pronunciation Instruction

Assuming that we are dealing with learners who have had prior exposure to English, we also need to examine the amount and type of prior pronunciation instruction students have had. In EFL settings, instruction may have taken the form of repetition drills led by a teacher whose own pronunciation is different from the target norm. Alternatively, in an ESL multi-skills class, pronunciation may not have been explicitly dealt with at all, and students may not have been fully aware of their pronunciation problems. Whatever, we need to recognize that in any pronunciation class at the intermediate or advanced levels of proficiency, we may be dealing with somewhat fixed or systematic pronunciation errors. Thus, the syllabus and techniques that we implement must be tailored to the types of problems we discern among our students. Therefore, if learners have had good pronunciation training before, this will help them. But, if they have had ineffective training or simply no training, they are at a disadvantage.

4.2.5. Aptitude, Attitude, and Motivation

Providing the fact that students are different from each other and some are inherently more capable of acquiring a good pronunciation than others, Skehan's (1989) overview of

Carroll's (1962, 1981) research on language aptitude is useful here. According to Carroll, there are four traits that constitute language aptitude:

a. **Phonemic Coding Ability**: the capacity to discriminate and code foreign sounds such they can be recalled.[12]
b. **Grammatical Sensitivity**: the ability to analyse language and figure out rules.
c. **Inductive Language Learning Ability**: the capacity to pick up language through exposure.
d. **Memory**: the amount of rote learning activity needed to internalize something (a new sound, a lexical item, a grammatical rule, the pronunciation or spelling of a word, etc.)

Our main concern here is the first trait, although the memory trait is also relevant. Some learners are in fact fairly balanced in these four traits, whereas others have very strong patterns of strength and weakness. Learners who are weak in phonemic coding ability would therefore have much more difficulty achieving a readily intelligible pronunciation than those with high aptitude in this domain. Teachers need to be sensitive to such learner differences and not expect all learners to achieve the same level of success in the same amount of time.

Snow and Shapira (1985), on the other hand, discount the importance of aptitude, pointing out that we have all demonstrated language learning ability via acquisition of our native language. One argument against assigning the determining role to aptitude, according to these researchers, is the fact that low-ability learners (as measured by language aptitude

[12] According to Skehan, (1989) phonemic coding ability is the language aptitude trait that relates least to one's general intelligence. This suggests that *having an ear for language* may be qualitatively different from other language aptitudes or traits.

tests) are in fact often able to attain fluency in a second language while some high ability learners are not.(Skehan,1989).

As should be evident by now, the network of factors influencing an individual's acquisition of second language phonology is a tremendously complicated one. Indeed, as Stevick (1976) suggests, we need to go beyond language aptitude and educational or cultural experience to see how individuals and their personalities affect the learning processor help in understanding learners' attitudes towards the target language and their motivation (or lack therefore) to acquire this language.

Guiora (1972) notes that personality, or in his words language ego, is at the very core of the language learning process, especially where the skill of pronunciation is concerned:

> *Speaking a foreign language entails the radical operation of learning and manipulating a new grammar, syntax, and vocabulary and, at the extreme limits of proficiency, modifying one of the basic modes of identification by the self and others, the way we sound.*(p.144).

Pointing out the often dramatic discrepancy between certain individuals' attainment in pronunciation versus their attainment in other skill areas, Guiora postulates that accent or pronunciation is a unique feature of language performance-one that can provide "*the key to the extent to which the individual is psychologically capable of stepping into a new system of communication"(*p.144).[14]

[14] Several studies have attempted to ascertain whether artificially inducing the language ego to become more permeable (i.e., via ingestion of alcohol or valium, or via hypnotism) will result in an increased degree of target like or native like pronunciation. The results are somewhat mixed, although the studies provide some evidence of a correlation between ego permeability and accuracy of pronunciation. Guiora (1972), Guiora et al.(1972) and Shumann et al.(1978).

Following in Guiora's footsteps is Schumann's work on the role that acculturation plays in the process of language acquisition. Schumann (1975) echoes Guiora's hypothesis that *ego permeability* (i.e., the extent to which the ego can be flexible and adapt) and personality factors are at the heart of second language acquisition. Schumann further states that in adults, the development of firm ego boundaries, along with individuals' attitudinal and motivational orientations, can place constraints on the cognitive process of language learning. Given such constraints, adults might well be hindered from attaining their biologically determined capabilities.

Schumann's *acculturation model* (Schumann 1986) clearly delineates the role that social and affective variables may play in language acquisition. This model based on the premise that certain social and affective variables *cluster* into a single variable of acculturation, states that learners will acquire the target language to the degree that they acculturate. Two types of determining factors are:

a. Those concerned with the language learning of a group of people, or sociocultural variables (e.g., social dominance patterns, size of the foreign language population, amount of congruence and equivalence between the foreign and target language cultures).
b. Factors concerned with individual differences, or affective variables (e.g., ego permeability, personality, type of motivation, degree of culture shock).

Schumann notes that sociocultural variables do not prohibit successful second language learning. That is, individuals may learn languages successfully under sociocultural conditions that are not favourable, or vice versa –they may not learn a language under sociocultural conditions that are highly favourable. Thus although the two sets of variables always interact, affective variables appear to carry more weight than sociocultural ones in determining any learner's acquisition process.

In applying this model, Schumann differentiates between two types of successful acculturation. In the first type, the learner demonstrates *integrative motivation*- that is, a desire to be socially integrated in the target culture.15This second type of motivation appears to be akin or similar to that described by Graham (1985) as assimilative motivation, and implies a desire on the part of the learner to become an indistinguishable member of the target speech community (*assimilative motivation*, which is rare among adult second language learners, is what all children have when learning their first language.). Accordingly, one can hypothesize that this second type of learner would willingly embrace the target culture, and would therefore be more apt to acquire native-like pronunciation in the second language. *Instrumental motivation*: in which an individual learns a second or a foreign language to attain a certain goal, for example a job promotion, does not contribute to successful acculturation, according to Schumann. However, other researchers such as Lukmani (1972) argue that the intensity of motivation is often as important as the type of motivation to play. In other words, someone with extraordinarily high instrumental motivation (e.g., someone who wants to sound like a native speaker in order to function effectively as an actor for instance) may well achieve a better pronunciation than someone with integrated motivation which is quite positive yet less intense.

[15] This form of motivation has often been discussed in the literature as a positive force in language acquisition (Gardner and Lambert, 1972)

4.2.6. Role of the Native Language

Whether our students are from the same language group (for instance the case of EFL settings), or from diverse language backgrounds (as in most often the ESL classes), we need to consider their native language in deciding about pronunciation priorities. Therefore:

- The learner's native language affects the learning of pronunciation.
- Sometimes this effect is bad, but sometimes it is good. There can be both positive transfer and negative transfer.
- Unfamiliar sounds or sound combinations may be difficult to pronounce. Thus :

  -An L1 sound may be substituted for an L2 sound.

  -The phonological rules of L1 may be mistakenly applied to L2.

- Fossilization: old habits are hard to break, but it can be done with motivation to improve.

4.2.6.1. Spelling Interferences

The written language is a significant source of input to EFL students. The higher the spelling inconsistency is, the more negative the influence on pronunciation will be. The English spelling does not serve as a guide to pronunciation. On the contrary, it can often be misleading. For instance, '*s*' has got five different ways of pronunciation**:**

i) /s/ as in ***basic, past, say***

ii) /z/ as in ***because, rose***

iii)/S/ as in ***sugar, sure***

iv)/Z/ as in ***casual, usually, television***

v) /*mute*/ as in ***aisle, island***

The frustration of the EFL learner never ends. Let us consider the following examples:

i. ***bough*** /baU/

ii. ***cough*** //kQf/

iii. ***dough*** /d@U/

iv. ***rough*** /rVf/

v. ***through*** /Tru:/

vi. ***hiccough*** /hIkVp/

Conclusion

In this chapter, we have been through a thorough historical overview of how pronunciation, as a skill, has been dealt with in language teaching over the past thirty years. Different types of teaching methods, approaches and techniques have been provided with a detailed description of each one's principles towards pronunciation. This is in addition to the illustrations of how teachers have provided learners with explanations and phonetic analysis.

Chapter Five

Chapter Five: Computer-aided Learning and Pronunciation Teaching

Introduction

The history of computer aided learning of the pronunciation of a foreign language goes back at least to 1964 (Vardanian, 1964), with the first attempt to display the learner's voice pitch on a screen in order to improve the perception and the realization of sentence prosody.

Pronunciation, for long on the periphery of applied linguistics research and pedagogy, continues to grow in importance because of its central role in speech production, speech perception, speech recognition, and even speaker's identity. Pronunciation related issues such as comprehensibility, accent, and the mutual intelligibility of the varieties of English are subject to many questions in applied linguistics. This calls for a sophisticated understanding of how technological tools that have long been used to shed light on phonological categories can be used and applied to teaching. Research into computer–assisted pronunciation teaching CAPT suggests that both researchers and pronunciation teachers increasingly make use of technology to answer key questions and to develop theories and practices that match a lot more acoustically.

We will review areas where computer technology and pronunciation intersect using CAPT in general and Praat software in particular in:

i. Appropriate pedagogical goals and the measurement of improvement
ii. The ability of CAPT to give useful, automatic feedback and
iii. The use of technology in diagnosing pronunciation errors.

Teachers and researchers have had high hopes for computer assisted pronunciation teaching, or CAPT, for several decades (Molholt, 1988). The use of computers is almost ideally suited to learning pronunciation skills. Computers can provide individualized instructions, frequent practice through listening discrimination and focused repetition exercises, and automatic visual support that demonstrates to learners how closely their own pronunciation approximates model utterances. In a foreign language teaching environment, in which few teachers receive adequate training in teaching pronunciation (Breitkreutz, Derwing, &Rossiter, 2002; Burgess & Spencer, 2000; Kawai &Hirose, 2000, MacDonald, 2002; Murphy,1997), and because of little or no time allotted to work on pronunciation and emphasis is given to other skills , CAPT seems to meet the pedagogical needs mainly with innovative applications .

A variety of experimental software applications and studies concerning how various features of pronunciation might best be taught, have demonstrated the flexibility and efficacy of CAPT. A wide variety of pronunciation features including mainly instant pronunciation quality, have been examined, for instance, speech rate, fluency, vowels and consonants, vowel lengthening and pitch accents, intonation and English stress timing.

Our review will necessarily take the perspective of a pronunciation practitioner, who is much more interested in what and how CAPT and all software in general can do and help students as well as teachers to better the teaching and the learning of an accurate pronunciation. We will first look at the pedagogical goals that can be reached once the applications are appropriately used and become more effective unless the teachers are not able to make effective use of applications because of both a lack of training in pronunciation and in the use of technology.

5.1. Pedagogical Goals

They are to:

i. Provide immediate, useful feedback, especially for those features that are most important for intelligibility.
ii. Provide a substantial amount of meaningful input, including the use of multiple models and accurate articulatory instruction.
iii. Give the students a reason and desire to practice through rule-oriented practice and realistic materials.

According to Pennington (1999, p. 434), most CAPT programs showed the range of acceptable variation in pronunciation and how to apply this knowledge to teaching. Some of Pennington's principles for the design of CAPT materials, which should according to her:

i. Establish baseline, reference accents for instruction.
ii. Set measurable goals and performance targets.
iii. Be designed to build skills from easier to more challenging exercises.
iv. Link pronunciation to other aspects of communication.
v. Raise users 'awareness of how their L1 phonological systems differ from the system of the target language.

Some studies explore new ways to teach the exploit of the computers' strengths. Wang and Munro (2004) suggested expanding the range of pedagogical techniques to include those typically employed for research but not usually for teaching. In training learners to recognize difficult vowel contrasts, they used identification tasks (rather than discrimination of minimal pairs) and synthesized speech to expand learners' perceptual space, making learners more sensitive to the range of phonetic variation in the pronunciation of vowel phonemes. Such an expansion of perceptual space requires, hence, the use of multiple models of voices, which is a practice advocated but many researchers; yet, in case there is an absence of multiple voices in CAPT applications, then the speech synthesis would be very useful. These multiple speech models, provide variation in input in some recent software for English, such as *Connected Speech* (ProteaTextware, 2007[4]), *Streaming Speech* (Claudwell, 2002)...etc.

Pronunciation teaching is subject to two overlapping and conflicting foci (Levis, 2005): a focus on accuracy (involving instruction on all possible aspects of pronunciation, with little attempt to prioritize attention) and a focus on intelligibility (involving instruction for only the elements of pronunciation that are critical for communicative success, while ignoring those that are not). This leads us to speak about the errors of pronunciation, which are not all equal in their impact on listeners (Jenkins, 2000).

Because some errors influence listeners' understanding more than others do, it is now generally agreed among pronunciation researchers that the goal of instruction is not perfection but rather being good enough to achieve communicative goal. '*Good* enough' here is usually referred to as *intelligibility, which* is whether a speaker is understandable (Dalton & Seidlhofer, 1994; Morley, 1991). Hence, *intelligibility*, in a general sense, however, includes two main types of understanding, which researchers call *intelligibility* and *comprehensibility*. In other words, and in a specific sense, intelligibility refers to whether listeners can perfectly and adequately decode the words pronounced in context by a speaker. On the other hand, *comprehensibility* refers rather to whether listeners can grasp and understand the message communicated by a given speaker (Munro &Derwing, 1999).

[4]ProteaTextware. (2007) Retrieved July. http://www.proteatextware.com/

5.2. Improvement in Using CAPT

Many questions have been raised as far as pronunciation teaching is concerned, for instance does the CAPT help improve students' pronunciation? Should pronunciation be taught? Moreover, can students achieve a native–like accent? And so forth.

As far as a native–like pronunciation is concerned, a question that we may ask is that should language learners be required to learn a native –like pronunciation of English?

Celce-Murcia et al. (1996) suggested that in the communicative approach to teaching, the teaching of pronunciation was urgent and important because the non-native speakers need to have a threshold level of pronunciation "*If they fall below this threshold level, they will have oral communication problems no matter how excellent and extensive their control of English grammar and vocabulary might be"* (p.7).

Otlowski and Fraser (1999) concurred with the research that goals of pronunciation teaching should not necessarily be acquiring native–like pronunciation, but it should aim at "*Developing functional intelligibility, communicability, increased self-confidence, the development of speech monitoring abilities and speech modification strategies for use beyond the classroom"*(p.3).

In support of this, Harmer (2001) also noted that "the degree to which students acquired *perfect pronunciation* seemed to depend very much on their attitude to how they speak and how well they hear". A number of psychological issues may affect how *foreign* a person sounds when she/he speaks, and so teachers need to consider intelligibility as the prime goal of pronunciation teaching. Hence, Harmer found *listening* as the key to intelligibility.

Concerning the first question asked above, CAPT seems to be effective in improving pronunciation accuracy. For a wide variety of pronunciation instructions and skills, learners improve by well-designed CAPT instruction.

Hirata (2004) examined how learners of Japanese acquired pitch and duration contrasts and is an excellent example of how CAPT can be used to suggest answers for research issues and instructions for teaching. Japanese is a *pitch accent* language referred to as *a tone language*, with words distinguished by their patterns of *H* and *L* pitches. It is also a language with phonemic lengthening. One experimental group learned the contrasts at the word level, while another learned the same words embedded in sentence-level practice. Both groups improved their production of these contrasts, with improvement being greater for those who learned the contrasts embedded in sentence–level practice. She, then, stated that *"developing abilities to produce and perceive L2 contrasts in sentences instead of only in an isolated word contest, is an important factor for L2 learners coping with real world situations"*(p. 372).

In another study of Japanese, Kawai and Hirose (2000) found that subjects more successfully made phonemic contrasts between short and long segments, with the greatest improvement on the short segments. Wang and Munro (2004) trained learners to pay attention to vowel quality rather than durational cues for the perception improvement of three vowel contrasts among learners of English, /i/ ,/ i:/, /ʊ/ , /u:/ and /æ/ , /ɛ/ **.**They found that subjects increased their perception performance on all the contrasts.

Another study by Haridson (2004) examined whether computer assisted prosody training leads to improved prosody, and ultimately, not only to improved segmental accuracy but also lexical recall. Intensive practice of sets of 30 French sentences was involved, with varied intonational contours. Rather than being models for the subjects, native French speaker sentences were used as feedback to the subjects' initial production. The intensive practice, yet, led subjects to notice segmental features that were not in focus. The prosody also estimated substantial recall of the lexical content of the test sentences, suggesting that the subjects' lexical memory was improved by their prosodic memory built through the training. The study offered implications for the value of suprasegmentals in pronunciation teaching. He found that learners, who practiced French intonation with native speakers' models provided as feedback, improved their intonation production.

Hirata (2004), found evidence that improvements in perception can lead to improvements in production; and production improvements also can lead to improved ability to perceive contrasts; hence, it is possible that improvement in one area spills over to improvements in the other.

Haridson (2004) found that not only did learners improve their intonation, but they also showed improvement in two areas that were not expected: lexical recall and segmental accuracy. She suggested that "*as the subjects became more confident with this aspect "prosody" of their language production, they were able to notice other elements such as liaison and their production of specific sounds"* (p.48).

5.3. Visual Immediate Feedback

The main interest and aim of CAPT is the provision of adequate feedback. This can be done in various ways, the most common methods are visualization and through *Automatic Speech Recognition* (ASR). Therefore, visual feedback includes "*the graphical display of a native speaker's face … [and] the vocal tract"* both of which seem to improve learners' ability to identify words and produce more-native like speech timing (Ehsani &Knodt, 1998, p. 63).

The best known CAPT visual displays, however, are spectrograms, waveforms, and pitch tracing. These visual displays have been advocated and supported for a long time, and they are still being used in research and teaching. Pitch tracings, for instance, is relatively iconic on a spectrogram, with rising, falling, and level lines on the display usually corresponding to rises, levels, and falls in a speaker's voice. Although this requires some training to interpret, they appear to be useful pedagogically.

According to Hardison (2004), who argued that CAPT has certain research and teaching benefits over traditional classroom teaching, first, CAPT is untiring. Teachers simply cannot provide the level of practice and feedback needed for many students to improve. Second, CAPT is consistent; it is always the same in its presentation of stimulus material and in the kind of feedback provided as compared to teachers, who are not. Third, CAPT provides variety in both the numbers of voices used as models and in opportunities for visual feedback, especially in areas like pitch movement and range. Finally, CAPT offers the chance to meet varied individual needs more easily than any teacher can. It promotes learner's autonomy in working on pronunciation, a critical factor in success. Hardison, yet, is not arguing that teacher–led instruction is unimportant. There is sufficient empirical evidence that both instructor-led teaching and CAPT lead to pronunciation improvement. We cannot deny the evidence that no pronunciation teaching is going on in classrooms that one would hope and that most teachers are and feel unprepared to teach pronunciation.

Nonetheless, spectrograms are still being advocated in recent works; for instance, Coniam (2002) used spectrograms to raise language teachers' awareness of the relatively stress-timed nature of American English, and by extensions, other English circles. The teachers that Coniam described were native speakers of Cantonese in Hong Kong, and the variety of English spoken there, Hong Kong English, is relatively syllable-timed. Coniam, in his work, used spoken dialogues from local TV shows involving famous characters speaking both Hong Kong and American English. Therefore, the utterances he used for analysis were analyzed through spectrograms, which demonstrated that Hong Kong English has a greater frequency of energy peaks than American English speech. The visual representation of the energy peaks in the spectrograms was successful in helping teachers understand this difference in the two varieties. Coniam did not ask much of the spectrogram, focusing much more on the relatively iconic visual displays of syllable energy while also providing a significant amount of *scaffolding*[12] that identified the location of the words associated with each peak. That the subjects were language teachers was also significant. His aims and goals were that these expert users start to understand the nature of rhythmic differences rather than produce those differences. These restrictive goals contributed to the successful use of spectrograms.

Lambacher (1999), indicated success with teaching English consonants to Japanese learners using spectrograms. For difficult consonant features and contrasts (i.e., aspiration, nasals, contrasts between [r/l],[s/ʃ],[f/h] and [s/Ɵ], he illustrated how the contrasts look on visual displays built into the CAPT program. He advocated not only the use of the spectrogram but also through standard pronunciation tricks, for instance, he taught aspiration with reference to a spectrogram and by blowing away a piece of paper.

Pitch tracing, or models of the fundamental frequency (F0) of speech, appear to be much easier to use successfully, yet, requiring training to interpret them.

[12]In education, *scaffolding* refers to a variety of instructional techniques used to move students progressively toward stronger understanding and, ultimately, greater independence in the learning process. The term itself offers the relevant descriptive metaphor: teachers provide successive levels of temporary support that help students reach higher levels of comprehension and skill acquisition that they would not be able to achieve without assistance. Like physical scaffolding, the supportive strategies are incrementally removed when they are no longer needed, and the teacher gradually shifts more responsibility over the learning process to the student.

There are several available Software; some of them are commercial, such as *Kay Elemetrics VisiPitch* or the more technically oriented one, the *Computerized Speech Laboratory (CSL)*; on the other hand, we have many free programs for instance *WASP* and *PRAAT*, all of the previously mentioned programs, provide the technology to represent intonation visually. They also allow learners to compare their own production to a model utterance by overlaying their utterances' pitch tracing on that of the model. *Streaming Speech* uses the discourse intonation model of Brazil(1980) to teach the intonation and pronunciation of natural connected speech, and discourse uses of pitch are advocated by Chun (1998; 2007).

5. 4. Automatic Diagnosis of Pronunciation Errors

Before dealing with automatic diagnosis of pronunciation errors, we will first introduce the Automatic Speech Recognition (ASR), which is a central question in CAPT feedback.

Automatic speech recognition (ASR) can be defined as the independent, computer-driven transcription of spoken language into readable text in real time (Stuckless, 1994). In a nutshell, ASR is a technology that allows a computer to identify the words that a person speaks into a microphone or telephone and convert it to written text. ASR allows software to interpret the meaning of a speaker's utterance. This technology has come so far that in recent years it has been widely used in commercial systems such as travel reservation, weather reports, or sports score reporting.

The earliest attempts to devise systems for automatic speech recognition by machine were made in the 1950s.Much of the early research leading to the development of speech activation and recognition technology was funded by the National Science Foundation (NSF) and the Defense Department's Defense Advanced Research Projects Agency (DARPA)in the

USA. Much of the initial research, performed with NSA and NSF funding, was conducted in the 1980s. Having a machine to understand fluently spoken speech has driven speech research for more than 50 years.

Although ASR technology is not yet at the point where machines understand all speech, in any acoustic environment, or by any person, it is used on a day-to-day basis in a number of applications and services.

The ultimate goal of ASR research is to allow a computer to recognize, in real-time and with 100% accuracy, all words that are intelligibly spoken by any person, independent of vocabulary size, noise, speaker characteristics or accent. Today, if the system is trained to learn an individual speaker's voice, then much larger vocabularies are possible and accuracy can be greater than 90%. Whether ASR can effectively provide immediate feedback that enables the learners to know which parts of their pronunciation are correct and which are not, many questions have been asked as far as ASR is concerned. For instance, do the words and phrases pronounced by non-native speakers (NNSs) match the models of native speaker's speech on which the ASR is based? Can the NNS be given global feedback on the mispronunciations? (For instance, if you get a score of 72%, this means that you have words that are mispronounced) or specific feedback on some given sounds or prosodic categories? And at last, how should the feedback about the nature of the mispronunciation be given? Assuming these questions are successfully answered, there is, yet, another concern which is whether ASR can be configured to provide accurate instruction to remedy errors. The last question is a CAPT design issue and not really unique and proper to ASR.

Neri, Cucchiarini, Strik, and Boves (2003) wrote that "*Ideal systems should always include an option to provide feedback by means of ASR technology so that the user can receive immediate information on his/her performance*" (p.458).

Because ASR applications have been successful with native speakers of English, especially in the use of voice recognition for word processing, it might be assumed that they could be adjusted to provide feedback nonnative speakers.

According to Van Compernolle (2001):

> *In fact, much of the progress in the last 15 years in acoustic modeling [for ASR] is based on more detailed modeling, creating sharper and sharper distributions for narrower and narrower classes. This is diametrically opposite of the tolerance and robustness required for non-natives* (p.76).

The interest continues the Research on improving ASR effectiveness for non-native speech. The automatic diagnosis of pronunciation, related to specific errors, is a fruitful application of ASR technology. One study states that *"speech recognition technology is a key to automatic evaluation of pronunciation quality"* (Neumeyer, Franco, Digalakis, &Weintraub, 2000, p.83).The trick is to know which gate is opened by the key, a door to general pronunciation scores or a door that specifies the errors made by the nonnative speakers. Yet, research indicates that the first door can be opened, while the second one remains still shut in search of the adequate key.

Since the key has not yet been found for the door that opens to automatically evaluating particular pronunciation errors, this would be far more useful to language teachers than global measures of pronunciation quality. On the other hand, progress is made between how humans perceive speech and how ASR perception models can make more than incremental or increasing changes in accuracy. To back our statement, Ladd (1980) pointed out that "*Machines don't hear like people because people hear things that aren't there, but the machines do hear very well all the factors which induce us to hear what isn't there"* (p.26).

In pinpointing some specific errors, Cucchiarini et al. (2000) correlated human ratings of Segmental Quality with ASR measures (which worked well for predicting human ratings of fluency and speech rate). They found that "*Segmental quality…is predicted most poorly on the basis of automatic scores, which is not a positive finding if we consider that segmental quality is the best predictor "for human raters" of overall pronunciation*" (p.116).

Truong et al (2005) constructed an ASR system that automatically measured the accuracy of three Dutch phonemes commonly mispronounced by learners of Dutch, /a/, /Y/ and /x/.They found that errors in the pronunciation of /x/ were automatically detected most successfully, but that the ASR system was less successful identifying difficulties with the vowels as compared to consonants. Nonetheless, the approach seems on target for automatic evaluation. Human raters do not need to hear every instance of a phonemic error to make the judgment that there is a systematic problem. In addition, even though learners do not always mispronounce sounds that are difficult, a system that can identify frequent errors could be very useful for learners and teachers. What is clear is that a lot of work needs to be done defining the parameters of the segmental in a language and including them in ASR systems before precise feedback on individual segments is possible.

The success of an ASR system depends on the uses to which it is set. Measures of pronunciation may be useful for some contexts, such as testing (e.g., Versant Testing, the test previously called Phone Pass and Set -10). For the classroom teacher and the self-directed learner, however, the is a strong need for ASR systems that can pinpoint specific errors and provide help in addressing the errors (Neri et al,2003)

Hardison (2005) has shown how beneficial inclusion of video can be for pronunciation training. As will be shown in the Tell Me More software, video can also be valuable in extending the knowledge gained to real world situations. Indeed, the incorporation of real – world natural speech rather than scripted sentences is another direction that is important. As part of that development, it will be important for software to be able to deal effectively with discourse-level input. In one view, *the use of computer technology has furthered the dominance of sentence-level practice rather than promoting the use of discourse intonation.*

For serving the cognitive domain in pronunciation training, Morley (1991:501) recommended providing adult and adolescent learners with two kinds of information:

> *Language information focuses on the production and modification of specific features through descriptions and explanations; and procedural information helps learners understand what they will do, how and why through explicit directions and guideline.*

For adult learners in particular, correction of their pronunciation in front of others may even discourage them from continuing their study. In pronunciation learning software, there is generally a Playback and comparison facility which enables the user to compare his/her production through listening, while comparing a model for correct pronunciation provided by the system.

Pennington (1999) suggested that developers of CAPT software should define at the outset what performance is counted as having made progress towards or achieved a desired target. Pronunciation teachers commonly emphasize that ear training and production of the sounds go hand in hand in pronunciation instruction. The assumption has been that target- like pronunciation would follow as a result of '*language use in meaningful settings*" (Pennington & Richards,1986 p.: 217)

Raux and Kawahara (2002) reported that recent computer-assisted pronunciation learning focuses on two major areas: evaluation of the learner's current pronunciation and instruction of the different aspects of pronunciation such as stress and intonation, both of which can be enhanced with the use of Praat as an instructional tool.

Neri, Cucchiarini, and Strik (2002) pointed out that digitized pronunciation software allows students to individually access unlimited and realistic L2 input through different channels and provides individualized feedback automatically and instantaneously (p.179.188).

As far as Praat Software is concerned and will be used in our work, Setter and Jenkins (2005) in their state of the art review of pronunciation teaching, point out that being able to successfully interpret formant plots produced by Praat requires *a sophisticated level of understanding* on the part of both teacher and learner (p.10).

Wilson's study (2005) refers to the advantages of Praat to teach segmental and suprasegmental pronunciation such as vowels and diphthongs. He says some features like vowel length differences before voiced and voiceless stop or voiced onset time VOT or intonation and stress can be shown and measure by this software. After being trained by the teacher on the use of Praat, students are able to record and analyze their own pronunciation.

Wilson further explains:

> *Although Praat is used by many pronunciation teachers and students, its interface is designed more with the scientist/researcher in mind. Nonetheless, it is extremely useful in pronunciation classes and is currently being used as both a teaching tool and a pronunciation aid in Phonetics and pronunciation courses at the University of Aizu. After being trained by the teacher on the use of Praat, students are able to record and analyze their own pronunciation. Although pronunciation is often judged and taught solely through the oral/aural medium, this use of Praat opens up analysis to the visual medium as well.*

All in all, it seems that Praat, when used in a language classroom, especially in an EFL context, where pronunciation should be emphasized further, could be really effective in teaching pronunciation and prosodic features of English language. Praat, being open- source software for the acoustic analysis of speech is easily downloadable and accessible for both students and teachers.

Conclusion

We have reviewed a wide range of literature, which showed the paramount importance of the use of CAPT in language learning in general and pronunciation in particular.

For any teacher, who thinks that pronunciation is essential, Praat is immensely promising. We admit that pronunciation is taught infrequently and unsystematically; this is due to several reasons. A lack of teachers' training leads to untrained and unconfident teachers; the varied needs of learners make group instruction irrelevant for some, especially in EFL contexts, where classes have learners of many L1 and varied pronunciation needs.

Pronunciation as a scope, in competition with other language skills, has simply a lesser priority in many classrooms and is often not taught at all. Yet, CAPT promises a way out of this bind, allowing teachers to have access to pronunciation teaching that hopefully goes beyond their own skills, providing individualized and offering additional instructional time in a language laboratory or outside of class, since learners can freely install the software and exercise at home.

Chapter Six

Chapter Six: Pronunciation Teaching Programs and Software

Introduction

Pronunciation, which tends to be neglected in our classrooms, due to several reasons among which are mainly lack of trained teachers and lack of time allotted to phonetics (one hour and half per week) and the scarcity of language laboratories. That is why pronunciation would be better taught and with better results when using a variety of interactive software packages and programs for providing students with the opportunity to perceive and practice pronunciation. The most used and well known software and teaching programs are overviewed, explained and illustrated below.

6.1. Multimedia Programs and Software

The most important and most efficient software are: Tell Me More, Rosetta Stone, Reflex English Cambridge, DLL Digital Language Lab, Winpitch (LingLab), Audacity, SpeedLingua, Speech Analyser, WinSnoori and Praat.

6.1.1 .Tell Me More

It is, by far, the first and one of the leaders in the language learning software based on advanced speech recognition and on the communicative approach in language teaching. It was founded under Auralog in 1987 after the company brought together a team of technical and linguistic experts. This method is based on the communicative approach rather than a simple memorization of grammatical rules and lists of words. This educational approach encourages *interaction* and the use of language skills in *real situations*. It is recommended by expert linguist experts from around the world and is used in some of the largest universities.

Foremost, it is suitable for all learners, beginners, intermediate and advanced ones. It features vocabulary, grammar, reading, writing, listening comprehension and conversation practice lessons. Its installation on a computer is simple and takes about 2 minutes to finish the whole process.

One of the nicest features is its use of speech recognition, which makes the program highly interactive. The speech recognition functionalities make it easy for the learner to interact with the program and empower the learner to practice proper pronunciation. Another nice feature is that, in case a word is not understood, by a simple click on it provides its meaning, translation or pronunciation.

It offers three study modes (Guided, Free-to-Roam or Dynamic). All of them automatically adjust themselves to your individual level in your personal language learning process.

6.1.1.1. What does it include?

The *Tell Me More* is a full multimedia language course. Included in the courses are videos which display every day scenarios that are likely to be encountered in the visited country. It allows practicing conversations with voice recognition and provides the opportunity to interact with the material. *Tell Me More* covers all aspects of language learning including speaking, writing, reading, grammar and cultural knowledge. There are also a number of additional features included such as 24/7 access to an online course advisor, weekly news which uses vocabulary, video and grammar lessons to teach a different news story each week, and progress and placement tests. Tracking tools provide a visual display of your progress with both charts and graphs (https://www.languagesoftware.net/tellmemore-review.html).

Figure 6.1: Tell Me More Software

6.1.2. Rosetta Stone

Rosetta Stone is another worldwide leader in computer-based language learning software. Its comprehensive, immersive program has been called '*the next best thing to living in a country*' by the *Wall Street Journal*,[1] and it is used throughout the world by individuals, businesses and governments looking to gain language proficiency for themselves or their employees.

Rosetta Stone is a computer program that helps people of all levels learn a new language or improve their current language skills. Individuals can commit to a timeframe for learning that fits with their own schedule and time commitments.

[1] rosetta-stone-language-learning-program.html

The *Rosetta Stone* software is an intelligent program, as it tailors its courses for each individual user. The program employs an interactive and immersive approach to learning, combining images, voice, and text to convey information without the need for translations. The course covers all aspects of language learning, with an emphasis on speaking and listening. The lessons are presented in a varied and entertaining format to ensure that the user not only learns their target language but has fun while doing so.

Rosetta Stone is a proprietary Computer-Assisted Language Learning (CALL) software published by Rosetta Stone Inc. The software uses images, text, sound, and video to teach words and grammar by spaced repetition, without translation.

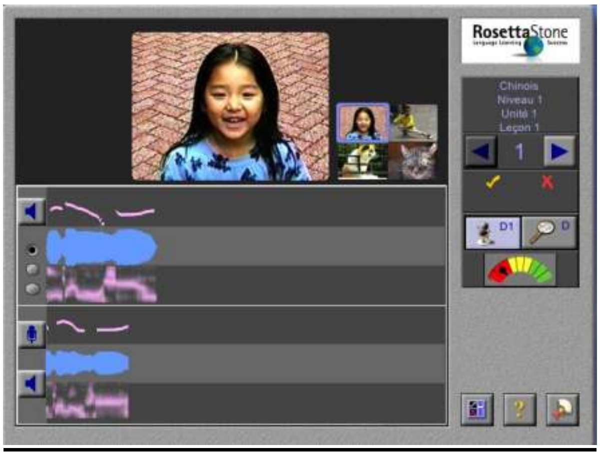


Figure 6.2: Rosetta Stone Software

6.1.3. Reflex English Cambridge

Speech recognition only works on the PC version. *Reflex English Cambridge* is an interactive English course for the new generation. Based on the *Language in Use* exclusive method Published by Cambridge University Press, *Reflex' English Cambridge* effectively combines education and multimedia technologies. From the very user-friendly graphical interface to pedagogy that fits various needs, Reflex' English Cambridge meets all the essential elements to ensure a comprehensive and effective learning. Several working modes give the opportunity to follow the method or customize the journey to better suit the needs.

Personal lessons can also be created with the items of interest, the learning time may be organized at any desired pace. The new voice recognition system (only works for the PC version) very accurately assesses pronunciation and indicates errors.

Pronouncing complete sentences, phrases, words or sounds may be practiced. The program features thousands of interactive exercises, recordings, photos, illustrations, videos, staging, interactive dialogues, narratives and language games to facilitate understanding and language mastery. It includes a multitude of tests to verify the effectiveness of learning in each language skills: grammar, vocabulary, pronunciation, reading comprehension and listening comprehension. The results of practice are provided along with comments and instructions to guide revisions. *Reflex' English Cambridge* is essential to effectively prepare for exams such as KET (*Key English Test*), PET (*Preliminary English Test*) and FCE (*First Certificate in English*).

Figure 6.3: Reflex English Cambridge Software

6.1.4. DLL-Digital Language

DLL is a language learning software that provides all the features required by teachers for language learning such as pronunciation practice video presentation, audio broadcasting, quiz and exercise.

The content for DLL is highly flexible. Teaching materials can be tailor-made with the aid of DLL content builder. With DLL, language learning becomes easy, flexible and interactive.

Figure 6.4: DLL Language Learning Software

6.1. 5. Winpitch Linglab

It is a new version of the pronunciation software program WinPitch. It was founded by Mr. Martin, our teacher of *Traitement de Signal* and *Phonétique Expérimentale*. This completely redesigned version is much easier to use by the teacher and learner alike and contains the regular function of prosodic real time display, variable speed playback, prosodic morphing, on screen teacher comment display, etc. New features include video capabilities (the program can process multimedia files in most formats) as well as automatic alignment of the learner's imitation of the teacher's models. This allows for an automated comparison and explanation of the differences analyzed at the segmental and suprasegmental levels. The new version of WinPitch LTL (Language Teaching and Learning) has a built in program for the preparation of lessons in any language (the program is Unicode compliant) allowing easy navigation by the learner between examples contained in each unit. Segments of prosodic curves can be highlighted in any color and text easily added for on screen explanation of specific melodic or rhythmic properties of the models being studied.

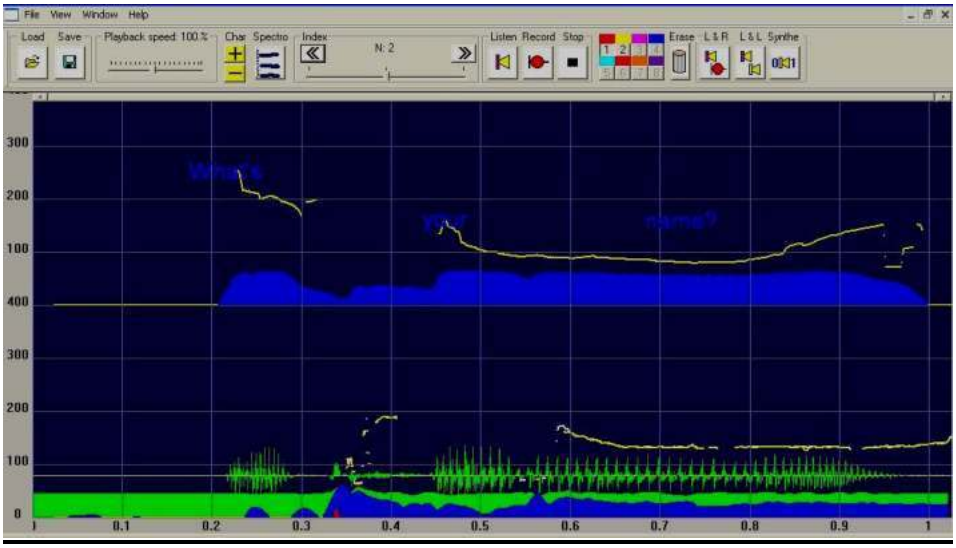


Figure 6.5: Spectrogram with WinPitch Software (What’s your name?)

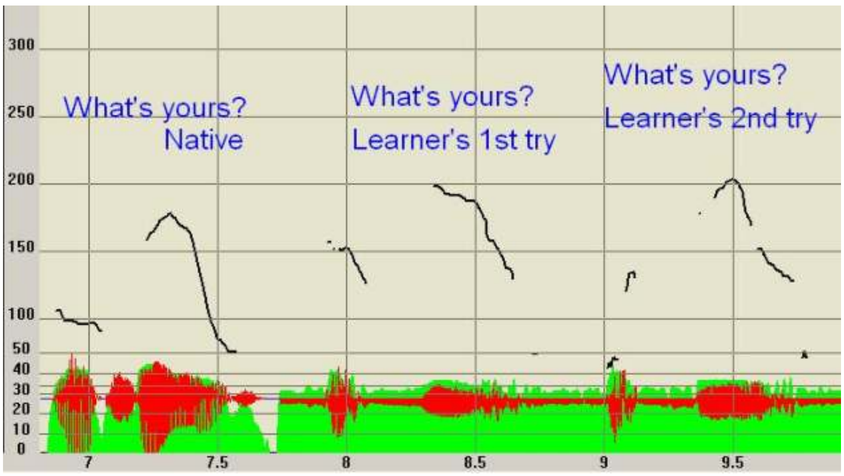


Figure 6.6: Intonation with WinPitch (What’s yours?)

6.1.5.1. Winpitch Illustrations

The following spectrograms are some illustrations of recorded sentences; practiced showing the waypitch can change from questions to statements using the Winpitch software:

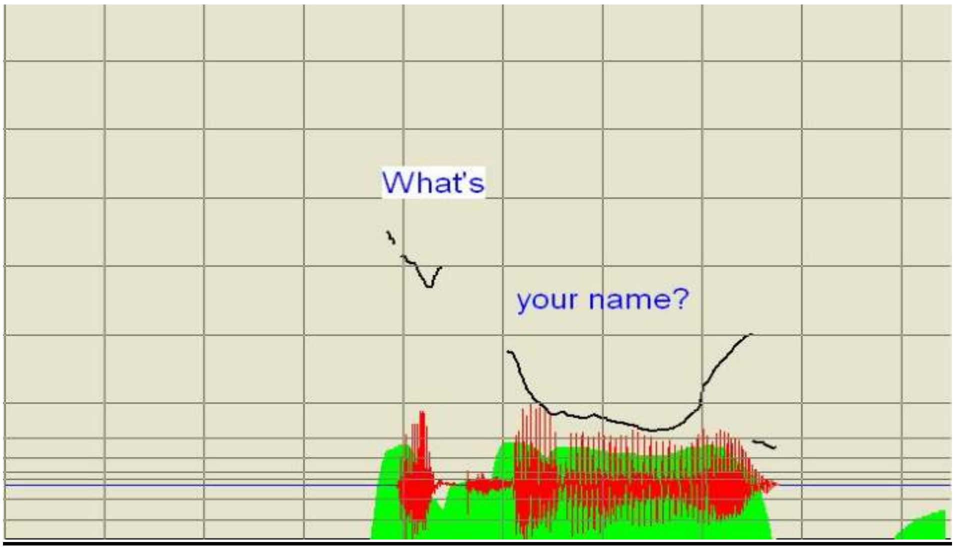


Figure 6.7: WH question: What's your name?

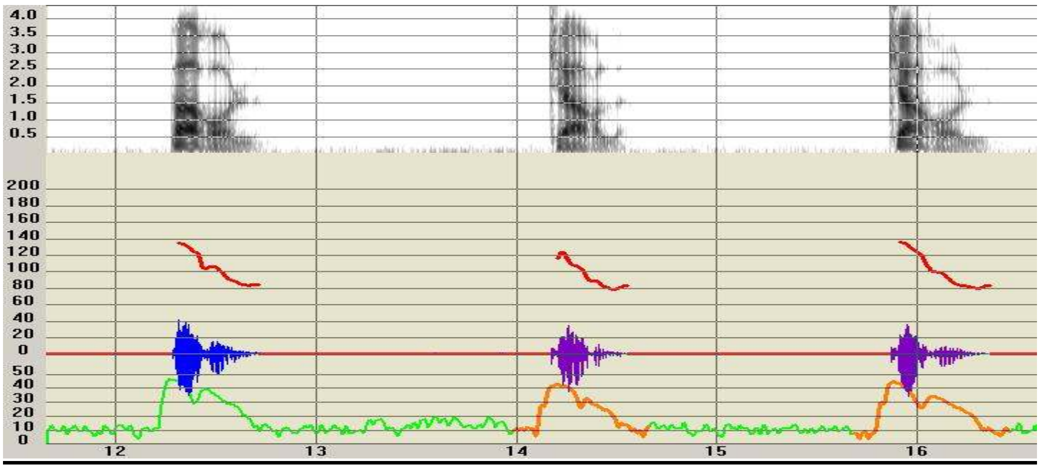


Figure 6.8: Statements

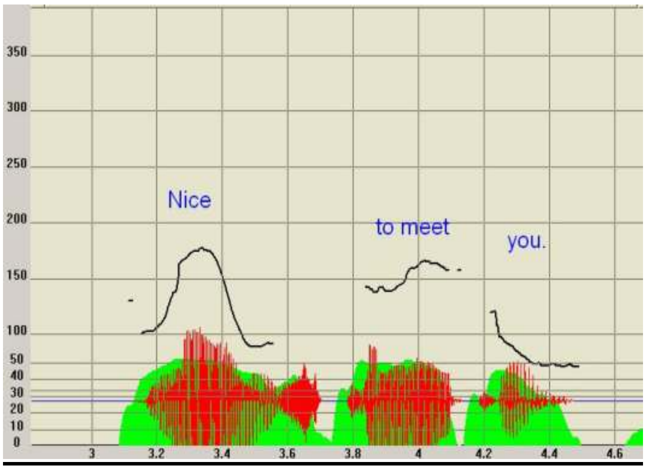


Figure 6.9: Statement: Affirmation

6.1.6. Audacity

Audacity is a free open source digital audio editor and recording Computer software application, available for Windows, Mac OS X, Linux and other operating systems. Audacity was started in May 2000 by Mazzoni and Dannenberg at Carnegie Mellon University. As of 10 October 2011, it was the 11th most popular download from Source Forge, with 76.5 million downloads. Audacity won the Source Forge 2007 and 2009 Community Choice Award for Best Project for Multimedia.

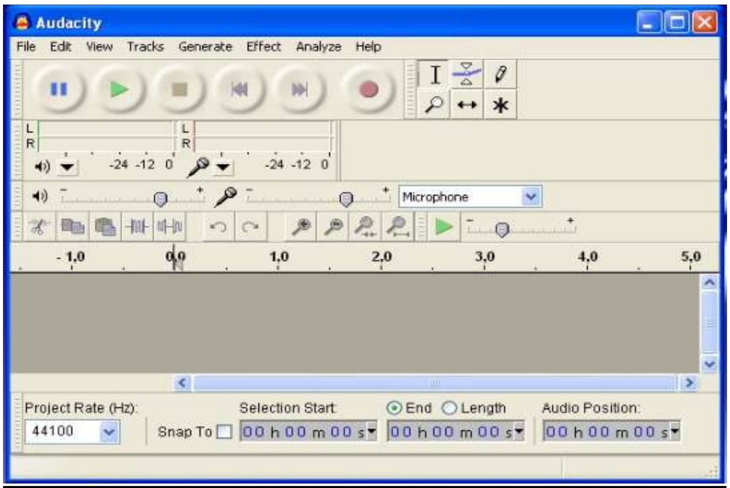


Figure 6.10: Audacity Software

6.1.7. SpeedLingua

This software efficiency may be summed up as follows:

a. **What can SpeedLingua offer for Students?**

It can offer blended learning modality in perfect harmony with the training objectives. The content is based on the most recognized language teaching frameworks. Thanks to the ease of use and the power of the revolutionary interface, specific courses can be created in minutes.

b. **Rapid Progress in Listening and Speaking**

The patented technology makes learners quickly progress on all oral skills acquisition: from listening to production. With SpeedLingua, they are more active throughout their learning experience.

c. **Unmatched Monitoring of Learner's Progress**

It comes with a flexible and easy interface to follow learner's progression: from a global overview to the smallest details.

d. **Motivated and Satisfied Learners**

Through various exercises, tests and motivation tools, the learners rapidly progress. They, thus, enjoy communicating, and they are even more motivated to continue learning.

e. **Ease of Use for All Participants**

For the learners, teachers or the training managers, the simple and intuitive interface is accessible even to the less experienced user.

Figure 6.11: Speedlingua Software

6.1.8. Speech Analyzer

In addition to the above mentioned software, Speech Analyzer is also one of the various computer programs used for acoustic analysis of speech sounds. It is distributed as freeware and easy to be downloaded.

It can be used to do the following tasks:

a. Perform fundamental frequency, spectrographic and spectral analysis, and duration measurements.
b. Add phonemic, orthographic, tone, and gloss transcriptions to phonetic transcriptions in an interlinear format.
c. Perform ethnomusicological analysis of music recordings.
d. Use slowed playback, repeat loops and overlays to assist with perception and mimicry of sounds for language learning.

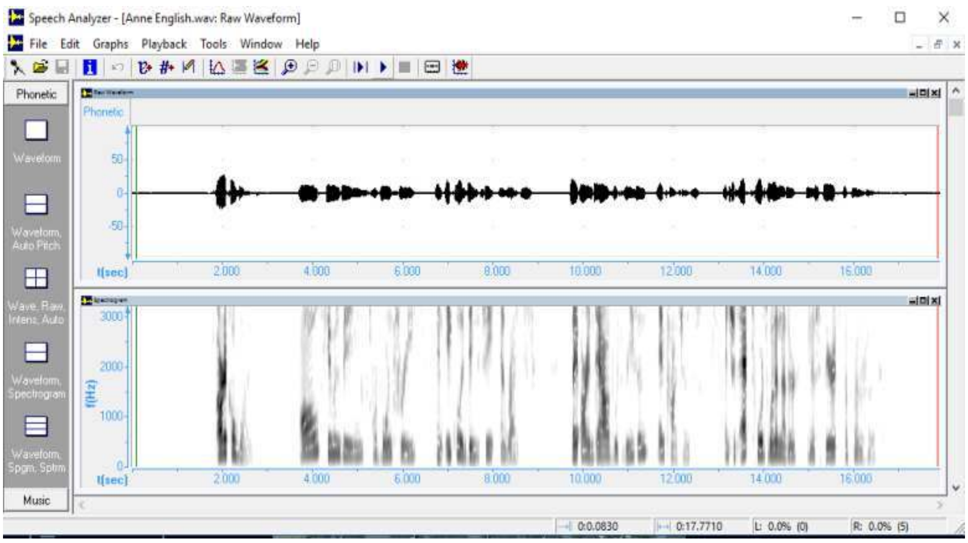


Figure 6.12: Speech Analyzer Software

6.1.9. Winsnoori

Figure 6.13: Laprie manipulating Winsnoori Software

Laprie is someone who is interested in speech analysis and is the main author of the WinSnoori software in 1987. This program offers a variety of speech analysis algorithms (pitch detection, automatic formant tracking, and copy synthesis for a parallel formant synthesizer). Tools derived from WinSnoori are now used in foreign language learning. It is an invaluable help to teach phonetics and more generally speech sciences. For several years, he has undertaken the development of the software WinSnoori that is for both speech scientists as a research tool and teachers in phonetics as an illustration tool. It consists of five types of tools by which the following may be achieved:

- To edit speech signals
- To annotate phonetically or orthographically speech signals. WinSnoori offers tools to explore annotated corpora automatically,
- To analyse speech with several spectral analyses and monitor spectral peaks along time,
- To study prosody. Besides pitch calculation it is possible to synthesise new signals by modifying the F0 curve and/or the speech rate,
- To generate parameters for the Klatt synthesiser. It is a user-friendly graphic interface together with copy synthesis tools (automatic formant tracking, automatic amplitude adjustment) allows the user to generate files for the Klatt synthesiser easily.

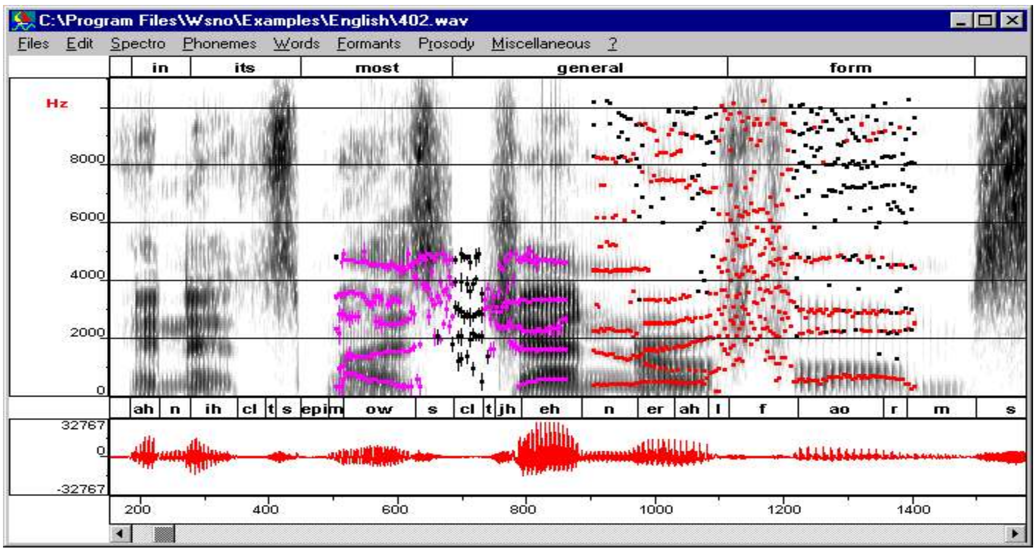


Figure 6.14: Spectrogram with Winsnoori Software

In the context of speech sciences, WinSnoori can, therefore, be exploited for many purposes, among which are illustrating speech phenomena and investigating acoustic cues of speech sounds and prosody.

6.1.10. Praat Software

Praat comes from the Dutch word for '*talk'*. It is a free scientific computer software package for the analysis of speech in Phonetics. It was designed, and continues to be developed by Boersma and Weenink, who are professors of the phonetic sciences at the University of Amsterdam. It can run on a wide range of operating systems, including various versions of Unix, Linux, Mac and Microsoft Windows (95, 98, NT4, ME, 2000, XP, Vista, 7 and 8). The program also supports Speech synthesis, including articulatory synthesis.

Praat is an excellent speech software which makes it possible to immediately visualize the EFL pronunciation teaching. Such specific values as formant, fundamental frequency, intensity and duration have been extracted for a better analysis of sound waves in respect of segmental, accentuation and intonation. Therefore, pronunciation errors are clearly identified. The application of Praat revives classroom teaching and helps improve the reliability and validity of pronunciation evaluation. A new idea for self-study and self-testing is further developed. This is mainly as students can record their attempts to match the wave form of a sound produced by a native speaker.

Praat can also read sounds recorded with the program or audio files recorded in intensity, volume and other complex details. Praat is able to isolate certain sound bites or filter frequencies either manually or using scripts. Here are some examples of spectrograms displayed using Praat software.

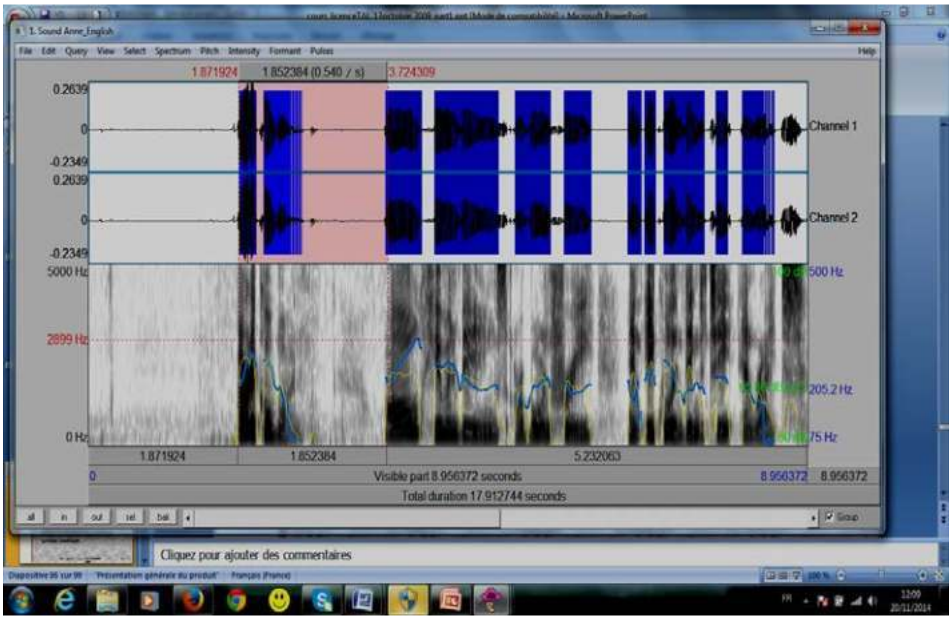


Figure 6.15: Spectrogram with Praat (Pitch and Intensity)

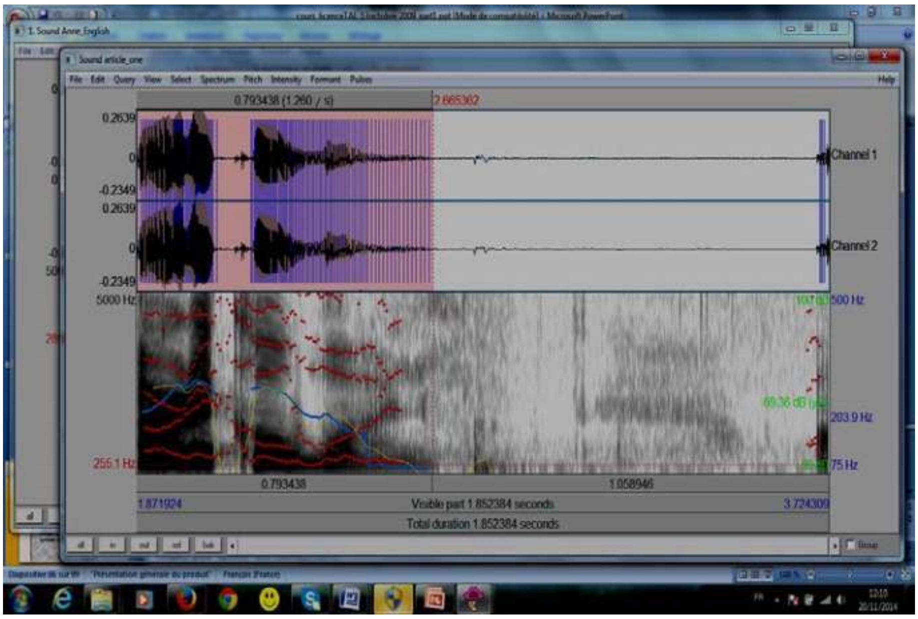


Figure 6.16: Spectrogram with Praat (Pitch, Intensity and Formants)

Conclusion

Pronunciation is one of the least taught aspects of language learning in language classrooms. Teachers wonder whether it is worth devoting class time to achieve general improvement in pronunciation. They claim, it is difficult to establish a link between pedagogic practice, learner characteristics, and achieved goals. Speech technology, fortunately, is able to fill these gaps and support the learning of pronunciation.

In this chapter, we have been through several different and various multimedia programs and software, which we highly recommend to the phonetic and oral expression teachers. Using them in their classes will allow them to better teach pronunciation and hence refine their students' performance.

PART TWO: FIELD WORK

Chapter Seven

Chapter Seven: Methodology

Introduction

This study was meant to help bridge the existing gap between the traditional method of teaching pronunciation and a new innovative one. This fieldwork introduces the idea of using a software program called Praat to help the students, who are considered as foreign learners, practice their speaking and precisely their pronunciation. In the previous chapter, we have been through a variety of software, some of which focus on the articulation of sounds, while others offer learners an opportunity to listen to pre-recorded materials. The digitized pronunciation software, such as those for pronouncing dictionary definitions, are supported with high quality sound recordings, showing speakers video clips, which permits the learners to look at the different articulatory movements used in sound production. In contrast, Praat, which is a program developed by two Dutch phoneticians, allows the user to record and analyze speech in the computer through displaying immediate spectrograms focusing on acoustic phonetic analysis, rather than providing pre-recorded examples of native speakers. Therefore, Praat allows students to analyze their own visual patterns in order to distinguish and compare how it differs from the ones of natives, considered in our case as models of both American and British English.

After the program is freely downloaded from the internet, we can use this program in several different ways such as making students practice phonology by matching sounds as the software contains tests of perception, discrimination and production or make them work on pronunciation by recording longer or shorter speech passages. Hopefully with the usage of Praat software, the students will be excited about learning pronunciation, and through meaningful practice outside of the classroom, at least, they will gain confidence when they speak English.

7.1. Research Questions

Pronunciation tends to be neglected in the department of English, Batna University, and is not given priority, though it is of a paramount importance. Interviewing a sample of teachers surprisingly revealed how much they are reluctant to teach pronunciation in classroom. There is, then, a great need not only to teach pronunciation but also to find out new methods of teaching it, which enable students to practice, develop their speaking skill and improve their pronunciation to be understood and be able to communicate effectively.

Inspired by the use of computer technology for the teaching of pronunciation using software with immediate visual feedbacks, many questions came to the researcher's mind, amongst them are the following.

a. How can innovative techniques and devices to teach pronunciation make phonetics sessions more interactive and more interesting?
b. To what extent would the use of Praat software improve students' pronunciation?

7. 2. Method

The teaching of the pronunciation of any foreign language must encompass both the segmental and suprasegmental aspects of speech. Thus, investigating the problem of mispronunciation requires the use of an experimental method. The experiment will display students' errors and through an experimental method, we can classify the errors according to their origins and correct them through practice.

The experimental method is, therefore, the most appropriate and adequate one in this work for it enables controlling all the essential factors. It is mainly used in natural and physical sciences, and recently it has been of great use in social and human sciences, and hence, in the field of education.

The experimental method is the only method of research that can truly test hypotheses concerning cause-and effect relationships. It represents the most valid approach to the solution of educational problems, both practical and theoretical, and to the advancement of education as a science (Gay1992).

Experimental methods have obtained satisfactory results when the design is carefully used, and since it includes experimentations, the researcher is supposed to obtain the same results in case the experiment is repeated.

7. 3. Subjects

The subjects who participated in this study were students enrolled in the Department of English at Batna University in the academic year 2011/2012. They studied to graduate and be awarded a BA degree after four years. Many of them aspired to become future teachers of English. The department offered four distinct proficiency levels for students from first to fourth year before the newly embraced Licence Master Doctorat (LMD) system, which reduced the number to three years. Unfortunately, each level lacked instruction in almost the five areas of language learning: reading, writing, speaking, listening, and grammar. Actually, no courses were devoted strictly to pronunciation, although some colleagues mentioned that they occasionally correct their students' pronunciation in classrooms.

7.3.1. Subjects Selection

The population of this study includes all first, second and third year students of English enrolled at Batna University during the academic year 2011/2012. Yet, it is neither possible, nor desirable to study the entire population, since, according to Deldime& Demoulin (1975); sufficient data can be obtained through the study of a proportion of the population: *a sample*.

We avoided the use of a random sampling for it is likely to make us fall onto subjects of the same features and characteristics, which are not essential in our work. Moreover, random sampling cannot be used in our work because our subjects should be selected according to certain variables like their previous grades in the modules of phonetics and oral expression, as well as their motivation to learn and so on. A systematic sampling technique in selecting learners would give us the opportunity to have a group of mixed levels. This can raise our chance to have a heterogeneous group of diverse characteristics and abilities and thus, diminish the risk of bias resulting from selecting students randomly.

The students consisted at the outset of 183 who were enrolled in first, second and third year graduate classes at the Department of English since the module of phonetics is not taught in the fourth year level. Yet, students from first and second levels were excluded from this study because not only it was felt that the students' performance was not at the level of neither fluency nor accuracy,since fluency and accuracy are complementary and considered to be as the basic principles of pronunciation, but also, because third year students had a more advanced level in phonetics as compared to the intermediate level of students along the first two years of the theoretical phonetic syllabus. In this syllabus, only simple questions were tackled, related to articulatory phonetics, including an introduction to speech organs and speech production, places and manners of articulation, primary and secondary cardinal vowels.., etc.

As almost all of these students were our students in second year level, we just emphasized on teaching the remaining suprasegmental features, for instance, sentence stress and intonation, which had not been dealt with in the two previous levels, mainly, as almost all teachers, did not have enough time to finish the theoretical program 'syllabus', needless to mention the various reasons, which were given, among which the strike movement was the most convincing and persuasive one.

Therefore, the population consisted of only 32 third year students of English at Batna University. All agreed to participate in this study. They were systematically selected according to their previously obtained grades in phonetics; which were computed ranging from 10 to 15 and the mean score was 12.5 computed as follows:

$$\frac{10+11+12+13+14+15}{6}$$

$$75 \quad = \quad \frac{12.5}{6}$$

$$\textbf{So:} \quad \mu = 12.5$$

The students whose grades were one standard deviation above the mean were selected as being intermediate students who participated in our study, where the number of female and male participants was respectively 16 aged between 20 and 21. All the participants were native speakers of Arabic except two, one is a Malian and the other is a Nigerian. None of them had travelled or studied in an English speaking country, nor had pronunciation classes before.

7. 4. Procedure

Despite the various difficulties and constraints encountered throughout this study, and thanks to the opportunity we had in teaching the same groups of students in the two modules 'oral expression and phonetics', this enabled us not only to put into practice the lectures of phonetics in the oral expression sessions, but also raised the chances of getting concrete results about students' pronunciation.

A pre-test and post-test regarding phonetics prosodic features, were given to the two groups 'control and experimental' in the form of many informal recordings. A pre-test was administered for the sake of estimating how much students knew and mastered the previously taught phonetic lectures during their first and second years; a post-test was administered in order to see to what extent did Praat software affect pronunciation teaching displaying the prosodic features of English language, namely stress and intonation, among the students.

After the administration of the pre-test, the selected sample, comprising these students with different levels, was then sub-divided randomly into a control group and an experimental group with 16 students in each.

During the first semester, the control group followed a traditional and theoretical course on the fundamentals of English prosody during regular scheduled sessions, using some books from the library as a reference; on the other hand, the experimental group had lectures divided into three: Presentation, practice and production dealing with the review of the suprasegmental features and then practiced them through the use of Praat software installed on their computers and at last they were formally recorded.

While the majority of subjects expressed a high degree of interest in improving their pronunciation of English, the students had no prior experience studying pronunciation in a classroom setting, using authentic materials.

7.5. Instruments/Data Collection

The results of our work were obtained from various ways of collecting data about the students' pronunciation which included: a questionnaire and recordings of students' pronunciation, which were then analysed. The questionnaire helped us collect students 'own perceptions and opinions about pronunciation and ways of teaching it. It helped us to know the origins of mispronunciation. The questionnaire was given to students at the end of the academic year 2011.

Recording students helped us not only to obtain an immediate feedback of their speech thanks to the spectrograms, measuring their pronunciation but also to evaluate the students' improvement after exercising with the Praat software. Furthermore, the recording enabled the students -mainly those having difficulties- to improve their pronunciation.

These two tools were especially selected for their usefulness, their practical use and for they complement each other. The questionnaire helped us to assess the students' perceptions of pronunciation skills and the second enabled us to analyse the students' productions.

To support our data collection, some teachers, mainly those of oral expression and phonetics were interviewed and recorded in order to explore in depth their attitudes concerning pronunciation.

7.5.1. Questionnaire

The aim of this work is to find out how Praat software can enable students' refine their pronunciation and to investigate the students' self-perceptions and attitudes towards pronunciation and its teaching. Questionnaires are one of the most widely used research methods to investigate phenomena in language teaching and learning due to a variety of factors mainly as they provide a substantial amount of data and require the least amount of resources in terms of time and money. In addition to that, participants feel at ease in answering questions since they can fill out the questionnaire at their own pace.

We have developed a questionnaire, which was completely anonymous. After explaining the purpose of the questionnaire, it was handed in to all the students and not mailed. The questionnaire consists of 21 items. There are some basic questions about the students' personal information including gender and age. Other questions which aim to assess the students' perceptions of pronunciation in general and their attitudes towards their own pronunciation in particular. 60 out of 85 copies were retrieved.

7.5.2. Interview

In addition to the questionnaire that was especially for students, and in order to support the data collected, another instrument was used in our work, which is the interview. It was conducted in our office using a *Digital Voice Recorder* named *OLYMPUS* for it offers a very good quality of voice recording. The participants were some of our colleagues who are all full time teachers, holders of an MA degree, mainly those teaching oral expression and phonetics, in order to probe their ideas and explore their point of views and practices with regard to pronunciation.

7.5.3. Corpus

During the whole year, we attached more importance to how to design pronunciation activities for students by presenting and practicing with a great deal of emphasis placed on different prosodic aspects using a set of samples of recorded native speakers based on the notion that they cannot pronounce what they cannot hear, and, hence, listening would come before production. Some of the examples include, for instance, one of the speeches of George W. Bush about Iraq for the sake of explaining the use of spectrograms which clearly shows the different pitch contours with the different peaks. This enabled the students to match the theoretical lectures of intonation using practice (see Spect. below).

In order to gather information about the students' pronunciation, a sentence adapted from the declaration of the human rights, article one, was used as a test sentence though it is a statement which could then reflect different types of intonational behaviour in case pronounced differently.

> *All human beings are born free and equal in dignity and rights. They are endowed with reason and conscience and should act towards one another in a spirit of brotherhood***.**

The article is transcribed as follows:

British: |ɔːlˈhjuːmənˈbiːɪŋz ə bɔːn friː ənd ˈiːkwəl ɪnˈdɪgnɪti ənd raɪts | ˈðeɪ ər ɪnˈdaʊd wɪð ˈriːzən ənd ˈkɒnʃəns ənd ʃəd ækt təˈwɔːdz wʌnəˈnʌðər ɪn əˈspɪrɪt əv ˈbrʌðəhʊd |

American: |ˈɒlˈ hjuːmən ˈbiːɪŋz ər ˈbɔːrn ˈfriː ənd ˈiːkwəl ɪn ˈdɪgnəţi əndˈraɪts | ˈðeɪ ər enˈdaʊd wɪθ ˈriːzən ənd ˈkɑːnʃəns ənd ʃəd ˈækt təˈwɔːrdz wʌnəˈnʌðr̩ ɪnəˈspɪrət əvˈbrʌðərˌhʊd |

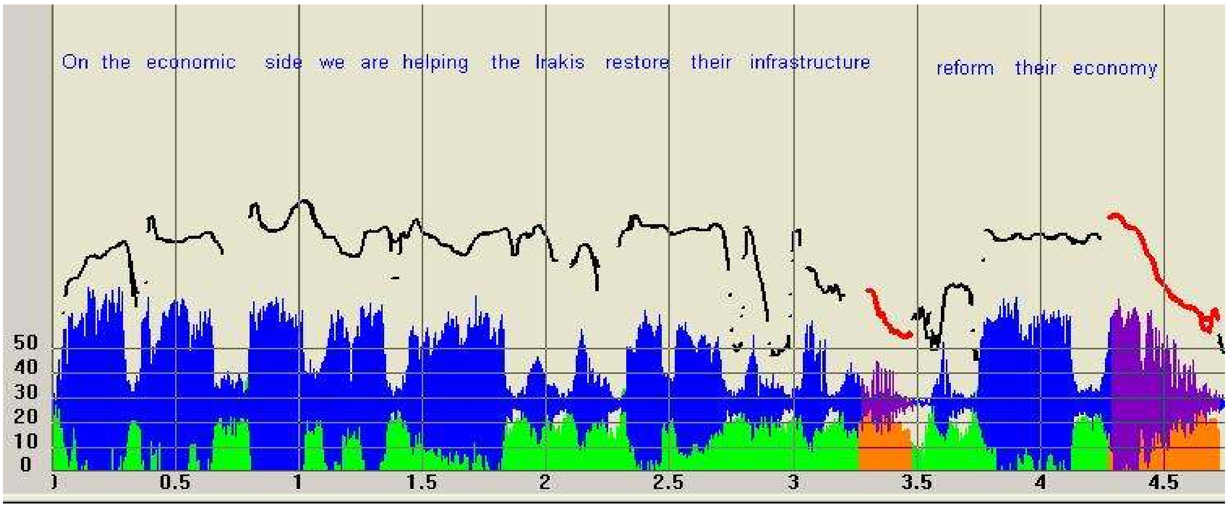


Figure 7.1: Spectrogram of George W.Bush with Winpitch Software

7.5.4. Recording

Recording students is our main pedagogical tool for our pedagogical work. It is hence a good methodological practice used to analyse and interpret the students' pronunciation and improving the reliability of research.

We took advantage of the fact that we taught students both phonetics and oral expression to put into practice the lectures of phonetics in the oral expression module. Therefore, before we started to record the students' pronunciation, we introduced Praat software which is free of charge and can be downloaded from the following web page: http://www.fon.hum.uva.nl/praat/ and using the next link:

http://www.fon.hum.uva.nl/praat/download_win.html.

The program is small and very useful to put on a small flash drive in case you have no internet access and want to work at home or work elsewhere if you do not have a laptop.

Then, we explained the process of recording with illustrations and used many informal recordings for the sake of training students to get familiar with the recording process. After that, we invited the students participating in this work to our office, recorded their pronunciation one by one by using our laptop and a microphone, which is a device designed to capture the sound waves, catch the pressure fluctuations and transform them into an electrical signal.

7.5.4.1. Procedure of Recording

In order to record any speech sound in *Praat*, we just follow the steps in order. Images are used to help memorizing the recording process.

i) We double click on *Praat* after extracting it; a small icon will be displayed on the desktop, and two windows will pop up. One is *Praat* objects, which corresponds to where the files we have opened in *Praat* appear, including sound files and TextGrids. The other is Praat picture, which is used for plotting graphs (see below).

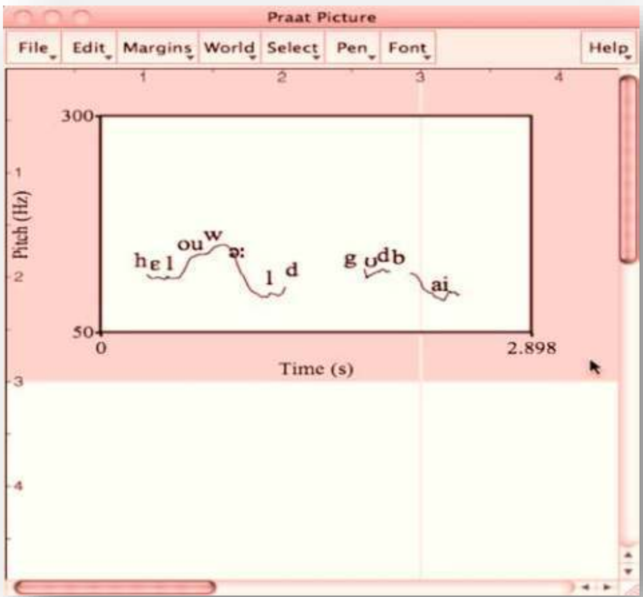


Figure7.2: Praat Picture

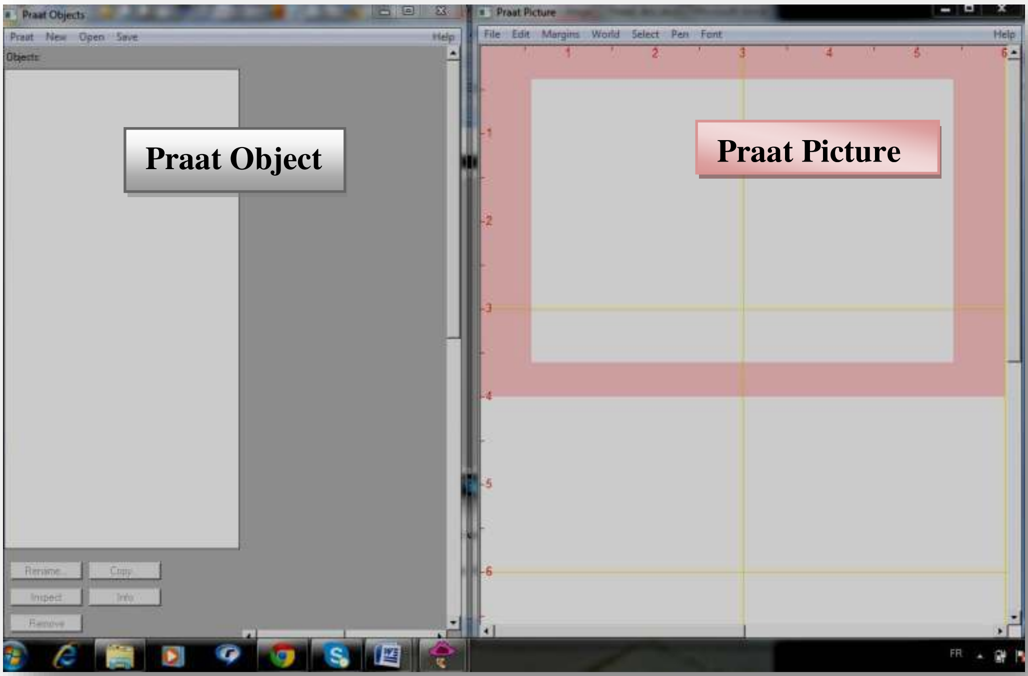


Figure 7.3: Praat Object and Praat Picture

ii) From the *objects* window, we select *New* then *Record mono sound* (for most speech recordings it is not necessary to record stereo)

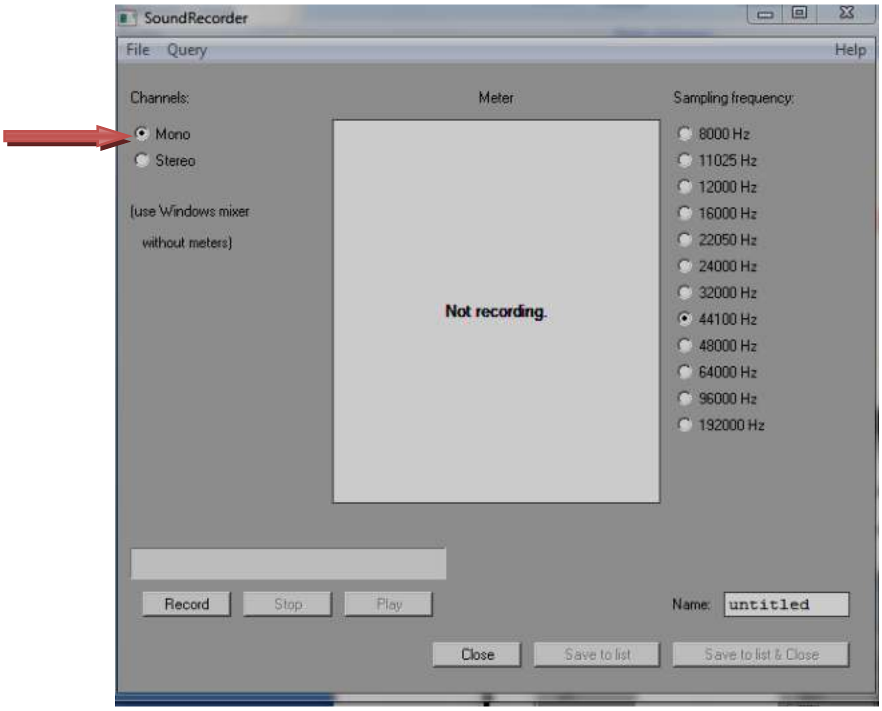


Figure7.4: Sound Recorder (Mono or Stereo)

iii) Now we have to decide about the sampling frequency, which affects the recording quality. For speech it is enough to choose a sampling frequency of *22050 Hz* though it is set to *44100 Hz* as default.

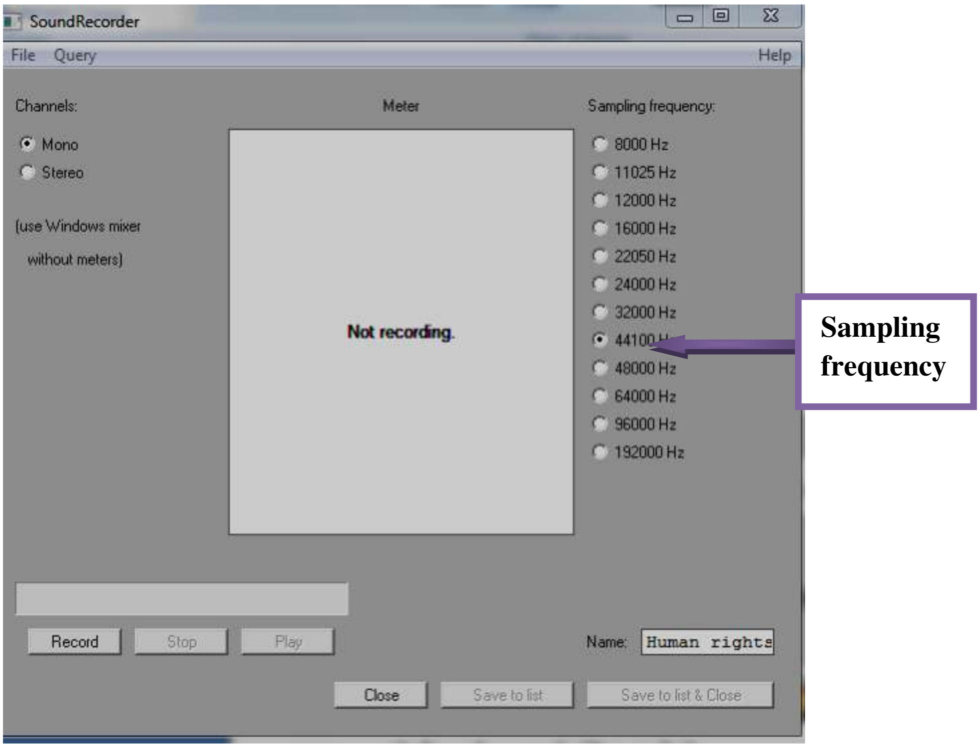


Figure 7.5: Sampling Frequency

iv) We click on *Record* in order to start recording (as shown above). We utter the sentence "*All human beings are born free and equal in dignity and rights. They are endowed with reason and conscience and should act towards one another in a spirit of brotherhood*[3]" into the microphone.

v) We *play* the sample again. In case we cannot hear anything, we just make sure that the microphone is properly connected and check its settings by clicking on Volume Control.

[3]The Universal Declaration of Human Rights: Article one

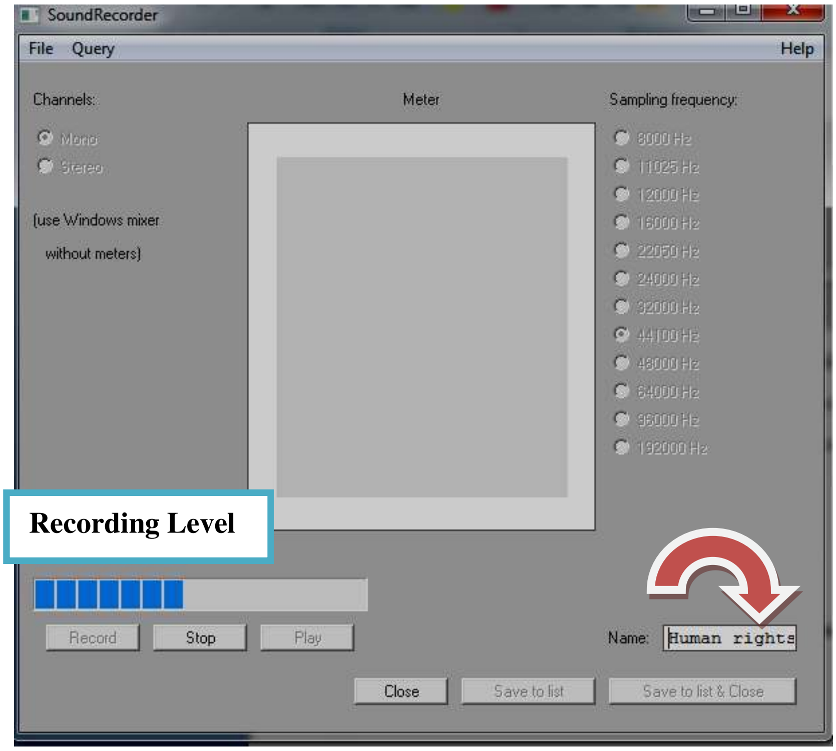


Figure 7.6: Recording Level

vi) After having recorded and played the sample again, we need to give it a name, by typing it in the box of the right corner found below. For instance, *Human Rights*.

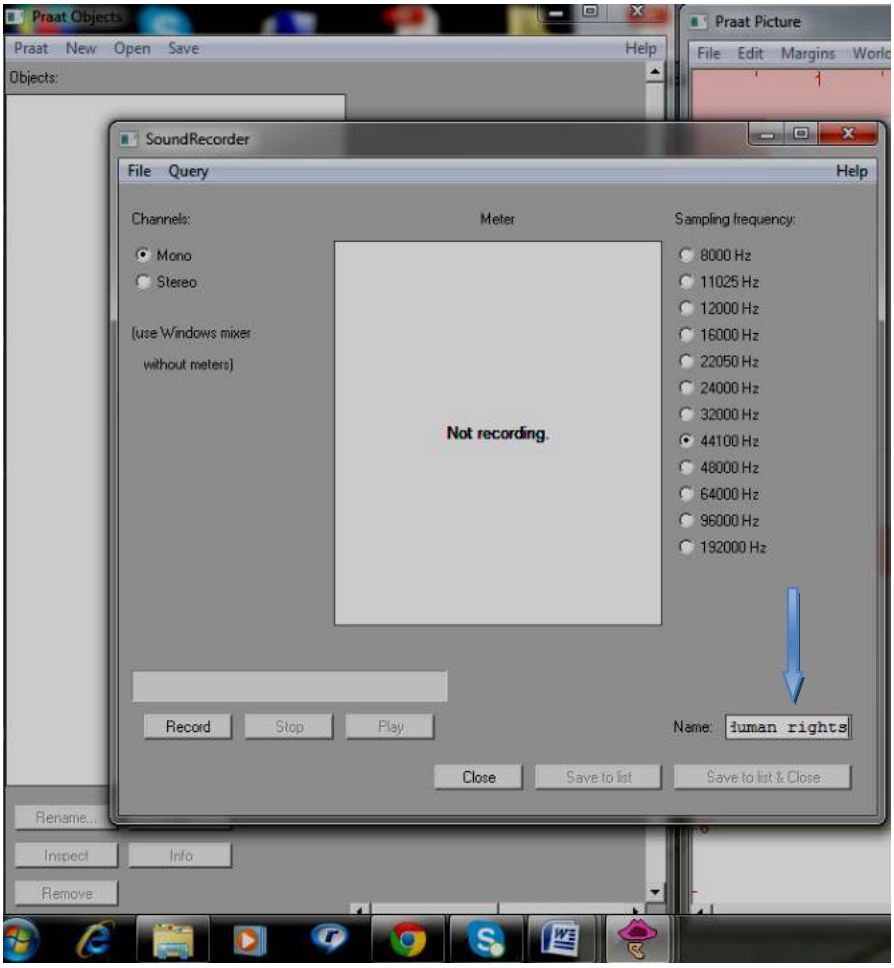


Figure 7.7: Giving a Name to a Sound.

vii) We save the sample by clicking on *Save to list.* It will immediately appear in the *Praat Objects* window which indicates the file where the sound is recorded.

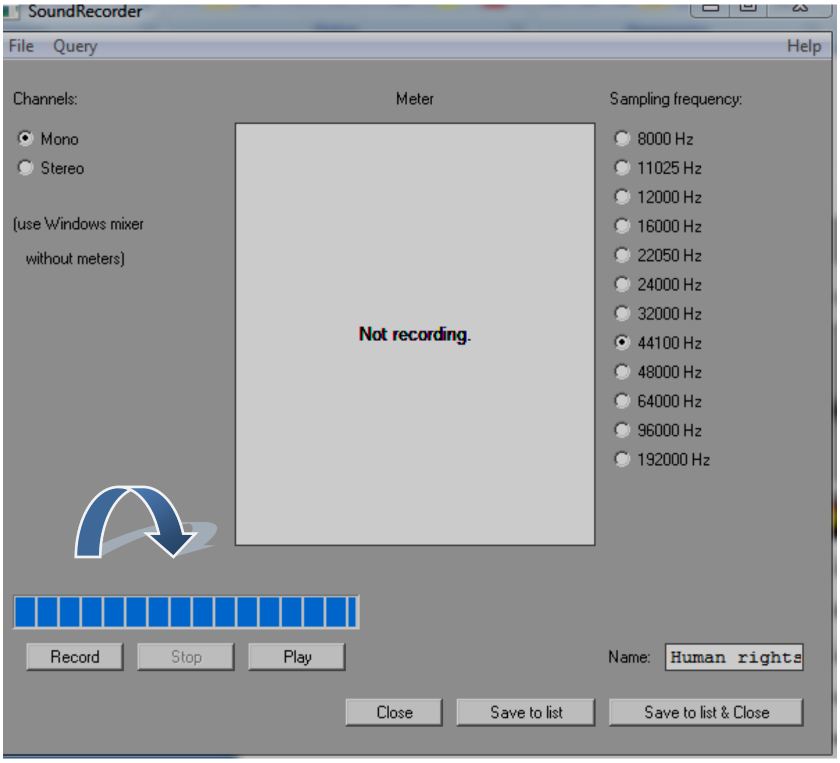


Figure 7.8: Saving File to List

viii) Now we have the speech sample in the objects window as shown below. However, if we quit Praat, we will lose the sample. In order to prevent this, we click on the sample in the objects window and go to '*Save*', then Save *WAV file* (as shown above); it is possible to save the file also with other formats as shown in the image below. From here we can save the speech sample as a *.wav file* to the computer.

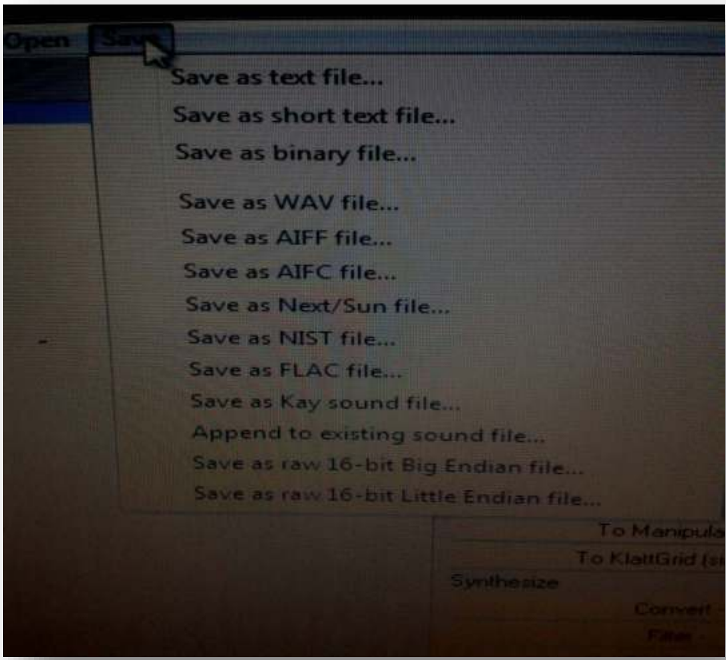


Figure 7.9: Saving Procedure

ix) Now we can open the speech file at any time by going to *Read* then *Read from file* and browsing to the location of the file (we can make a folder where we save all the speech files). Another option is that From *Read* we can of course also open other speech samples that you have in your computer. As shown in the following window:

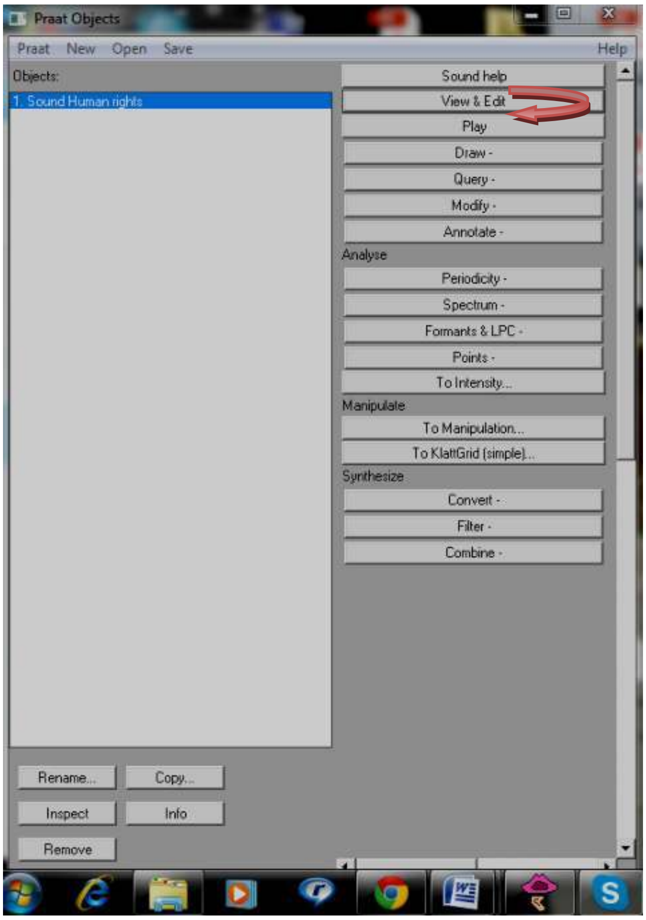


Figure 7.10: Read from File

x) After that, we click on the *View & Edit* button in order to see the waveform of what we have just been recording. The following window will appear.

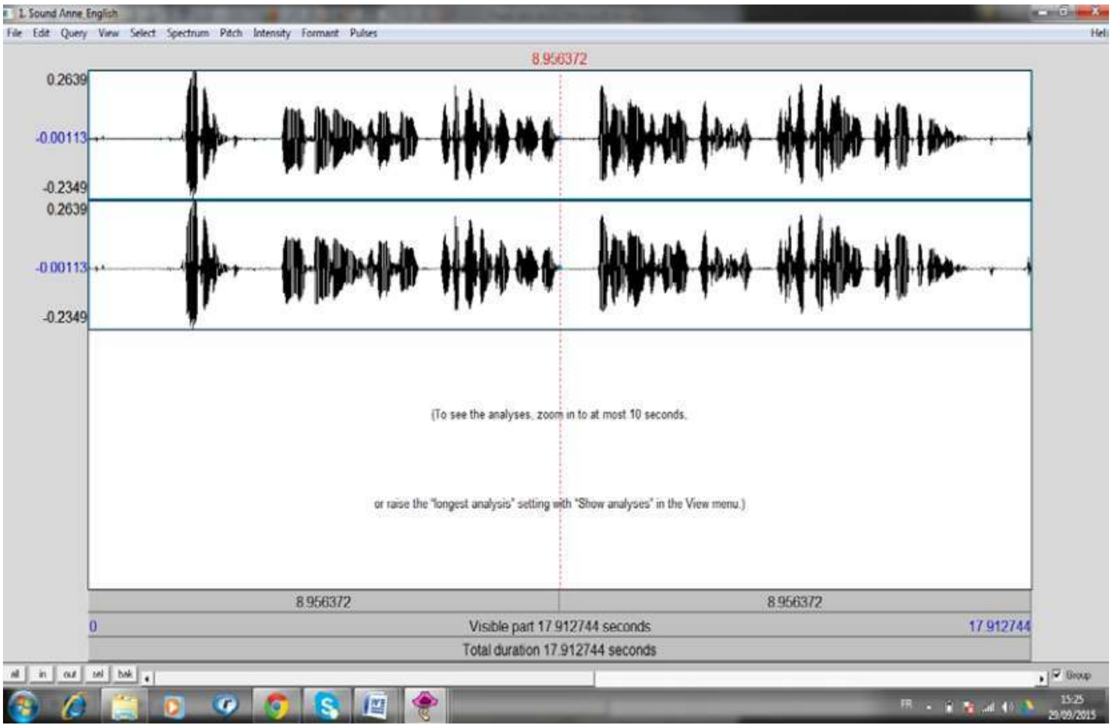


Figure7.11: View & Edit

xi) And to simplify the display, the followings should be selected (if they are not already set by default).

- *Spectrum* then select *Show spectrogram*
- *Pitch* then select *Show pitch*
- *Intensity* then select *Show intensity*
- *Formants* then select *Show formants*
- *Pulses* then select *Show pulses*

Here is a window of our spectrogram after having selected the aforementioned buttons:

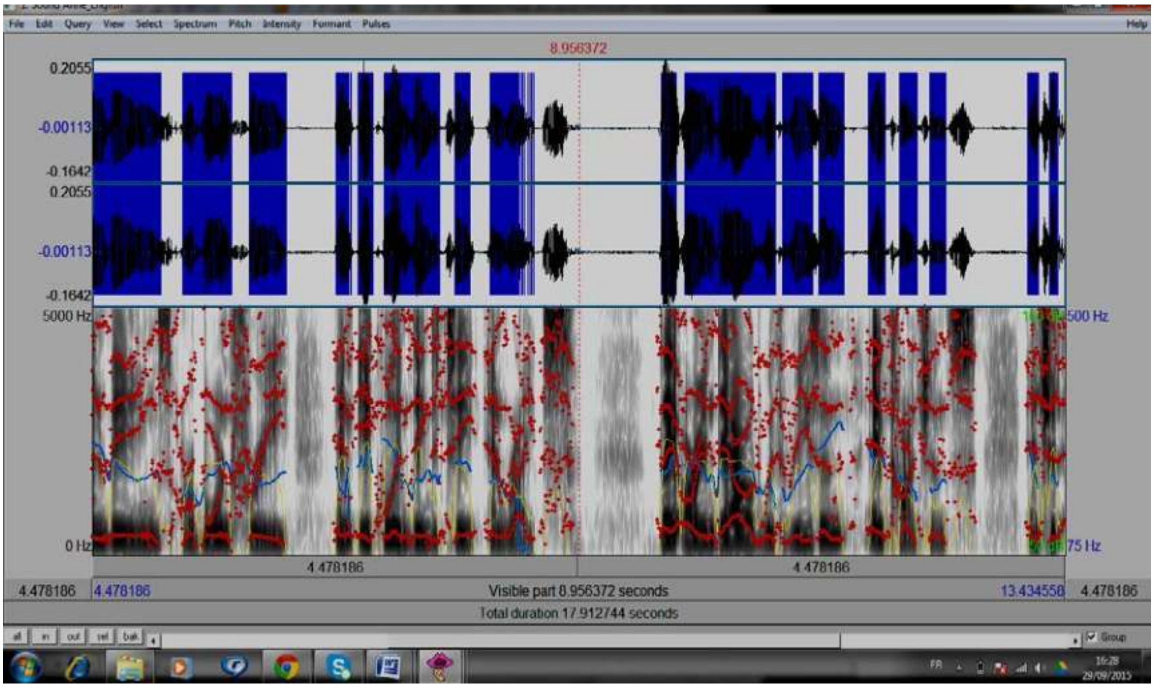


Figure 7.12: Spectrogram Appearance

- The **dark blue** colour represents the *pulses (a waveform)*
- The **red** colour represents the *formants*
- The **blue lines** represent the *pitch*
- The **yellow lines** represent the *intensity* **dB**
- The horizontal axis represents the *time* in seconds
- The vertical axis represents the *frequency* in Hertz Hz

7.6. Segmentation and Annotation

It is important to know that *segmentation* is a word which is derived from *segments* which are *vowels* and *consonants*. The segmentation then is to separate words acoustically by setting marks or borders in the speech signal by establishing a link between the spelling words *orthographic words* and their phonetic counterpart, which is referred to as *annotation*.

Therefore, annotation generally includes describing, classifying and organizing speech material by systematically adding symbolic labels to its parts.

Praat enables us to both see and play the speech signal recorded and the phonetic transcription in the form of tiers, where words are displayed and separated -as we have already mentioned- by markers. The latter can be easily moved to the right by using the mouse. The object type that handles annotation is the *TextGrid*. A TextGrid is a collection of *tiers* (this rhymes with *cheers*, not with *liars*) while a tier is an ordered sequence of texts, each of which is connected to a point in time or to a stretch of time, as explained below.

7.6.1. Procedure of Segmentation and Annotation

In the written form of the language, we have a set of punctuation to distinguish between the different parts of speech. Yet in speaking, we have pauses as a counterpart of punctuation. The speech itself is a continuous stream of ideas without any borderlines set between words. As mentioned earlier, Praat enables us to quantify the duration of not only the entire speech but also the one of separated words and segments. In order to know the way segmentation is done, we will explain the different steps to follow with several image illustrations.

- We first double click on *Praat* icon found on the computer's desktop. We will have two windows opened; the one to the left is the *Praat Objects* and on the right is the *Praat Picture* (steps mentioned earlier).
- We click then on the *Open* button in the Menu. We will have plenty of options: (*Read from file*, *open long file*, *read separate channels from sound file*, *read from special sound file*...etc.) Look at the figure below.

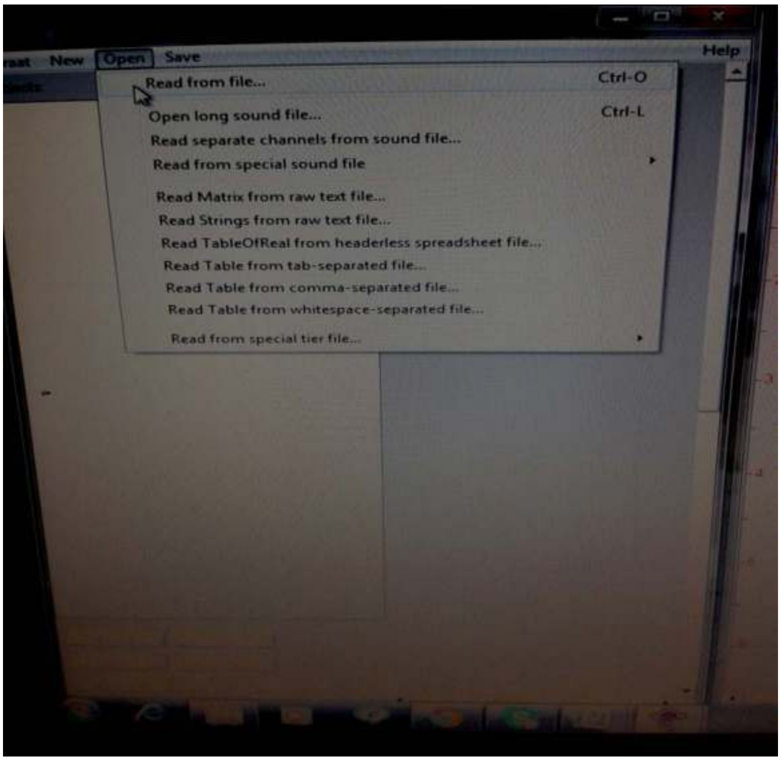


Figure7.13: Opening a Recorded and Saved File

- We click on *Read from file* or the keyboard shortcut which is *Ctrl-O*
- A window will open leading us to the path where the sound file is stored and will be opened.

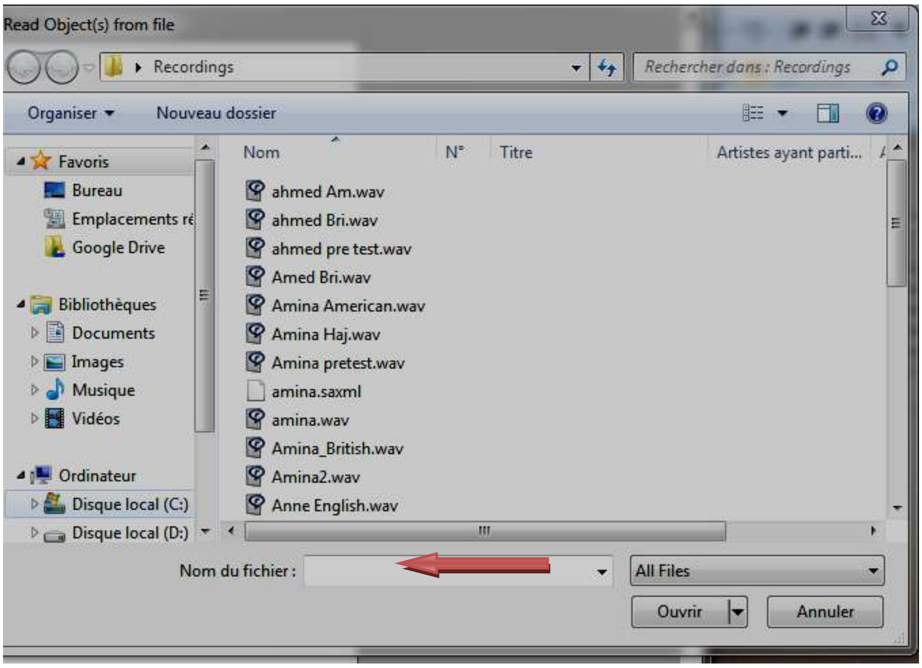


Figure 7.14: Path for File Saving

- We click after that on the sound file as we have given it a name. For us it is *Anne English.wav* –as shown with the red arrow above.
- A window will open in the *Praat Objects*. It includes the sound on the left selected in blue and many other options on the right side. For instance: *Sound help*, *View & Edit*, *Play*, *Draw*, *Query*, *Modify*, *Annotate*, *Periodicity*, *Spectrum*, *Formants &LPC*, *Points*, *To intensity*…etc. (As in the image below).

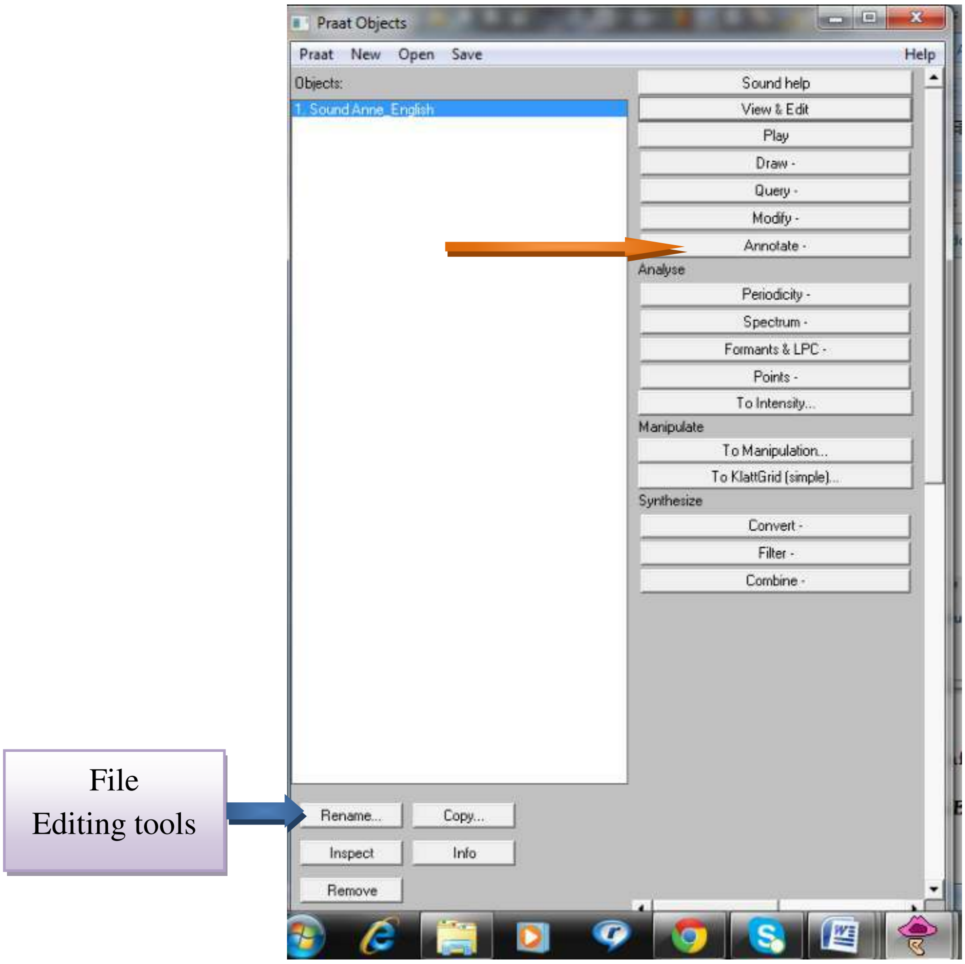


Figure 7.15: Annotate and File Editing Tools

- After that we click on *Annotate*, then *to TextGrid. Then we will* get a dialogue box in order to name the tiers and decide which of the tiers are point tiers

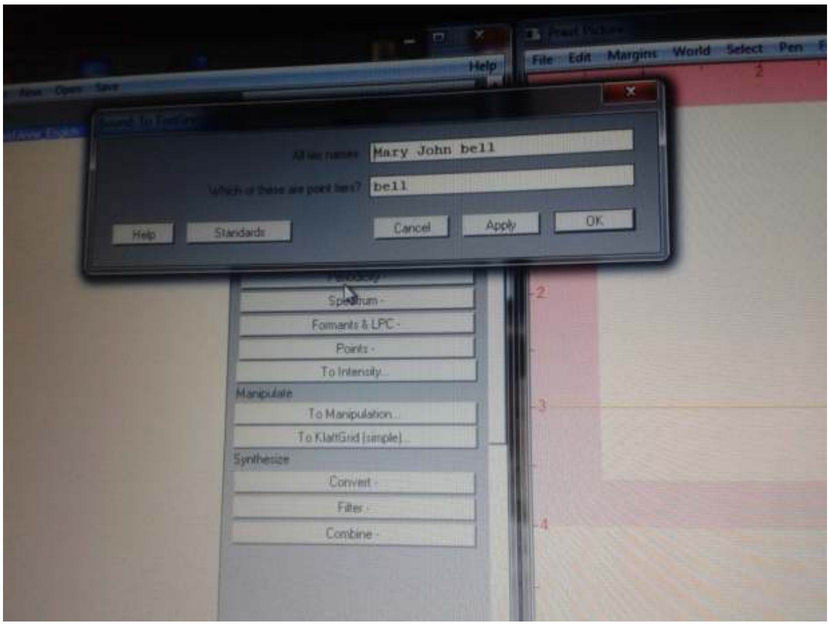


Figure 7.16: Creating Tiers

We need to create two tiers named *Anne* for this will be enough and suitable for our purpose to adjust a phonetic transcription to the adequate sounds and words. When we create a TextGrid, we need to specify the names of layers or tiers depending on what we want to segment; for instance words and phonemes. Hence in our case, we specified *Anne* for *All tier names* and left the '*which of these are point tiers'* just empty.

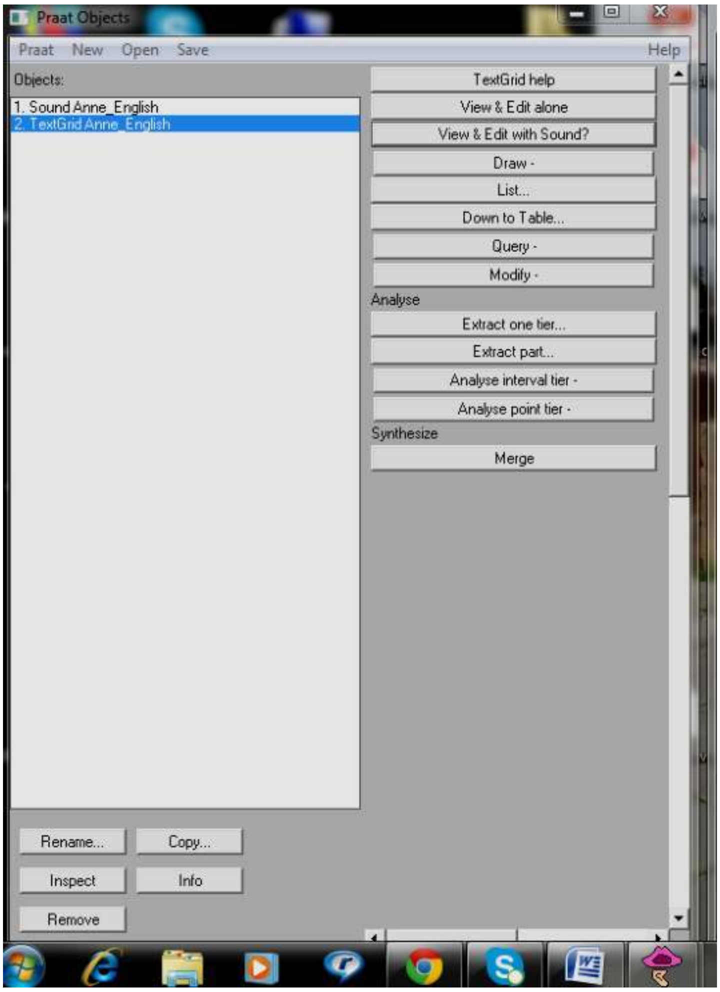


Figure7.17: Creating a TextGrid

We open then the sound file named *Anne* and *TextGrid* together by highlighting both. Then we click on *View & Edit* in order to open both files in the *Edit window.*

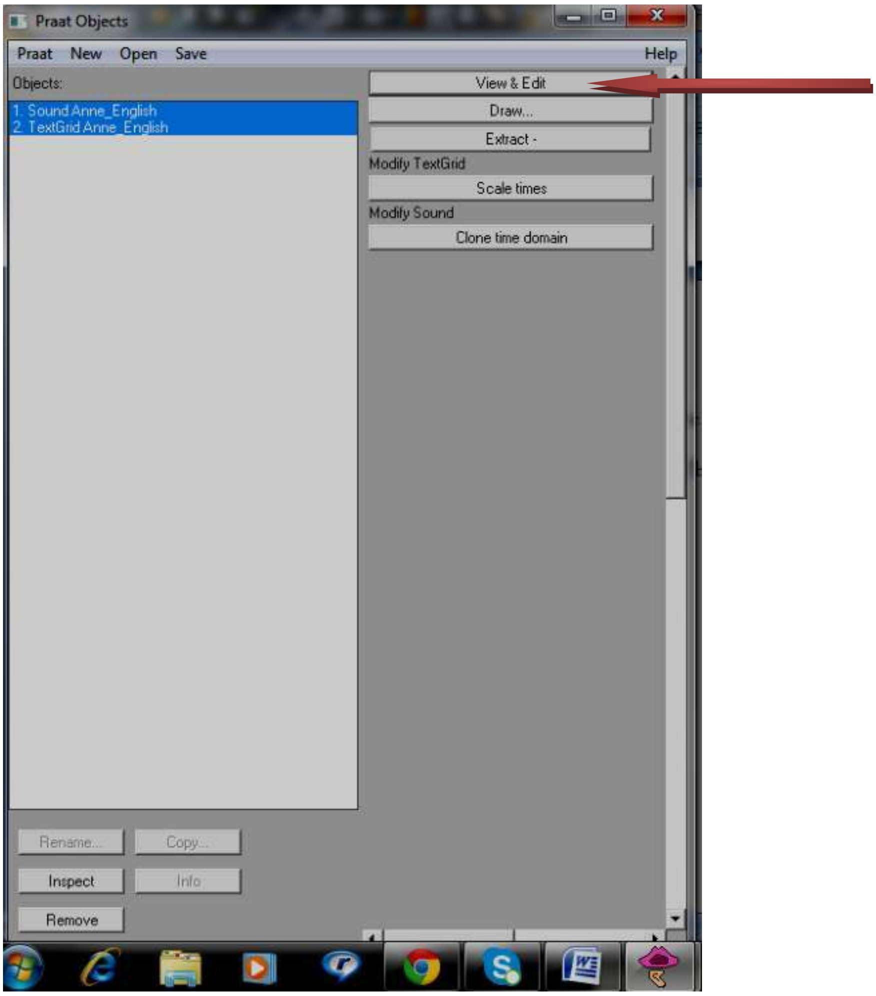


Figure 7.18: Selecting Both Files

What appears on the screen is in the picture below. We can see now the waveform situated at the top, the spectrogram in the middle and the *TextGrid* at the bottom, which correspond to our sound file we named *Anne English*. Once again, we can rename our file. *See File Editing tools.*

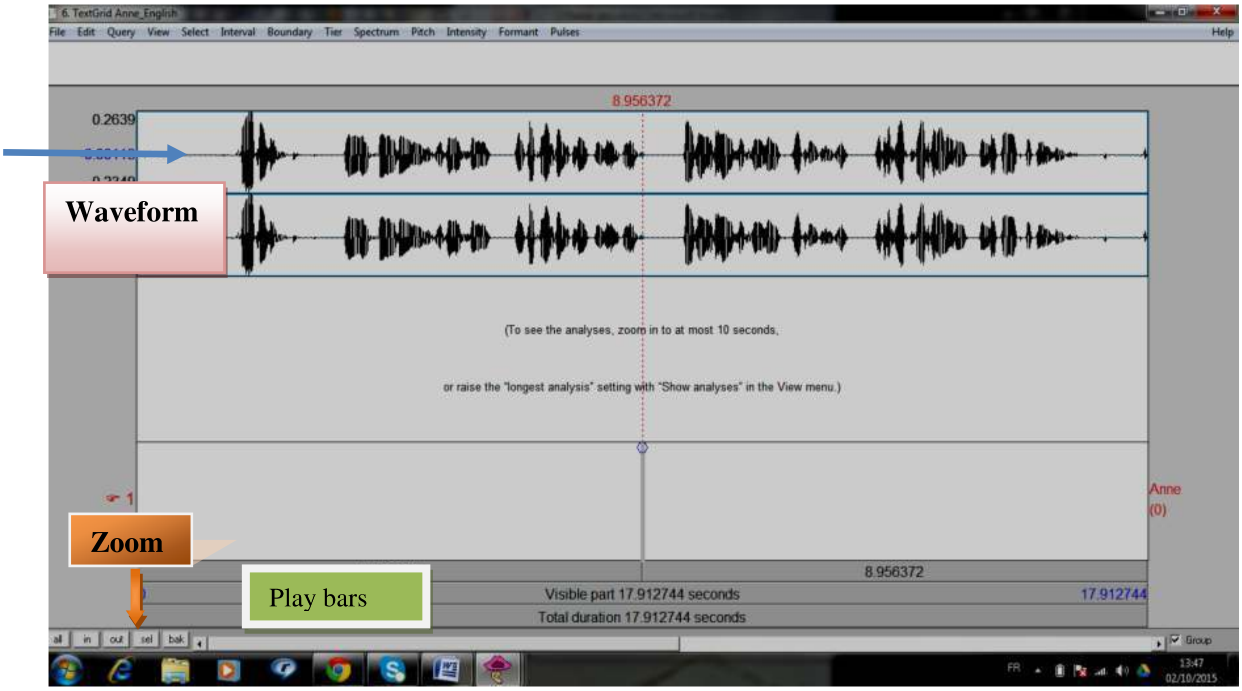


Figure 7.19: Waveform, Play and Zoom Bars.

The recorded sound file can be played from different options, whether by clicking on the *View* button found on the menu bar; then we hit the *play* or *play window*, or simply click on the *play bar* at the bottom (see the figure above).

Since the sound file is somewhat long, and we need to make use of the five Zoom options found on the left at the bottom, we have the following options named from right to left *Back, select, out, in* and *all.*

When we need to see any portion of any sound file we zoom into a specific portion of the sound file, but before that we select it first by clicking on *select in* the zoom options panel on the left at the bottom (see the picture above). We can also use the *in* and *out* buttons whether to *zoom in* or *zoom out* within a sound file at the center point of the window whatever the position of the cursor is.

The *TextGrid Editor* shows a spectrogram with various options; for instance: *pitch contour*, *formants*, *intensity*, *time* and *frequency*.

- To create a new boundary or point in a tier, we listen to the sound file and look at the spectrogram because this would give us hints asto where each selected portion starts and ends.
- We click inside the cursor circle in the same tier, and a boundary line will show up. Other shortcuts can be used for the same operation by choosing one of the commands in the *Boundary/Point* menu to insert a boundary at the cursor time on the selected tier (shortcut: *Enter*) or on any tier (shortcuts: *Command-F1* through *Command-F9*).
- To create a new boundary or point in a tier, we click inside the cursor circle in that tier; the original text in the interval that is split is, then, divided up between the two resulting intervals, depending on the position of the text cursor in the text window. When we click on the blue circle, a second boundary appears, and the text '*human*' moves to the left of that boundary, because the text cursor in the text field is after the last letter of '*human'* (see picture below).

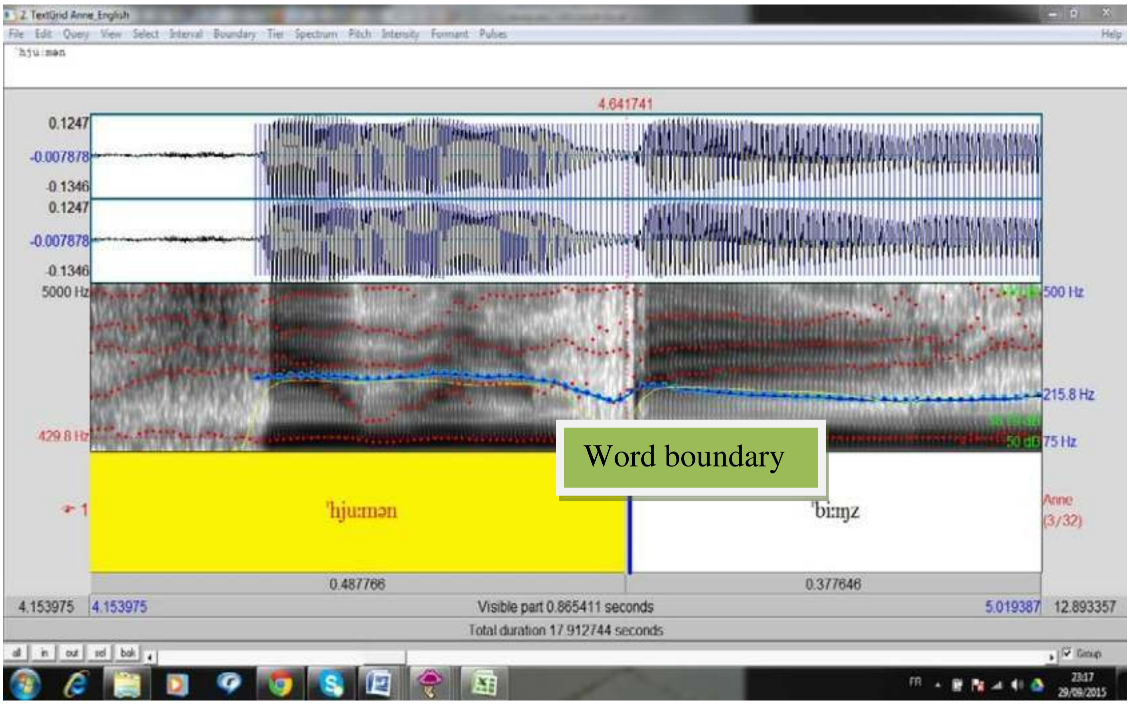


Figure 7.20: Word Boundaries

- The figure below shows that the text that we type will show up in the text field above the waveform as well as in the selected interval using a phonetic transcription.
- After creating the *TextGrid* matching sounds with their phonetic symbols, we need to save the written *TextGrid* to disk using the buttons *Save as text file* or *Save as short text file***.** Once it is saved, we can read it later into Praat using the same button introduced before which is *Read from file*.

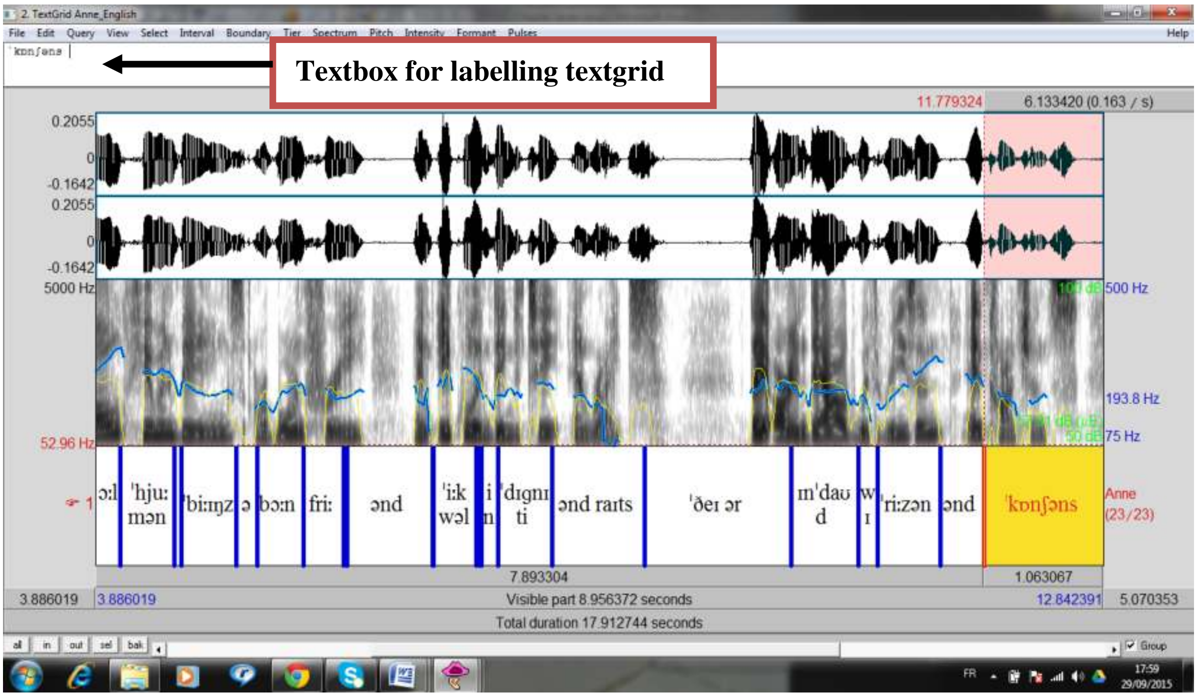


Figure 7.21: Textbox for Labelling Textgrid

- **Selecting a Boundary:** We can select a boundary by clicking in its vicinity (yellow part on the screen). The cursor will move to the exact end time of the interval labeled *conscience* 1.063067seconds, such precision is really perfect as it is very exact (see figure above).
- **Adding Two Boundaries at a Time:** In the figure we have created the interval by selecting this whole sentence in the spectrogram and then pressing the *Enter* key. This adds two boundaries: one at the start of the selection and one at the end of the selection.

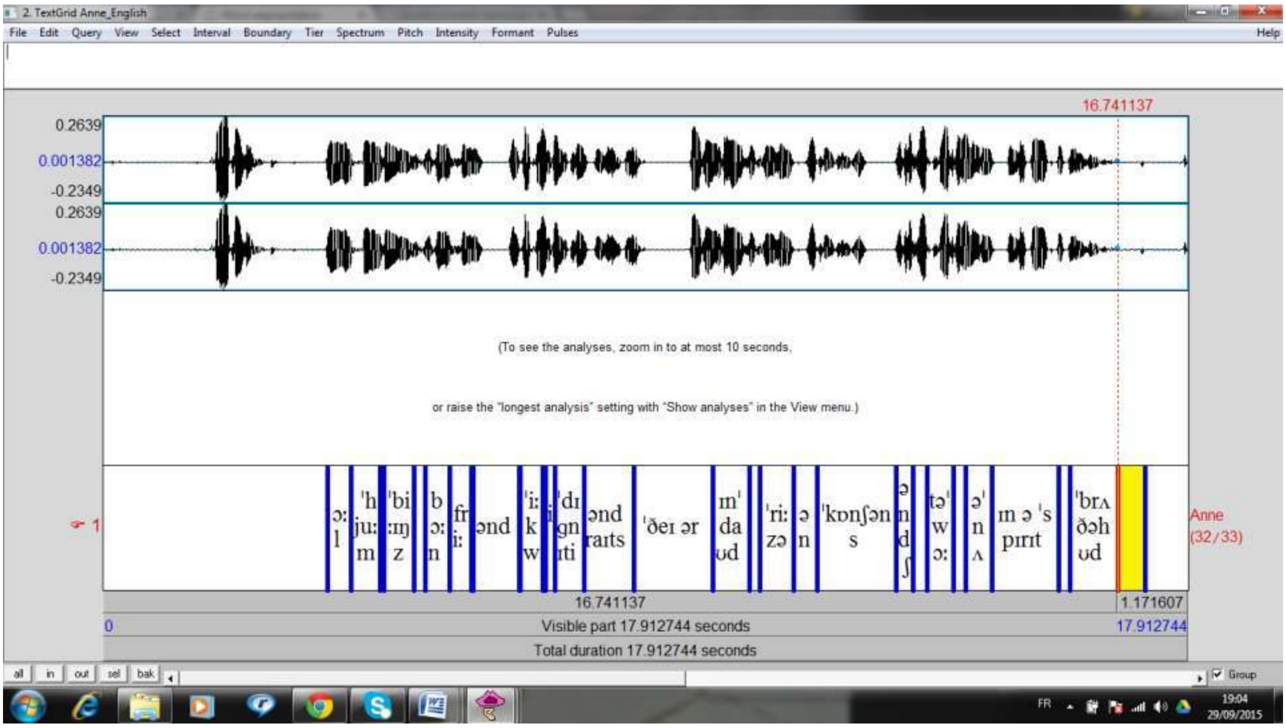


Figure 7.22: Entire Test Sentence Annotated and Segmented

In the figure above, we can see the entire sentence annotated and segmented.

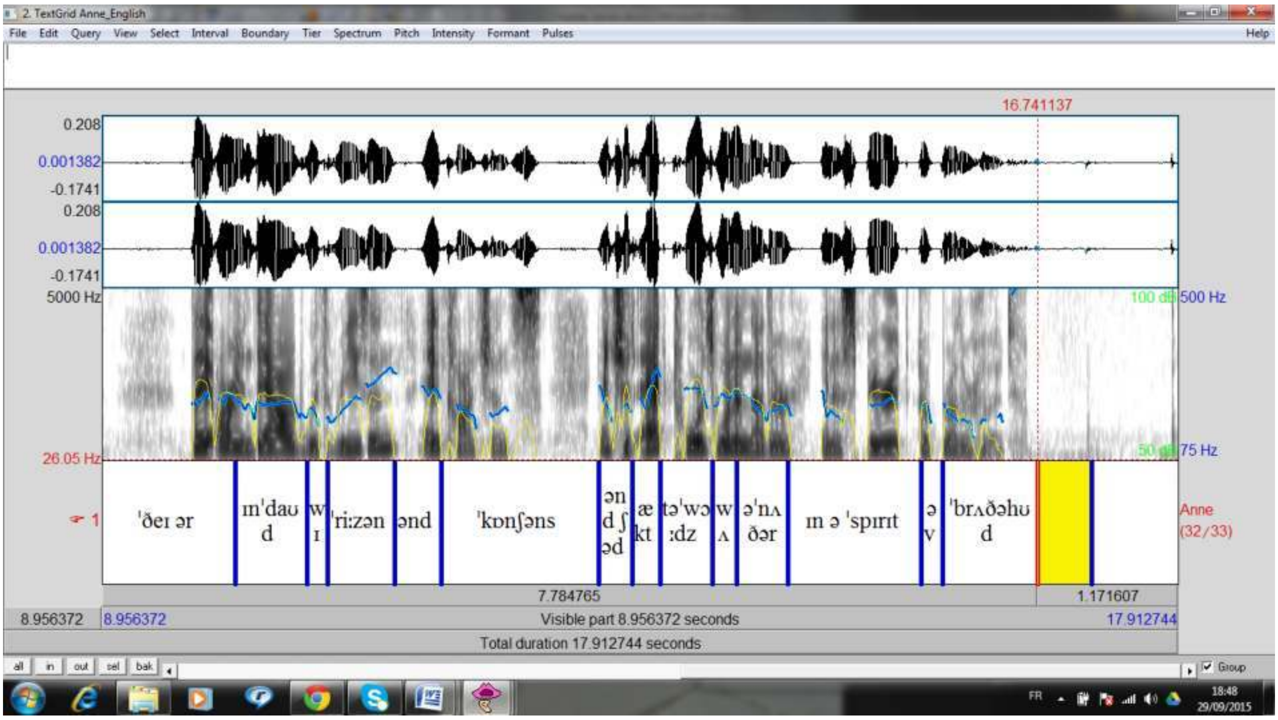


Figure 7.23: Part of the Test Sentence Annotated and Segmented

May be one of the important windows for those who want to know more about Praat is the manual window. When you click on '*Intro*' from the '*Help menu*', the following window will appear:

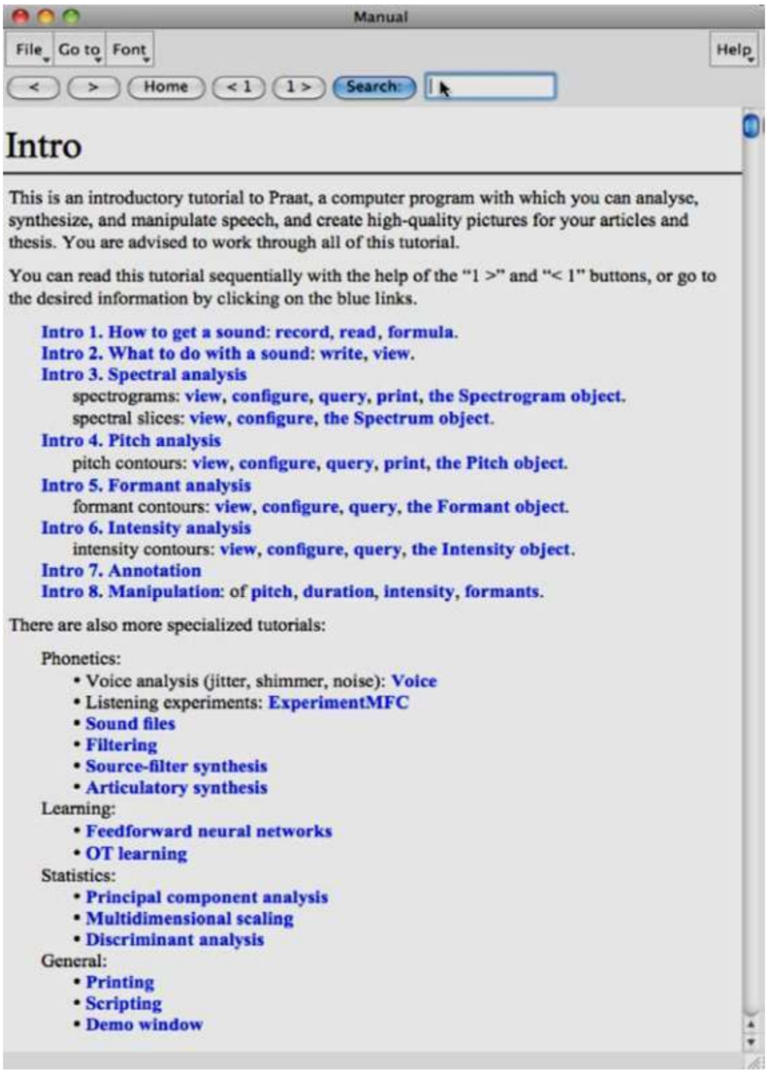


Figure 7.24: Praat Help Manual

Conclusion

In this chapter, we have been through a detailed overview of the way this work has been conducted including the experimental method that was utilized, the research questions, the population meant by in this study, the different instruments used to gather data, with explanations and demonstrations of both the recording and the segmentation of the speech using Praat software.

Results and findings from recordings, teachers' interviews, students' questionnaires and classroom observations will be revealed in the next and final chapter.

Chapter Eight

Chapter Eight: Data Analysis

Introduction

As data gathering tools for our analysis, we used various means. We administered questionnaires to students, conducted interviews with teachers, recorded students' pronunciations and compared students' grades, performances and attitudes.

The analysis and discussion of the data gathered by these different instruments are presented in the students' responses (questionnaire), the teachers' responses (interview) and the analysis of the corpus (recordings).

The questionnaire and the interview included 60 third year students and 5full time teachers. The responses are discussed under categories, where every category encompasses responses about items.The undertaken experiment involved sixteen students, using the Praat software.

The amount of the data collected in this study is qualitative; hence, the students' recordings, the teachers' interviews transcripts as well as the open-ended items in the students' questionnaire are all analyzed through mainly qualitative procedures.

Cohen, Manion and Morrison (2007) describe this type of analysis as *"[involving] organizing, accounting for and explaining the data; in short, making sense of data in terms of the participants' definitions of the situation, noting patterns, themes, categories and regularities"* (p. 461).

8.1. The Students' Questionnaire

The questionnaire administered was mainly to investigate the students' perception and attitudes towards pronunciation learning and teaching; based on *Likert Scale,* various issues were explored such as the desire for pronunciation classes, preferences about accents, and error correction.

The items in the questionnaire were classified into groups such as items meant to assess students' self-perceptions of their pronunciation and their attitudes towards the way pronunciation is taught at the department.

The questionnaire presents and discusses the answers provided by a group of sixty students of similar backgrounds to a series of questions. The provided answers are presented in the appendices. It is worth noting that instead of the complete, full-sentenced textual answers, only the relevant keywords, conveying the idea of the answer were extracted.

The purpose of each question is revealed along with the analyses of the corresponding answers.

8.1.1. Questions 1 and 2 *(Gender and Age)*

The students are of a similar age and, once analyzing the answers of the few exceptions (i.e. much older students), no significant deviation was detected.

Also, upon analyzing the answers, no perceptible difference was detected between the answers of male (42% of the students questioned) and female (58%) students. Therefore, a separate study of males' and females' answers was unnecessary.

8.1.2. Question 3: Students' satisfaction about their pronunciation

Are you satisfied with your English pronunciation? Why?

Yes ☐

No ☐

The answers revealed that 70% of the students were dissatisfied with their pronunciation. It is therefore clear that efforts must be made to improve the teaching of pronunciation in the department of English. Moreover, further analysis showed that the dissatisfaction rate did not decrease with the students' ages and their academic levels.

Regarding the reasons of the dissatisfaction, the main cause reported by the students (62% of them) was the awareness that they make pronunciation mistakes when they speak. 43% admitted that they expected to pronounce better at that point of their studies. The inability of distinguishing between British and American accent was also a point mentioned.

Surprisingly, no explicit criticism, neither towards the teaching process and the curriculum nor the teachers, was detected. Some students (7%) were displeased with the fact that they could not practice their pronunciation more, but this comment was more oriented towards the general environment of the students (i.e. home, friends, town, etc.) than the teaching environment of the university. We can conclude that the students do not expect more effectiveness from the currently available methods of teaching, believing that not much more could be done in that matter. This fact emphasizes the need for new teaching methods to be urgently introduced.

8.1.3. Question 4: Students' Preferences about Pronunciation

Which pronunciation do you prefer? Why?

American ☐

British ☐

We observed that the pronunciation preference was evenly split in half. Furthermore, no noticeable variation between the answers of American-preferring and British-preferring students was observed. This fact relieved us from considering any kind of data subdivision according to that preference.

The most interesting part of that question was related to the reasons of preference. The relevance, the variety and the soundness of the reasons mentioned confirm the admirable interest of the majority of the students in the pronunciation of English. Many (30% of the students questioned) based their preference mainly on strong personal liking. Less common answers mentioned the historical prime of the British pronunciation (8%), its broader use worldwide (7%). The enthusiasts of American pronunciation stated that it is less formal (10%) and easier. Both groups mentioned that their preferred pronunciation was easier in their view (32% of the total).

Very few students left this question unanswered, which confirms their enthusiasm for English pronunciation and contributes to our motivation to offer them new means of improving their skills.

8.1.4. Question 5: Importance of Pronunciation

How important is pronunciation to you? Why?

Very Important	Important	Less Important	Not Important

As a validation to our observation above, 67% of the questioned students stated that pronunciation is ‘very important’ to them; the remaining 33% considered it just ‘important’. None of the students chose the answers ‘less important’ or ‘not important’.

The comparison between the percentage of the students dissatisfied with their own pronunciation (70%, in question 3), and the percentage of those who considered it at least important (100%) clearly indicates a deficiency in the current teaching methods and a need for other options.

Regarding the reasons of considering pronunciation important, the students tended to agree that it is an essential part of a language (mentioned by 67% of the students), and that a good pronunciation reflects good language mastery (mentioned by 38%). Another common opinion is that good pronunciation reinforces self-confidence when speaking (18%). Once again, the soundness of the opinions confirms that the students are truly aware of the importance of pronunciation.

8.1.5. Question 6: Language Practice

Do you practice your English outside the classroom? Why?

Yes ☐

No ☐

The answers to those questions were split, with a slight leaning towards the negative. 52% of the questioned students declared that they did not practice their English outside the classroom. This result might seem in contradiction with the previous answers, which tend to agree on the importance of pronunciation. However, the contradiction disappears once the reasons of not practicing become known. More than 68% of those who did not practice mentioned a lack of English speakers with whom they could converse. Many of the non-practicing students (23% of the questioned) also declared having no adequate environment (home, family or friends' circle, etc.). 26% admitted being uncomfortable speaking English in front of people. Only 6% did not practice due to a simple lack of motivation.

Initially, the purpose of the question was to identify the causes of not practicing English outside class. We were, however, curious about the reasons of those who did practice it. The reasons were diverse, but always sound. Many (55%) of the practicing students mentioned that it is essential for improvement. Another recurring reason (25% of the practicing) is simply the fact that they enjoyed speaking English.

One reason worth mentioning despite a modest percentage (14%) is that some students practiced for professional motives, either for their current or contemplated career. This reminds us that besides the quest for good grades and a diploma, the professional future of young people is being forged and decided in the classrooms. Again, the 70% dissatisfaction rate comes to mind.

8.1.6. Question 7: Pronunciation Improvements

What do you do to improve your pronunciation?

This question was addressed to both those who did practice their own pronunciation and those who did not. We were assuming that students who did not practice still managed to find a way of improving their pronunciation, and we were correct. No difference was observed in the answers of practicing and non-practicing students.

Unsurprisingly, watching television and listening to songs were by far the most popular means used (mentioned by 58% and 48% of students, respectively). The other means mentioned, which were rather evenly split, included the dictionary, watching Internet videos, listening to news, and using subtitles when watching movies and series.

At this point, we can make a number of interesting observations:

i. Only 7% of the students mentioned listening to teachers as a means of improvement. This tends to confirm the lack of expectation towards the current teaching methods, and the need for new ones.
ii. The computerized and multimedia means of improvement are far more popular among the students than the classic dictionary (mentioned by 15% of the students). This observation confirms the potential of using software and individualized teaching methods.
iii. Many of the students who have the opportunity of listening to, and conversing with English speakers do so. We may assume that those who admitted being reluctant speaking in public (26% of the non-practicing), would eventually start conversing with English speakers if some kind of method was offered to them to reach a certain mastery level first.

8.1.7. Question 8: Use of Phonetic Lectures

When you speak English, do you use what you learned in phonetics lectures?

If the answer is no, why not? If it is yes, how often?

Yes ☐

No ☐

Three quarters of the students questioned stated that they did use the lectures contents, which was good news after the displayed lack of confidence in the teachers expressed in the previous question.

Among those who used the lectures contents, 20% declared *always* using it, 44% *usually* used it, and 20% *occasionally* did. This confirms the importance of the phonetics lectures and their usefulness in the student's curriculum.

The gathered data regarding the *why not* question was deemed irrelevant, as many students gave too vague answers, or left it unanswered.

8.1.8. Question 9: Pronunciation Teaching Classes

Do you like to have pronunciation teaching classes?

Like	Undecided	Dislike

Again, the answers confirmed the merit of the pronunciation classes and the students' appreciation of it. 95% of the questioned students declared that they like to have a pronunciation class; the remaining 5% were undecided.

8.1.9. Question 10: Importance of Pronunciation Teaching

Do you agree that pronunciation teaching should be given more importance and must be included as an independent part of preparation programs at our universities?

Agree	Undecided	Disagree

100% of the students agreed with the statement, making any deeper analysis unnecessary.

8.1.10. Question 11:

Do you encounter difficulties when you speak or listen in English because of pronunciation problems?

Yes ☐

No ☐

Anticipating the high percentage of students dissatisfied with their pronunciation, we asked this question, along with question 12, to obtain a deeper vision of the problem. The result was that 85% of the students declared encountering difficulties. This leads us once again to question the effectiveness of the current teaching methods.

8.1.11. Question 12:

In which of the following situations do you encounter pronunciation difficulties?

A. When I watch movies ☐

B. When I chat with native speakers in English ☐

C. When I listen to foreign music in English ☐

D. When I listen to the radio and television broadcast ☐

E. When I watch the different news channels ☐

The answers to this question were rather split, without any predominant one. The students mentioned common daily activities such as watching movies, chatting/speaking, listening to radio and music, etc. Even without one main difficulty, we observed a high average of difficulties encountered.

8.1.12. Question 13:

Have you ever heard about a spectrogram before your teacher introduces it?

Yes ☐

No ☐

Overwhelmingly, 97% of the students questioned never heard about a spectrogram. Even though, as we have seen in question 7, students are open to computerized means of learning, they weren't aware of what a spectrogram is.

8.1.13. Question 14:

While dealing with the Praat software in classroom and explanation of its use as a new tool in teaching an accurate pronunciation, how do you evaluate it?

Very important	Important	Less important	Not important

62% of the students stated that Praat is a *very important* new tool. The remaining 37% described it as just *important*.

8.1.14. Question 15:

After having been introduced to the Praat software, do you keep exercising at home?

Yes

No

If the answer is yes, how often?

Every day	Once a week	When I have time	Before phonetics lectures

67% of the students declared using Praat at home for practicing. Considering that 97% of them did not even hear about it before the pronunciation course, this result along with the result of question 14 is highly satisfactory. It shows a large-scale acceptance and openness to new methods of self-teaching in general and to Praat in particular.

8.1.15. Question 16:

Given more time to practice, do you think this program would help you improve your pronunciation?

Yes

No

100% of the students thought that Praat would help them improve their pronunciation. This puts the results observed in question 15 in perspective: a quick computation allows us to deduce that 32% (100%-67%) of the students did not, or could not, practice with Praat at home while being certain that it would improve their speaking skills.

According to that, we can deduce that the introduction of Praat as a tool used in class could be highly beneficial to these 32% of the students even more than to the others, as they did not practice at home.

8.1.16. Question 17:

After practicing several times in classroom and at home (if at all), are you able to be aware of the correctness or incorrectness of your pronunciation?

Yes ☐

No ☐

Similar to question 16, 100% of the students stated that Praat allowed them, after appropriate training, to correctly evaluate their pronunciation.

Even if they seemed too optimistic, or too eager to please to the author of the questionnaire, the students' optimism by itself indicates their readiness to use Praat as a means of learning as well as a means of self-evaluation. With adequate initial supervision, the students would gradually use Praat on their own.

8.1.17. Question 18:

When learning English pronunciation, I keep working until I reach the goals that I make for myself.

Strongly agree	Agree	Partly Agree	Slightly disagree	Disagree strongly	Disagree

65% of the questioned students *strongly agreed* with the statement, and 35% *agreed*. That strong statement made by the students confirms that, given adequate and modern tools to improve their skills, the students possess the sufficient responsibility and self-awareness to use them at best.

8.1.18. Question 19:

When you are conversing with English speakers, do you ask for feedback on your English pronunciation?

Yes ☐

No ☐

As expected, the majority of students questioned (73%) did not ask for feedback on their pronunciation. We did not ask the exact reasons, because it was not relevant to the main issue of our work. Additionally, thanks to our experience in teaching, we knew that these reasons are usually an embarrassment, unwillingness to interrupt or to change subject, insufficient familiarity with the interlocutor, etc.

The fact is that the students severely lack a reliable means to provide them with feedback on their pronunciation. A computerized tool is a very suitable answer to this need, as all the reasons mentioned above become irrelevant.

8.1.19. Question 20:

When you aren't sure of your pronunciation, what additional methods do you use to correct yourself?

The students provided various answers. While, the exact percentages are unimportant, the analysis of their responses allowed us to make three relevant observations:

i. As previously observed, computerized and technological methods were preferred over the classic ones.

ii. Because of embarrassment and discomfort in front of their classmates, almost all the students preferred the means of correcting themselves. Many had more than one. This confirms their eagerness to learn and to improve themselves.

iii. Methods involving other people (asking classmates) were far less popular than private methods (dictionary, internet). This is evidence again that students who sense a lacking in their pronunciation skills are reluctant to ask other people for help.

8.1.20. Question 21*:*

When you do not know how to pronounce a word in English, do you ask your teachers for help?

Yes ☐

No ☐

The teachers remaining the main source of knowledge for the students, the latter tended to ask them for help (68% of the questioned). Even the students who were reluctant to ask for help, or to converse with English speakers for practicing declared asking their teachers for help. This must not be seen as an argument for the current methods of teaching, but rather for keeping the teacher in the center when new methods are introduced. Our argumentation in favor of introducing innovative methods of teaching, in general, and the Praat software, in particular, is by no means a denial of the central role played by the teacher in the learning process.

8.2. The Teachers' Interview

As mentioned in the previous chapter, in addition to the questionnaire that was for students, another instrument was used in our work, which is the interview. It was conducted with five of our colleagues, who teach mainly oral expression and phonetics to be a representative sample to probe their ideas and explore their point of views and practices with regard to pronunciation. They were selected as a representative sample of the teachers, in terms of age, gender and educational background. Moreover, it was simply more practical to work with those teachers as they were readily accessible being at the same department.

The interview explored the teachers' attitudes and opinions focusing on the students' speaking skill, assessment of pronunciation, the curricula, the teachers' skills, knowledge and confidence, types of classroom activities practiced, etc. As mentioned earlier, we divided the interview into three main categories, and every category encompasses questions including: (I) Pronunciation, (II) Teaching content and (III) teachers' perspectives, feelings, opinions and practices.

It is worth noting that due to the length of the interview, we tended to provide summarized answers of every teacher, in the tables below, keeping their main ideas. Some particularly interesting quotes left unchanged, were illustrated. The full interview script is available in appendix II.

8.2.1. Category I: Pronunciation

***Q1**: Do you assess your students' speaking skills?*

Teacher1	Teacher2	Teacher3	Teacher4	Teacher5
Yes. Speaking skill is considered as a measure to know whether a student masters English.	I do.	I certainly do.	Yeah of course, it's important to assess their speaking skill.	Yes, I do.

Table 8.1: Teachers Assessing the Speaking Skill

We have noticed that all the teachers answered positively. This leads us to say that they are aware that assessment is a measure of students' speaking progress.

Q2: *What exactly do you assess and how?*

Teacher1	Teacher2	Teacher3	Teacher4	Teacher5
-I do assess them without following a real certain criteria. -Being available -Providing a feedback and usually the feedback must be done in an encouraging discrete way. -Record assessment sometimes by an asterisk, by + sometimes by a letter which I know what does it sand for.	-Four elements: first one is fluency, second accuracy third vocabulary and pronunciation.	-I assess the speaking skill right after the student finishes his or her answer.	-what is important in the speaking capacities is fluency. -Listening to what they say how they say it. When it's a big mistake of pronunciation or they are not using the appropriate word for the appropriate message, I feel obliged to correct them.	-Use my mental capacities to check how they answer and how they talk in class, I don't use any means of assessment.

Table 8.2: Speaking Skills Elements Assessed

The interviewees had varied answers concerning how and what exactly they assess in their students' speaking skill. The majority answered that they actually assess, in one form or another, students whether by placing an asterisk next to students' names, taking into consideration for instance 'fluency' and avoiding correcting students unless the mistake is flagrant. One teacher indicated that she did not use any means of assessment and she only relied on her 'mental capacities' to check how students talk, which still remains as a general and unspecific pattern of pronunciation assessment.

Q3: *What speaking activities do you use in classroom?*

Teacher1	Teacher2	Teacher3	Teacher4	Teacher5
Broadcast that will help them get into discussion	-Different types of activities, they could be labeled as interactive and communicative. -Listening tasks reading tasks speaking tasks that are based on discussion, and dialogues, role playing, formation gap, jigsaw activities.	-Give students the opportunity to narrate oral stories, fables to evaluate the way they tell the stories, the mastery of the language and the speaking abilities.	-Giving them opening the large space for discussion. Encourage them to speak in English	-I give free talks -I give them a text and ask them to read it -I ask them to watch a movie and come back with their opinions about it.

Table 8.3: Classroom Speaking Activities Used by Teachers

In order to improve their students' speaking skills, the teachers use several activities in classroom, which are almost all based on tips of stimulating students' discussion and interaction. One said that she asks students to watch movies and come back with their opinions. This is definitely encouraging; yet, not all the students can watch that movie or simply can afford it. This is in addition to the crowded groups which hinder the students' interaction and discussions.

Q4: *What do you do with students whose speaking is very difficult to be understood?*

Teacher1	Teacher2	Teacher3	Teacher4	Teacher5
Encourage them to work hard.	I pay attention to those kinds of students, motivate them. Show them how to pronounce the word correctly and fluently. Provide them with tips to overcome their shyness.	I merely ask them to repeat what they have said.	I ask them to say it again, in English if they fail why not in the native language.	Usually I correct them in class directly after they make a mistake. I drive their attention to the mistake when it is a flagrant. Use the board to write the word ask them to repeat and memorize it.

Table8.4: Teachers' Solutions to Students' Unintelligible Speaking

When teachers encounter students whose speaking is somehow difficult to be understood and in order to avoid making them feel embarrassed or uncomfortable because of their mispronunciation, they provide them with tips to overcome their shyness and motivate them. They also drive their attention to the mistakes, using phonetic transcription, and encourage them to freely express themselves, even by tolerating the use of their mother tongue 'Arabic' in the classroom.

The first issue which emerges from these comments is that the students' pronunciation is solely noticed when the teacher is unable to understand them. The use of the mother tongue is permitted though this may not help students improve their pronunciation if they use it each time they face communications difficulties.

8.2.2. Category II: Teaching Content

Q1: *While planning the speaking activities, can you alone, as a teacher, decide the content of teaching or have other considerations?*

Teacher1	Teacher2	Teacher3	Teacher4	Teacher5
The teacher is the only one who knows what fits his students. No external considerations.	The teacher decides alone the content of teaching.	I alone decide what to teach	In collaboration with my students that we have decided on how to have this opportunity of speaking English in classroom.	Decides alone as well.

Table8.5: Teachers' Independent Decision Making about the Course Content

The teachers' responses indicated that while planning the speaking activities, the teachers design their own curricula and decide what to teach to their students without any form of guidance or helpful framework available at the Department of English. One teacher said that she "*wished teachers would have been provided with what to teach, since we have more experienced people, who may be of a great help in this domain by providing or by suggesting a specific content to teach our students*," She added that "*she has experienced this year a collaboration with her students in deciding together what to teach and create opportunities for speaking English in the classroom*".

Q2: *To what extent is pronunciation an important feature?*

Teacher1	Teacher2	Teacher3	Teacher4	Teacher5
Pronunciation is vital in learning any foreign language. Help students figure out some phonetic rules that will help them pronounce better.	Practice is the best thing	Frankly don't give it too much room.	Pronunciation is very important.	Medium importance is given to the pronunciation

Table8.6: Teachers' Estimation of Pronunciation Importance

Some teachers agreed that pronunciation is of a paramount importance in learning any foreign language; this would be achieved not only by mastering the phonetic rules but also by practicing. One teacher said that "*frankly pronunciation is not given too much room*". This clearly shows how neglected is pronunciation in the department.

Q3: *How much time is devoted to pronunciation? Why?*

Teacher1	Teacher2	Teacher3	Teacher4	Teacher5
Normally all that lesson is going to be devoted for pronunciation The overcrowded groups Students will not welcome being corrected all the time.	/	No more than five minutes	Pronunciation should be present through the whole course not at specific moments. Not to correct the students unless the mistake is so terrible that you cannot ignore.	Not very much time. 15 minutes.

Table8.7: Teachers' Devoted Time to Pronunciation

Even though teachers are aware that more time should be devoted to pronunciation, they admit that currently phonetics is insufficiently taught in their classrooms due to several factors. The main causes mentioned are:

-The large groups of students

-The lack of time.

-Some students' embarrassment or discomfort caused of being corrected in front of the class.

-The fact that pronunciation should be the subject of continuous improvement rather than intermittently taught in dedicated lectures.

8.2.3. Category III: Teachers' perspective, Feelings, Opinions and Practices.

Q1: *How important is pronunciation?*

Teacher1	Teacher2	Teacher3	Teacher4	Teacher5
Pronunciation is vital Should be given enough attention by teachers of oral expression and phonetics as well.	Very Important I really appreciate talking a lot about pronunciation	It's so important in fact in learning a foreign language.	Pronunciation is very important.	Very important because most of the time if you don't pronounce correctly the message might not get transmitted.

Table8.8: Teachers' Estimation of Pronunciation Importance

We have asked again this question, to examine the teachers' feelings and opinions about pronunciation, still the majority acknowledged that pronunciation is actually very important in any language learning and should be given more priorities.

Q2: *Do you teach pronunciation? Why?*

Teacher1	Teacher2	Teacher3	Teacher4	Teacher5
I do not teach pronunciation as but I do never miss an occasion to correct my students' mispronunciation. Because another teacher is doing the job (phonetics)	I don't teach it. -Show them how to pronounce the word correctly and fluently.	I don't teach it. Because I see that it's more acquired than learned.	I don't teach pronunciation as a separate element.	I used to teach it with intermediate levels but since then don't have enough time to teach everybody to pronounce correctly.

Table8.9: Teachers' Choice to Teach Pronunciation

As mentioned above, all the teachers admitted that pronunciation is of a great importance for the role it plays not only in English language but also in any foreign language learning. Yet, once they were asked about whether they teach pronunciation or not, the majority answered that they do not teach it in a systematic and planned way. Their main reasons include:

-The absence of pronunciation as a separate element in the curriculum.

-The amount of time and efforts devoted to it.

-Other teachers are supposed to teach it. (Phonetics' Teachers)

-It's more acquired than learned.

Q3: *If you teach pronunciation, on which aspects do you focus?*

Teacher1	Teacher2	Teacher3	Teacher4	Teacher5
I do not have a special area that I would focus on	Interactive and communicative at the same time.	/	Correct pronunciation. -Correct English whether American or British. Their intonation	Focus on the tone how the students use the tone and the correct pronunciation of every letter.

Table8.10: Aspects of Pronunciation Stressed by Teachers

Although teachers indicated earlier different reasons which make them reluctant to teach pronunciation, they tend to avoid dealing with some fundamental elements of teaching phonetics such as the focus on the prosodic features of language. Prosody here refers to intonation, tone of voice, emphasis, pauses, and changes in speech rate. They mentioned that they just concentrate on *correct pronunciation* whatever the accent is which is referred to as *intelligibility*.

Q4: *Do you feel confident when you teach it?*

Teacher1	**Teacher2**	**Teacher3**	**Teacher4**	**Teacher5**
I do not teach pronunciation”	Why not ?	I don’t teach it.	I don’t teach pronunciation as a separate element.	Of course I am very confident.

Table8.11: Teachers’ Confidence when Teaching Pronunciation

Considering that none of the teachers reported explicitly teaching pronunciation, no significant data could be extracted from the answers to this question. This leads us to hypothesize that teachers’ lack of confidence is basically due to the lack of suitable materials, skills and knowledge in the field of phonetics; this results not only from teachers lack of interest but also from avoiding to deal with it. These may be amongst the reasons why pronunciation is ‘marginalised’.

Q5: *Which pronunciation do you use: American, British or a mixture of both and why?*

Teacher1	**Teacher2**	**Teacher3**	**Teacher4**	**Teacher5**
British	I do vary from American to British	British	I cannot identify myself as speaking the British accent or the American accent as far as it is an **Algerian** accent a combination of all accents.	I use American English at a 100%

Table 8.12: The Pronunciation Variety Taught by Teachers and Why

Two teachers said that they use British English may be because it is considered to be historically original and sounds more formal. Two others wrote that they vary from American to British; one of the latter said that she uses only the American English because she spent around one year in the USA; she masters this accent very well. Surprisingly, the last teacher mentioned that she uses neither the English nor the American accent but rather what she referred to as *Algerian accent* which is, according to her, a mixture of all accents. We noticed that the choice of the pronunciation accent being implicitly taught is solely based on the teachers' personal preference. It can be one or the other or even a mixture of both. The fact that the teaching is implicit makes students totally unaware that the choice is being made for them. Students should be at least aware of which accent is being taught. Ideally, the students should be made aware of the major accents and be able to distinguish between them.

Q6: *Of which pronunciation materials are you aware of or familiar with?*

Teacher1	Teacher2	Teacher3	Teacher4	Teacher5
Download lot of broadcast on BBC Use as authentic material.	I have software taken from Cambridge on my dictionary.	I am aware that they exist but I ignore the types of these devices	No, I am not familiar with any.	No one!

Table8.13: Pronunciation Materials with which Teachers' are Familiar

Pronunciation is dealt with in classrooms by teachers who develop their own materials, do their best to help students improve their pronunciation, but we cannot deny the existence of a number of different approaches, materials, software used to teach pronunciation. The majority of the interviewees are not aware of their existence.

Q7*: Have you heard about the Praat software before, and have you got an idea about what a spectrogram is?*

Teacher1	Teacher2	Teacher3	Teacher4	Teacher5
No	I've heard about it from you actually in one of our informal discussions.	Ah never, I have never heard about this before. No, never. I have never heard about such devices.	No never. I think it's not a new term, but knowing exactly what is it No	No, no never A spectrogram is a device that measures the tone I think of speaking.

Table8.14: Teachers' Knowledge of Praat and a Spectrogram

All the teachers indicated that they have heard about neither the Praat software nor spectrograms before except one teacher who said that she heard about it in one of our informal discussions. This leads us to say that teachers, besides feeling reluctant and less confident to teach phonetics, have not got the necessary pronunciation materials and more precisely the Praat software.

Q8: *Do you have any knowledge in acoustic phonetics?*

Teacher1	Teacher2	Teacher3	Teacher4	Teacher5
No. Just a definition	Honestly I have not studied it. Studying it academically I mean.	No.	No.	No, Never.

Table 8.15: Teachers' Knowledge of Acoustic Phonetics

The question was asked to determine whether teachers had a prior knowledge in acoustic phonetics, which deals with the physical properties of the transmission of speech sounds. Having at least some fundamental elements in this domain, would help both teachers and students use software like Praat, which is a computer program that enables visualizing, playing, annotating and analyzing sounds in terms of their acoustic properties (e.g. frequency, pitch, etc.).Unfortunately all teachers indicated that they had never had a mere idea about this elementary field of phonetics.

Q10: *According to you, why are many teachers afraid of phonetics?*

Teacher1	Teacher2	Teacher3	Teacher4	Teacher5
Well, may be because it is more technical	I think the main reason, is because we were not exposed to native speakers environment, we do not master the appropriate rules of pronouncing correctly or appropriately and this is due to the lack of opportunities given to us to live in an authentic environment	Frankly, don't like it. I prefer literature	Because we don't have enough knowledge in teaching it.	We were not taught phonetics; We don't have specialists in phonetics; We don't have appropriate materials.

Table 8.16: Teachers' Apprehension of Teaching Phonetics

Phonetics is a sub-field of linguistics which is concerned with how sounds are produced, transmitted and perceived. It has three major categories which are the articulatory, the acoustic and the auditory ones. We can ascertain that in our department we only focus on the articulatory side of phonetics, which is the production of sounds and neglecting the other two branches except by providing their definitions. Their lack of interest or fear of phonetics has been justified as follows:

- The absence of specialists in phonetics
- They were not taught phonetics as it should be.
- Phonetics is not a simple subject to be taught in classrooms, it needs physical materials to be taught; we cannot teach it with a chalk and a board.
- Teachers are afraid of teaching phonetics because they know they do not master the appropriate rules of pronouncing correctly or appropriately.

- Due to the lack of opportunities given to them to live in an authentic environment, to spend at least a good time in an authentic environment may be.
- A lack of suitable teaching and learning materials of a high quality.

Q11: *What are your suggestions for teachers who teach oral expression and phonetics?*

Teacher1	Teacher2	Teacher3	Teacher4	Teacher5
To have a continuous coordination between both teachers.	Provide the appropriate environment for teachers and for students. Want to travel to a foreign country to be in contact with native-speakers. Want instruction in English pronunciation.	Well, to try to work together. Provide the appropriate environment for teachers and for students.	To coordinate both modules. Give both teachers the important devices and materials to teach these two modules adequately.	It is not a suggestion for teachers more, it is a suggestion for the administration to give both teachers the important devices and materials to teach these two modules adequately, some teachers are helpless.

Table 8.17: Teachers' Suggestions for Teachers of Oral Expression and Phonetics

Most of our interviewees seemed not really to be confident in their pronunciation abilities, but, nonetheless, they wanted to improve their students' pronunciation. Their suggestions for both teachers of oral expression and phonetics were as follows:

- Provide the appropriate environment for teachers and for students,
- Want to travel to a foreign country to be in contact with native-speakers.
- To have a continuous coordination between both teachers.
- Want instruction in English pronunciation.
- Give both teachers the important devices and materials to teach these two modules adequately.
- Teachers have teaching potentials and capacities, students are motivated and willing to experiment new equipment and instruments (Digital tools, ICT).

In this chapter, we will also present and analyse the results of an experiment completed within the scope of this works, which consisted of the practical use of Praat software by sixteen students.

8.3. Description of the Experiment

In the first phase of the text, sixteen students were asked to listen to two recordings of native speakers reading a thirty word long sentence. As in the previous work involving students, all were of similar age and background.

For the second phase of the experiment (the pre-test), which took place immediately after the first, we asked the students to read the same sentence and try to produce the closest pronunciation to one of the native speakers possible. They were free to choose the native speaker they preferred, American or British.

In the third phase (the post-test), the students were then asked to use the Praat software intensively during several weeks. Besides the main goal of globally improving their pronunciation skills, the task that was assigned to them in the scope of this pronunciation test was, again, to produce the closest pronunciation to the one of the native speaker they chose in phase two.

Between phases two and three, i.e. the period when the students were asked to practice using the software, we emphasized that the experiment was not an assignment, that no extra points would be granted for the participating students and that the results would not be used anywhere else outside the scope of this work. The idea was to minimize as much as possible the motivation to please the teacher or to obtain good grades, leaving personal curiosity and interest in the method as the main factors of motivation to complete the experiment.

The collected pre-test and post-test data were then gathered and analysed. The complete data set of the results is presented in the table in the appendices.

In some particularly interesting cases, the data related to the pitch variation of some of the participating students was collected through segmentation of spectrograms and analysed.

As an additional step, after the end-year examinations took place, the final grades of the students who took part in the experiment were calculated and compared to the ones of a the control group. The grades, the analysis and the conclusion are presented in this section.

Both groups were pre-tested on the *Independent Variable* which is the innovative method while the *Dependent Variable* was the teaching of the pronunciation. Finally, general impressions and attitudes of the participating students were monitored throughout the experiment.

8.3.1. Purpose of the Experiment

The purpose of the experiment was to assess the efficiency of the Praat software, as an innovative tool for self-teaching, in improving the students' pronunciation skills, in a context where teachers themselves admitted that pronunciation is not sufficiently taught in the classrooms.

Efforts were made to ensure that the students were solely driven by curiosity, interest in new methods and goals of genuine self-improvement. Therefore, if positive results were to be observed, that would confirm the efficiency of the tool both in improving the students' pronunciations and in gaining their interest.

Regarding the comparison of grades mentioned above, it had two purposes. The first was to use the students who did not participate in the experiment as a control group, and the second is to determine if the improvements observed when using the Praat software, if any, were inducing more general improvements of the students in their curriculum. The idea here is that self-teaching, innovative learning methods, and the satisfaction of obtaining positive results might increase the global motivation of students and improve their interest and skills in pronunciation, phonetics, and English language in general.

The aforementioned examples of pitch variation were extracted from the students' recordings (spectrograms) for the purpose of illustrating and confirming those students' improvements in reproducing the native speakers' pitch variation. They are representative of their improvements in resembling to how the native speaker sounds. We will show how the visual analysis of pitch variation can deal with errors of misplaced pauses and incorrect intonations.

8.4. Results of the Analysis

In table IV (1) in appendix IV, the phonetic transcriptions of both native speakers' pronunciations are provided at the top, after the literal transcription of the sentence. Next in the table, the phonetic data of each student follow, separated for each word of the sentence. It details pre-test and post-test phonetic transcriptions and an objective evaluation of whether the student:

a. Improved his/her pronunciation (noted '*better*' in dark green);
b. Had a correct pronunciation initially and kept it (noted 'same/gd.' in light green);
c. Had an incorrect pronunciation initially and kept it, or substituted a pronunciation error for another (noted 'same/inc.' in orange);
d. Degraded his/her pronunciation of the word, either by accentuating the initial mistake, or by making a worse one.

Two analysis approaches were possible, and both were performed: in the first, data were grouped by words pronounced, whereas in the second, data were grouped by student.

There are also general observations which are independent from either approach. They are the most obvious and the most encouraging:

- On average, during the experiment, 41% of all the words pronounced by all the students were better pronounced after using Praat. This number might seem unimpressive until we notice that 50% of the words were correctly pronounced from the beginning, leaving only 50% of the words mispronounced. Removing these 50% of correctly pronounced words would mean that the use of Praat is responsible of correcting, or reducing the extent of 82% of the pronunciation errors that occurred during the experiment.
- For all the words of the test sentence, the combined rate of both initially correct pronunciations and improved answers reached 91%. It means that only 9% of the students who used the software did not benefit from its use. This may be due to an incorrect understanding of how the software works, limited computer skills, or sheer lack of motivation. As there is no real way to track the students' real usage of the software, it even may be that some students did not use the software at home at all. Whatever the reason is, it is clear that teacher's supervision is still essential, even in the case of self-teaching methods and approaches.
- We also observed that during the pre-test recordings, some students did not respect the rules of using weak forms of function word, as they were supposed to, when pronouncing a sentence (connected speech). The analysis of the posttest transcriptions allowed us to conclude that, with the condition that the native speaker him/herself respects those rules, the students trying to imitate him/her would tend to start

following these rules even if they were not explicitly asked to do so. A good example of this improvement, taken from the real data we gathered, would be initially pronouncing "*should*" in its strong form (ʃʊd), then correcting by using its weak form (ʃəd).

It is worth noting that some readers might question the accuracy of the method we used to obtain the percentages mentioned above, in the sense that we did not take into consideration the simplicity of the words, their commonness, and their difficulty to pronounce, etc. Therefore, each word should not have the same statistical weight in the final rates. We can argue to that, pointing out that the test sentence used was carefully chosen because the number of easy, difficult, rare and common words is similar to what can be found in typical English sentences.

8.4.1. Analysis of Data Grouped by Words

The analysis focused on how each word of the assigned sentence was pronounced by the students as a whole. Summarizing values related to this approach are below the data table.

A thorough examination of the data collected led us to the following observations:

- Students' pronunciation mastery varies greatly from one word to another. A student might seem to master pronunciation due to his/her limited vocabulary, or a deliberate choice to use the easiest words to pronounce among possible synonyms. With some exceptions due to factors such as the open/closed form factor, a clear pattern shows that the simple and common words of the sentence were the ones the students correctly pronounced from the start (e.g. 'all'94%, 'free'100% and'act'88%). The more complex and less common words were, the fewer students knew the right pronunciation in the pre-test phase (e.g. 'brotherhood'19%, 'endowed'19% and'towards'13%). This indicates that in order to improve their overall pronunciation skills, students have to be challenged to pronounce words that they either do not use in their vocabulary or do not hear frequently enough in their daily activities. Praat software is a very interesting tool in that perspective, since the test sentence the student has to pronounce and reproduce the pronunciation can be imposed by the teacher. Homework assignments and classroom activities can be defined in the form of pre-decided sets of sentences to listen to and to pronounce, improving the students' pronunciation and enriching their vocabulary.

- One might intuitively presume that the more difficult and uncommon a word is, the less is the chance that its pronunciation would improve. Interestingly, it is the opposite that is observed: students clearly improved the pronunciation of the words that are considered more complex and less common (e.g. 'brotherhood'75%, 'endowed'69% and'towards'69%), and the occurrences for these same words to be still incorrectly pronounced during the-post tests were rather low (6%, 13% and 19%, respectively). On the other hand, in some cases of simple and rather common words, the rate of incorrect pronunciation persisting in the post-test was abnormally high ('are'25%, 'equal'25% and'should'38%). We conclude that when students are shown a word with which they are consciously unfamiliar, they are more open to listening, mimicking, and learning its pronunciation. Surprisingly, when the students listen to words they are very used to hear or to speak, even incorrectly, they tend to disregard the pronunciation of the person, or the recording to which they are listening. This might be related to self-confidence, and the exact reasons of that behaviour are outside the scope of our work. We can, however, conclude once again that providing constant and fast feedback or an efficient means of self-evaluation of the students' pronunciation is crucial to their learning. Once the students become familiar with a set of words, it will become increasingly difficult to influence and improve their pronunciation.

- For some words, it has been noticed that almost all the students made the exact same pronunciation error (e.g. the word 'endowed' with the diphthong /@U/ instead of /aU/). English spelling being not so phonetic, such error patterns reveal that students confuse written and spoken forms of the words. It means that the current teaching techniques fail to provide the students with adequate skills to distinguish between written and spoken forms. An advanced use of Praat would allow the detection of such patterns and help dealing with them.

The fact that variations in pronunciation mastery change so much from one word to another cannot be explained by the only two points above (i.e. the difficulty of the word and self-confidence). Other factors are certainly involved, such as an incorrect understanding of pronunciation rules taught to the students earlier during the curriculum. Further works are certainly required, and the Praat software is an ideal tool for that matter. Large groups of students can be assigned dozens of preselected test sentences to use in the software, and the large amount of collected data can be analyzed for patterns and recurring errors. A simpler approach would consist in a teacher initiating a discussion with students and evaluating their pronunciation. In both cases, patterns should emerge, revealing what can be efficiently self-taught using the software and what should be more focused on during class time and theoretical lectures.

8.4.2. Analysis of Data Grouped by Students

The analysis focused on how each student was involved in the experiment, with performance ratings evaluated for the test sentence as a whole.

The analysis of the data led us to the following observations:

- There is a disturbing variation in the students' mastering of English pronunciation, considering that all the students involved in the experiment were from the same group. The best student had a 90% rate of pronouncing words correctly in the first time, whereas the two students with the poorest performances had these rates at 23% and 30%. If we dismiss the mentioned best and worst two results, as the top student might have had some unknown advantage (e. g. lived abroad, had an English speaking relative, etc.), and the poorest results were those of foreign students with different linguistic backgrounds, we are left with rates ranging between 37% and 63%. This large variation is clear evidence that teaching pronunciation has to urgently be improved at the department, as the students clearly rely more on themselves, with various rates of success to learn English pronunciation. This problem was clearly mentioned in the students' questionnaires, and it is now confirmed in this part of the experiment. Again, we point out that the supervised use of the Praat software can be part of the solution.
- We noticed that after the use of the software, i.e. during the post-test, the students' mastery rates became much more homogenous. What we call mastery rate can be obtained by adding, for each student, the percentage of words of which the pronunciation improved, and the percentage of words that were correctly pronounced since the beginning. In this way, the mastery rate is a fairly accurate indication of the students'pronunciation skills. The fact that the mastery rate became more homogeneous is very encouraging, because it means that all the students who took part in the experiment were much closer to the same level. For instance, the two students with the poorest pre-test scores (mentioned above), were the two with the highest rate of improvement, which placed them much closer to the average of the group. The implication of this observation is that the use of the Praat software can be easily

integrated in the learning curriculum, and the goals of the courses, homework and classroom activities that would use it can be clearly defined and anticipated. Better still, that would mean that there would be much fewer students left with desperately poorer skills than their group, with all the negative consequences imaginable.

8.5. Spectrogram Analysis and Pitch Variation Interpretation

The figures 8.1, 8.2, 8.3 and 8.4 represent screenshots of Praat software showing spectrograms in grey, with pitch variation highlighted in blue.

Since the intonation contours and the pitch levels can be easily visualized by our students on the spectrograms, this visualization enhances the students'comprehension not only of the intonation patterns studied theoretically in classroom (e.g. falling, rising...etc.) but also the pitch levels (e.g. high, mid, low…etc.)

The following is a summary of the relevant points related to this step of the experimentation:

- As an example of the initially incorrect pitch that could be corrected through visual analysis, some students initially raised their pitch strongly towards the end of the sentence so that their realization did no longer appear to be a statement, but rather sounded like a highly astonished question (as in figures 8.1. and 8.2). However, during the final recording, we observed that the students were able to correct themselves thanks to the visual representation of the pitch variation. The pitch movement was respected with a fall at the end of the first taught group and at the end of the pitch with a considerable fall, as it is a statement (as in figures 8.3 and 8.4).

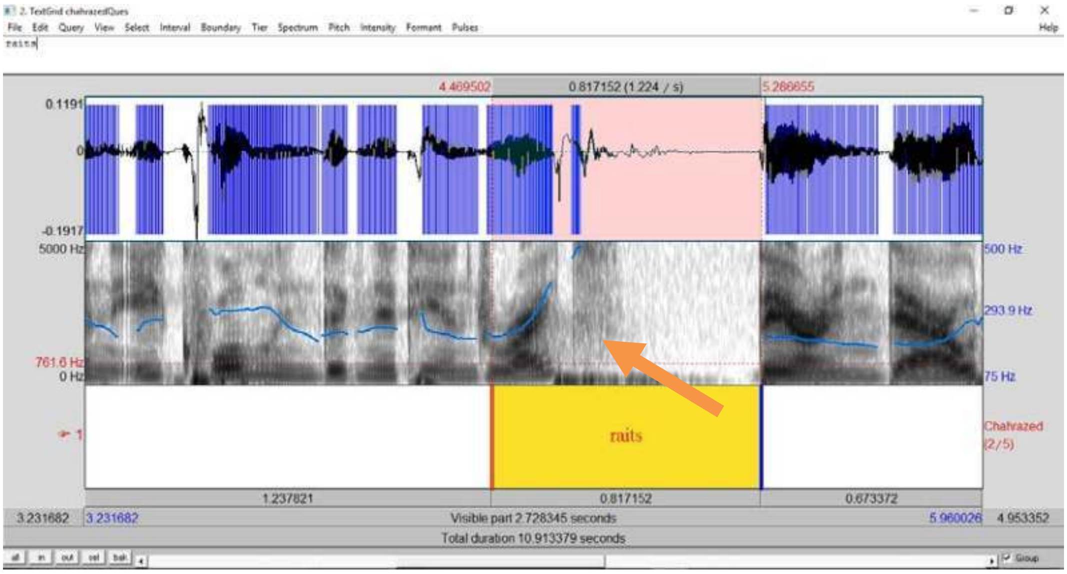


Figure8.1: Spectrogram of Pre-test Pitch Variation. Chahrazed A. – part 1

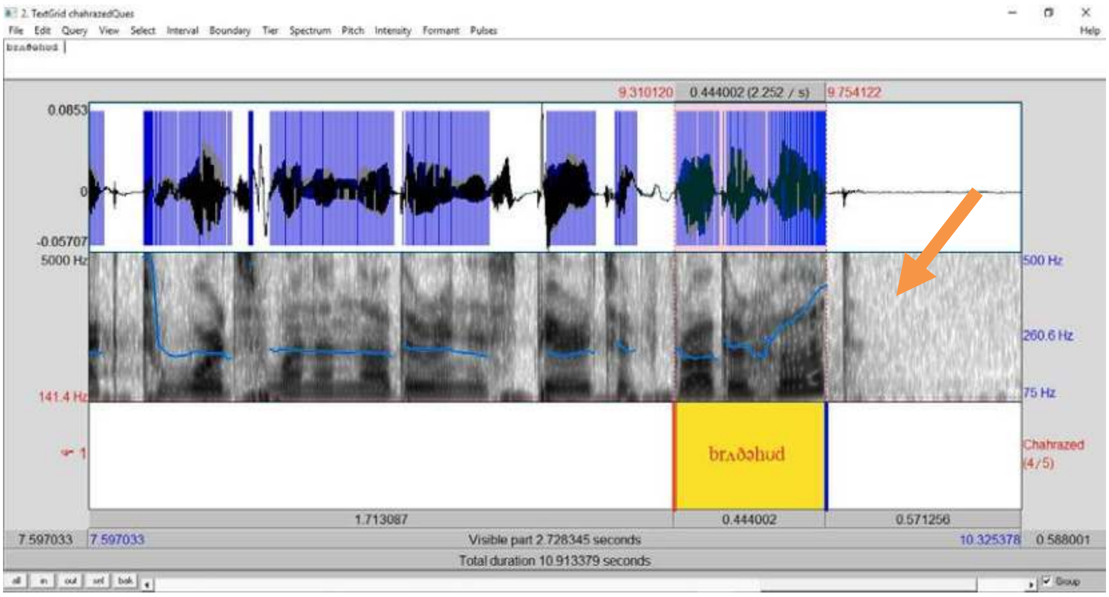


Figure 8.2: Spectrogram of Pre-test Pitch Variation. Chahrazed A. – part 2

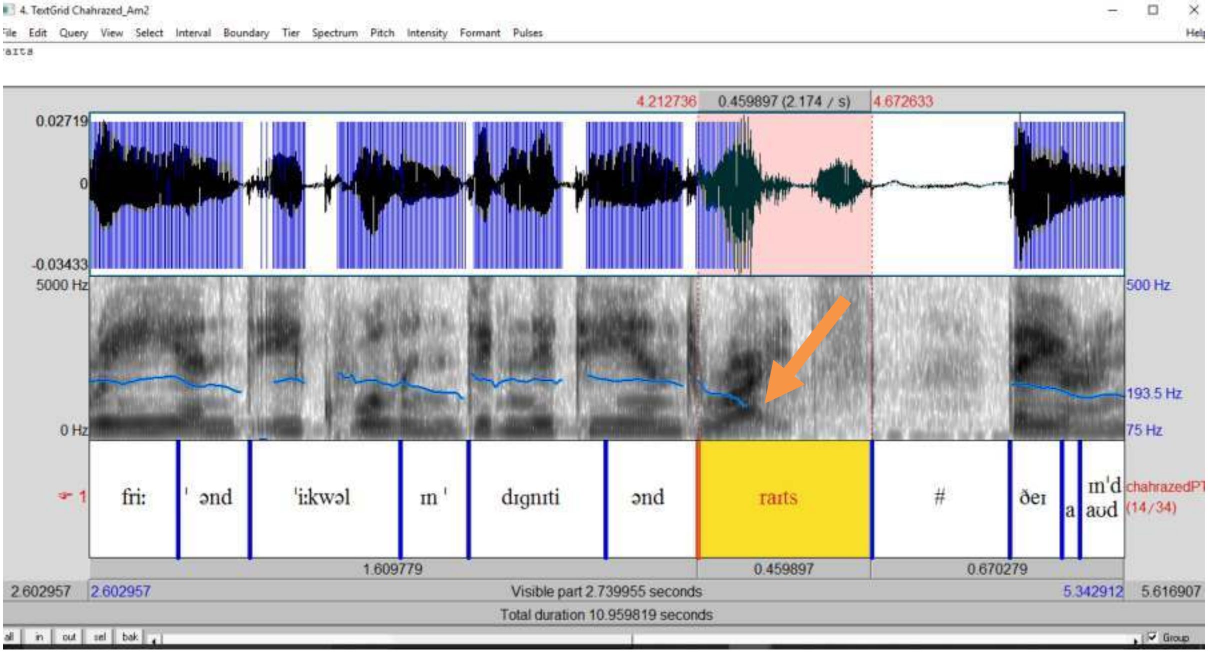


Figure 8.3: Spectrogram of Post-test Pitch Variation. Chahrazed A. – part 1

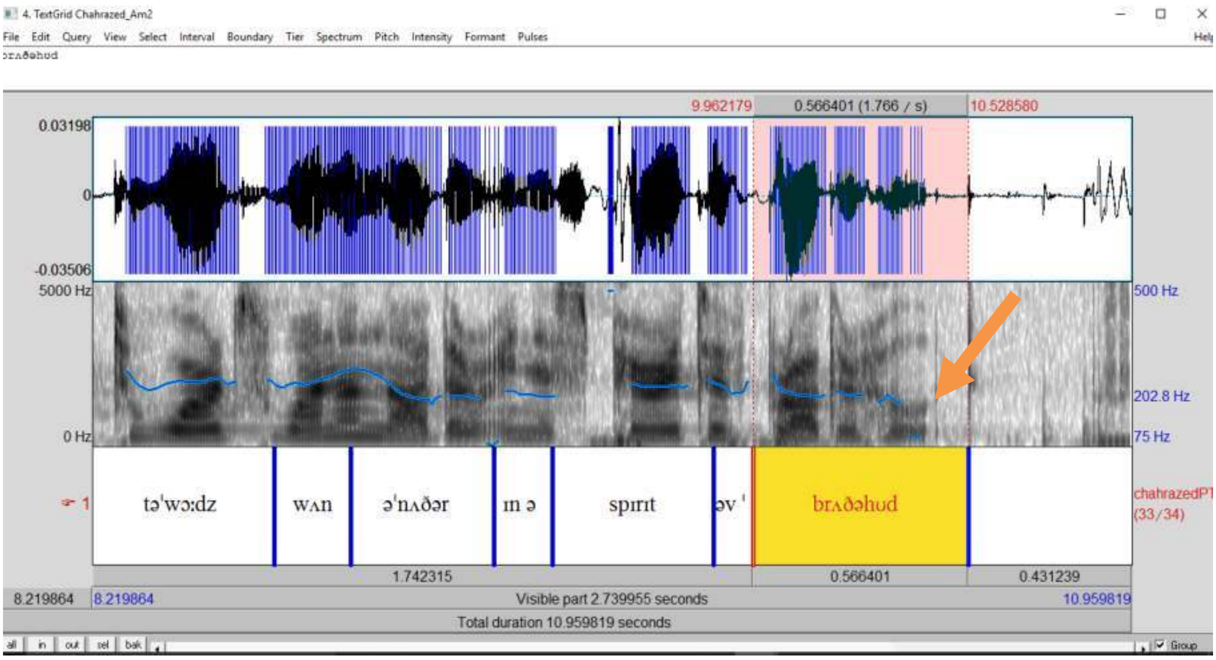


Figure 8.4: Spectrogram of Post-test Pitch Variation. Chahrazed A. – part 2

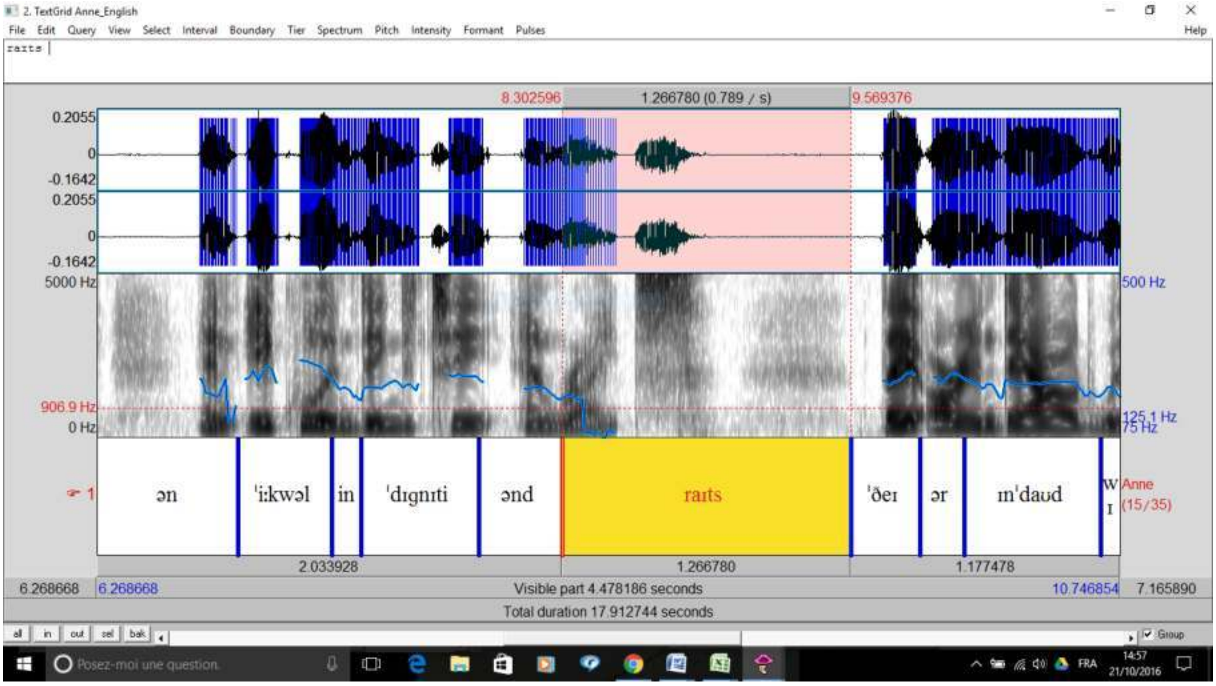


Figure 8.5: Spectrogram of Pitch Variation. Anne Marie. Part 1

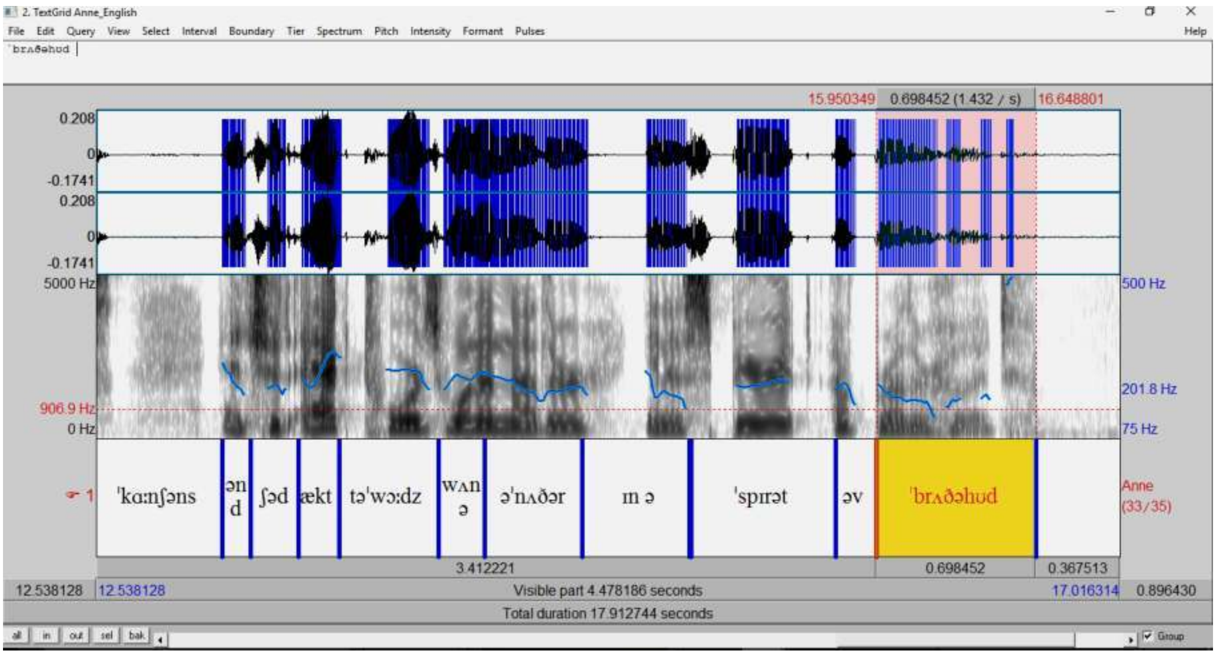


Figure 8.6: Spectrogram of Pitch Variation. Anne Marie. Part 2

- Since the visual examination of the pitch variation also allows the identification of pauses, silences and their duration, we were able to observe that students paused slightly more often (i.e. the mean length of these pauses was longer) than the native speakers during the pre-test. Again, thanks to the visual comparison made by the students, we noticed during the post-test recordings that the pauses were well respected and matched those of the native speakers'. Notice that, according to Duez (1982:13), the pause is defined as *any interval of the oscillographic trace where the amplitude is indistinguishable from that of the background noise"*

- In addition to the visualization of intonation and pitch patterns, the Praat software enables students to improve their pronunciation; other ways of testing the students' pronunciation would be the automatic and exact acoustic values that any student could practice until he/she achieves a pronunciation native like. This feature is practically useful for the students who fail to improve their pronunciation using the *listen and repeat* method. The followings are examples of the acoustic values (Intensity, formants, time and pitch), extracted from the Praat segmentation of one of the native speakers and one of the students of the experimental group.

Anne	**ɔːl**		**Omar B**	**ɔːl**
Intensity (dB)	62.90		**Intensity (dB)**	63.21
Formant1 (Hz)	510.34		**Formant1 (Hz)**	441.07
Formant2 (Hz)	919.56		**Formant2 (Hz)**	1422.01
Formant3 (Hz)	2870.75		**Formant3 (Hz)**	2388.06
Formant4 (Hz)	3665.23		**Formant4 (Hz)**	3308.82
Time (s)	0.395		**Time (s)**	0.230
Pitch (Hz)	272.08		**Pitch (Hz)**	154.37

Table 8.18: Acoustic Values after Segmentation

8.6. Students' Final Examination Grades

As stated previously, when selecting the group of the students involved in the experiment, we made efforts to ensure that the group was balanced, consisting of students with mixed abilities. This concerned the students of both the experimental and control groups.

The following table summarizes the gathered students' grades, their progress, and the degree of homogeneity of the final level of the groups.

Experimental group				
Students	First-term exam	Final exam	Variation (points)	Variation (percentage)
H. Amina	15.00	16.25	1.25	8.33%
A. Lyria	12.75	14.50	1.75	13.73%
B.Radja	12.50	15.00	2.50	20.00%
A. Chahrazed	12.25	17.00	4.75	38.78%
A. Amina	11.75	13.50	1.75	14.89%
G. Abir	11.50	15.25	3.75	32.61%
B. Ihcen	11.00	15.25	4.25	38.64%
B. Sofiane	10.75	13.50	2.75	25.58%
B. Amina	10.25	13.50	3.25	31.71%
B. Khadidja	10.25	15.75	5.50	53.66%
B. Kamel	9.50	12.50	3.00	31.58%
A. Ahmed	9.25	13.50	4.25	45.95%
B. Ghafour	9.00	14.75	5.75	63.89%
B. Moaiz	8.75	12.50	3.75	42.86%
A. O.Karim	8.50	12.00	3.50	41.18%
B. Haidara	6.00	14.00	8.00	133.33%
Mean	10.56	14.30	**3.73**	**35.36%**
Std. deviation	**2.05**	**1.37**		

a. Experimental Group

Control group				
Students	First-term exam	Final exam	Variation (points)	Variation (per centage)
A. Imene	13.00	9.25	-3.75	-28.85%
A. Sihem	12.50	9.50	-3.00	-24.00%
B. Walid	12.00	12.25	0.25	2.08%
B. Nabila	11.00	10.25	-0.75	-6.82%
B. Samya	10.75	13.25	2.50	23.26%
A. Imane	10,25	11.50	1.25	12.20%
B. Fatma	10.25	10.75	0.50	4.88%
A. Mahbouba	10.00	6.00	-4.00	-40.00%
A. Salima	10.00	8.50	-1.50	-15.00%
B. Rachida	8.75	11.50	2.75	31.43%
B. Yazid	8.75	8.00	-0.75	-8.57%
A. Zahia	8.50	12.50	4.00	47.06%
B. Faiza	8.50	10.00	1.50	17.65%
A. Nesrine	8.25	8.50	0.25	3.03%
A. Nedjoua	8.00	12.75	4.75	59.38%
B. Sihem	8.00	8.25	0.25	3.13%
Mean	9.91	10.17	**0.27**	**2.68%**
Std. deviation	**1.57**	**1.98**		

b. Control Group

Table 8.19: First-Term and Final Grades of the Students Involved in the Experiment

The comparison of the first-term grades of the two groups and the two mean values demonstrates that the students of both groups had a very similar level, as they were selected from the same class.

Yet, the final exam grades show that the students who practiced using the Praat software achieved grades that were on average more than 4 points higher than the grades of the control group. The average progress rate made by the experimental group is 35.36%, which is much superior to the 2.68% value of the control group. This, by itself, clearly confirms that the introduction of an innovative learning method in the curriculum will be highly beneficial to the students.

Moreover, we observed that the use of the Praat software by the experimental group increased the homogeneity of the group, in the sense that their final grades were much more similar and the calculated standard deviation decreased from 2.05 to 1.37. The control group, however, showed a decrease in the homogeneity of the grades, revealed in the form of an increase of the standard deviation from 1.57 to 1.98.

This observation revealed a second advantage of the tested method: increasing the homogeneity of a group of students will enable them to start the next phase of their studies roughly at the same level, avoiding cases of students falling behind or slowing down their classmates. Even when dealing with students in the final year of their curriculum, it is essential that they obtain their diploma with the same set of skills. Maintaining homogenous groups is a crucial aspect of learning, and the presented method is clearly very beneficial to that end.

The case of a foreign student, who had significant adaptation issues at the beginning of the year, is worth mentioning. Due to his linguistic backgrounds, both his pre-test performance and his first-term grade were below the average of the students. For instance, he had a great difficulty pronouncing the /ð/ sound, which was confused with the /v/, /d/ and /f/ sounds alternatively. The use of the Praat software allowed him to improve both his pronunciation to the degree that his performances in both the post-test and the final exam were on par with the rest of the group. Other foreign students with similar difficulties, who did not participate in the experiment, did not show such great improvement.

The followings points summarize the particularly relevant observations and descriptions of students' behaviours during the experiment carried out in classroom:

- What we have observed in classroom is that using computer software in teaching pronunciation provided a new and dynamic environment of learning for our students, as they felt excited, very enthusiastic and eager to learn. They experienced new technology to assist their pronunciation learning; this aroused their curiosity and will to explore more this software.
- During the use of the Praat software, students, mainly those shy and reticent ones, appreciated the ability of recording themselves as many times as they wanted, then listening and checking their pronunciation.
- It was noticed that students seemed encouraged, motivated and also challenged to carry on practicing their pronunciation using the Praat software.
- Students cooperated with each other in classroom using Praat, created an enjoyable and interesting learning atmosphere for them unlike the traditional way of teaching, where they had to inactively sit in classroom, listen to the teacher and try to memorize some rules for the exam.
- The number of errors significantly decreased in the second recording. At least students perceived the difference between their first recordings, the one of the

native speakers with their second recording. They were not phonologically deaf.

- After the completion of pronunciation teaching training using the Praat software, and after the fifth lecture, students were asked to complete a post-test which was a second recording of the same sentence. Comparing students' post-test scores to their pre-test scores enabled us to see whether students' training was successful in motivating them to use the Praat software and practice their pronunciation.

According to Boyd & Murphy (2002):

> *Software used for teaching pronunciation makes the invisible sound become visible and concrete graphics appear in front of the foreign language learners. The learners learn to pronounce the sound not only by listening, imitating and repeating, but also through receiving feedback. Therefore, learners may receive feedback without suffering embarrassment in front of other students. (p.37)*

Conclusion

The data gathered from the students' responses to the questionnaire allowed us to compile the following observations:

a. There is a high rate of dissatisfaction among students regarding their pronunciation skills.
b. The students do not believe that the current teaching methods will help them improve their level of pronunciation; they tend to seek for means of improvement elsewhere.
c. The students are highly motivated to improve their pronunciation skills for a variety of reasons.
d. The students encounter real life difficulties because of their low level of pronunciation skills.
e. The students are open to innovative, computerized and multimedia methods to improve their pronunciation.
f. The introduction of Praat is highly successful, and many students admit its usefulness.

The main conclusion to be drawn from the questionnaire is the fact that a massive introduction of the Praat software in the classroom could be of tremendous benefits to the vast majority of English students.

The interview with the five teachers, explored different issues which prevent them from dealing with pronunciation. The interviewees' responses indicate that it is possible to ascertain how widespread these attitudes are among other teachers in the department. In short, pronunciation does not appear to have a central and integrated part in the teachers' priorities.

Nevertheless, the comments made by this group indicate that there is a great need for ongoing development in the teaching of pronunciation among teachers.

Recommendations have been made with a view to overcoming teacher’s reluctance in this area and encouraging them to teach pronunciation confidently and effectively. This includes providing materials for classroom use as well as those for student self-access by integrating pronunciation software which enables them to gain time and manage students’ progress. The teaching of pronunciation could be then encouraged and supported.

Teaching pronunciation, with reference to prosody in a real, computerized and a student-centred environment would be more useful for the students, mainly because they will have more opportunities to practice these prosodic features and promote their performance.

We highly recommend the need for change in the curriculum by integrating pronunciation within formal lectures.

Pronunciation should be given increased prominence, become a central component of a number of the competencies that students need to achieve to obtain a high accomplishment level.

Besides determining the status of pronunciation teaching at the department, the analysis of the results of the experiment clearly confirmed the potential benefit of using that software, in the sense that it established a measurable improvement in the students’ pronunciation skills during the experiment, as well as in class.

Moreover, it was established that the technique of using this software and the results it allows to achieve make it highly compatible with standard university curriculums.

A significant improvement in the fluency of the students' speech was noticed; they considerably improved their pronunciation.

General Conclusion

General Conclusion

Many technological tools and software are being developed, which can be of a paramount importance for both teachers and students for their usefulness and effectiveness in teaching and learning English pronunciation. However, this requires finding ways to put into practice not only these tools but also teaching methods and activities which are more appropriate to meet the needs of both teachers and students.

Therefore, this study was meant to bridge a gap which currently exists in the Department of English, at Batna II University. In our work, we have introduced and used Praat software as a teaching/learning tool, which allowed us to record, visualize and depict our students' errors of pronunciation. Our aim is setting new learning goals and showing how to renew the learning process, based on improving students' listening skills combined with visible cues on spectrograms with sounds given by the Praat software. These practical skills are combined with basic theoretical lectures we taught our students in acoustic phonetics. Such an approach will help students understand how intonation and pitch work in sentences in English language.

The outcomes of this work are in conformity not only with the questions asked but also the literature reviewed.

Findings with the research question (1):

How can innovative techniques and devices to teach pronunciation make phonetics' sessions more interactive and more interesting?

It was proven that students in computerized learning environment appeared to be livelier, active, less anxious and enjoyed phonetic lectures learning as compared to the traditional classes.

Findings with the research question (2):

To what extent would the use of Praat software improve students' pronunciation?

Students were asked to install Praat on their computers which allowed them to practicing the software outside the classroom.

They started to work with the software at home to put into practice its different prosodic functions explained by the teacher in classroom.

Students cooperated with each other in classroom, which made learning so enjoyable and interesting for them.

The most interesting results, obtained in this research, were not only the students' performances and their significant improvements in bettering their pronunciation but also the benefit that they received from working with the Praat software. Therefore, students had plenty of opportunities for recording, practicing as many times as they wished at home without even the presence of the teacher.

What we can say is that the application of the Praat software revived our classroom teaching and helped improve the reliability and validity of the students' pronunciation evaluation. Praat helped students improve their pronunciation due to factors such as practice sessions in which they could take risks without stress and fear of error in front of the entire class.

In conclusion, Praat proved to be a useful software tool for pronunciation training. By using Praat, students were able to record and analyze their own intonation. Even if the students' post test results did not match the exact frequency of the native speakers, but at least this could develop the students' ability of self-correction. Thus, they could compare their voice to a native speaker, and they were able to track their progress over time after hearing and practicing.

The current work was an innovative approach to teaching phonetics and in particular pronunciation at the department of English at Batna University with an attempt to explore the use of Praat, a useful computer software in teaching several aspects of both segmental and suprasegmental pronunciation.

The results of our study showed that Praat could enable students refine and improve their pronunciation, mainly as they can record their own sound samples and have an instant and immediate visual feedback. This allowed the students to actually observe the pronunciation errors which may not be noticed only through listening.

Limitations of the Study

Although the research has achieved its aims, there were some unavoidable limitations in undertaking it. These limitations were eclipsed by the benefits of the study conducted.

Since we have introduced a new technique of pronunciation teaching using Praat software, time was a big constraint and challenge. The time allotted to phonetic sessions was insufficient to manage teaching and training students.

This research was conducted only on a small population, because we could not devote more time explaining to individual students how to work with the Praat software. Therefore, to generalize the results for larger groups, the study could have involved more students and possibly at different levels.

We concede the great difficulties encountered during the fulfilment of this work, mainly when recording students for it was almost impossible to find a quiet and appropriate place for a best quality recording.

Although some students brought their computers with the Praat Software installed, others did have neither computers nor internet access. This might have discouraged students' interest and motivation to practice their pronunciation at home.

The recording procedure took place in our office; as mentioned earlier, therefore, we could not avoid noisy corridors. An anechoic chamber would allow a very good quality of recording when designing acoustic treatments.

As, there was no research material available in the department, we have used our own materials. Various aspects of the students' training could have been covered if a language laboratory with some computers had been available.

Further research needs to be conducted implementing some tests of discriminations and identifications on the students' production and perception of some specific segmental or suprasegmental features.

The Praat Software is unquestionably a very beneficial computer tool, providing both articulatory and acoustic analysis of speech. It needs to be thoroughly explored mainly as it is freely available and easily downloadable for both students and teachers.

Suggestions and Perspectives

Our perspectives are that we would like to feel a tremendous sense of accomplishment by deeply exploring spectrograms readings, and how to master the Praat's scripts that actually provide a very large amount of cues and functions, which automatically analyse the pronunciation acoustically taking into account the sound frequency, its intensity, its duration, and its pitch (F0) etc. The Praat scripts not only facilitate measuring the formants (vowels) and their transitions but also offer phonetic annotations.

References

References

Abercrombie, D. (1967). *Elements of General Phonetics.* Edinburgh University Press.

Asher, J. J. (1977). *Learning language through actions: The complete teacher's guide book*, Los Gatos, CA:Sky Oaks Production.

Ausubel, D.P. (1964). Adults versus Children in Second Language Learning: Psychological considerations. *Modern Language Journal*, 48, 420-424.

Baltaxe, C.A.M (1978) *Foundations of distinctive feature theory*. Baltimore: University Park Press.

Bell, A. G. (1916) *The mechanisms of speech*. Funk and Wagnalls, New York.

Beckman, M.& Elam, A. (1997) *Guidelines for ToBI labeling*, version 3: Ohio State University.

Bloomfield, L (1933).*Language*. First published in Great Britain. New York University Press (P.79. 91)

Blair, R.W. (1991). Innovative approaches to Language Teaching. In M. Celce-Maurcia (Ed), *Teaching English as a second or foreign language*, 2nd ed. (pp. 23-45).New York: Newbury House

Boyd, B.L., &Murphrey T.P. (2002). Evaluation of a computer based, asynchronous activity on student learning of leadership concepts. *Journal of Agricultural Education*, 43 (1), 35-37

Bowen, J.D. (1972). Contextualizing pronunciation practice in the ESOL classroom. *TESOL Quarterly*, 6(1) 83-94

Bowen, J.D. (1975a). An experimental integrative test of English grammar. *UCLA Workpapers in Teaching English as a second language, 9, 3-15.*

Bowen, J.D. (1975b). *Patterns of English pronunciation*. Rowley, MA: Newbury House.

Bowen, T.& Marks, J. (1992) *The Pronunciation Book: student-centered activities for pronunciation work.* Burnt Mill, Harlow: Longman.

Breitkreutz, J. Derwing, T. & M. Rossiter (2002). Pronunciation teaching practices in Canada. *TESL Canada Journal*, 19, 51–61.

Brown, H. D. (2001). *Teaching by principles: An interactive approach to language*. New York: Longman.

Brown, H.D. (1994). *Principles of language learning and teaching*, 3rd ed. Englewood Cliffs, NJ: Prentice-Hall.

Bruen, J. (2001a). The parallel development of oral proficiency an use of language learning strategies. *Teaching German*, 34(2), 158-168.

Brumfit, C. J., & Johnson, K. (Eds.).(1979).*The communicative approach to language teaching*. Oxford: Oxford University Press.

Burgess, J. & S. Spencer (2000). *Phonology and pronunciation in integrated language teaching and teacher education*, *System*, 28, 191–215.

Carroll, J. B., (1962). The prediction of success in intensive foreign language training. In R.Glaser (Ed.), *Training, research, and education* (pp.87-136). Pittsburgh: University of Pittsburg Press.

Carroll, J. B., (1981). Twenty-five years of research in foreign language aptitude. In K.C. Diller (ed.), *Individual differences and universals in language learning aptitude*. (pp.83-118.) Rowley, MA: Newbury House.

Celce-Murcia, M., (1983) Teaching pronunciation communicatively. *MEXTESOL Journal*, 7(1), 10-25

Celce-Murcia., M., (1987) Teaching pronunciation as communication. In J. Morely (Ed.), *Current perspectives on pronunciation: Practices anchored in theory (pp. 5-12). Alexandria, VA: TESOL*

Celce-Murcia, M., &, Goodwin, J. (1991). *Teaching Pronunciation*: In M, Celce-Murcia (Ed.), *Teaching English as a second or foreign language, 2nd ed. (pp.136-153).* New York: Newbury House.

Celce-Murcia, M., Brinton, D.M., & Goodwin, J.M. (1996).*Teaching pronunciation: A reference for teachers of English to speakers of other languages*. New York: Cambridge University Press.

Chamot, A. U., & Rubin, J. (1994). Comments on Janie Rees-Miller's "A critical appraisal of learner training: Theoretical bases and teaching implications". *TESOL Quarterly*, 28, 771-776.

Chamot, A. U., & El-Dinary, P.B. (1999). Children's learning strategies in immersion classrooms. *The Modern Language Journal*. 83(3), 319 – 341.

Chomsky.N. (1957) *Syntactic Structure*. Mouton Publishers. The Hague.

Chomsky. N. (1965) *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.

Chomsky. N., & Halle, M. (1968).*The Sound Pattern of English*. New York: Harper & Row.

Chun, D. (1998) *Signal analysis software for teaching discourse intonation. Language learning and technology, 74-93*

Clark, J. &Yallop, C. (1995) *An Introduction to phonetics and phonology*. 2nd edition. Oxford: Blackwell (p.181)

Claudwell, R. (2002). Streaming Speech: Listening and Pronunciation for Advanced Learners of English. Birmingham: *Speech in Action*,. Windows CD-ROM

Cohen, L., Manion, L., & Morrison, K. (2007).*Research methods in education* (6th Ed.). New York: Routledge Falmer.

Coniam, D. (2002). Technology as an awareness raising tool for sensitising teachers to features of stress and rhythm in English. *Language Awareness*, 11 (1), 30–42.

Crystal, D., (1969). *Prosodic systems and intonation in English*. Cambridge studies in Linguistics. Cambridge University Press.Crystal, D., (1971). *Linguistics*. Baltimore: Penguin.

Cruttenden, A. (1986). *Intonation*. Cambridge, Cambridge University Press.

Cruttenden, A. (2008). *Gimson's pronunciation of English*.7th Edition. Hodder Education.

Crystal, D., (1969). Prosodic systems and intonation in English.*Cambridge studies in linguistics*. Cambridge University Press.

Crystal, D., (1971). *Linguistics*. Baltimore: Penguin.

Crystal, D., (1997). *The Cambridge Encyclopedia of Language 2ndEdition*. Cambridge University Press.

Crystal, D. (2003). *The Cambridge Encyclopedia of the English Language*. 2nd Ed. Cambridge University Press.

Cucchiarini, C, H., &Boves, L. (2000).Different aspects of expert pronunciation quality rating and their relation to scores produced by speech recognition algorithms.*Speech communication* 30, 145-166

Curran, C.A., (1976). *Counseling-Learning in second language learning*. East Dubuque, IL: Counseling-Learning Publications.

Dalton,C.&B.Seidlhofer .(1994). *Pronunciation.Language Teaching.A scheme for teacher education*.Oxford University Press.

De Lacy, P. (2002).The interaction of tone and stress in Optimality Theory.*Phonology* 19.1: 1-32

Deldime, R. & Demoulin, R. (1975).*Introduction à la psychopédagogie: à l'usage des étudiants: Guide Méthodologique, exercices, référentiel théorique*. Bruxelles: De Boeck.

Denes, P., & Pinson, E,. (1993). *The Speech Chain: The Physics and Biology of Spoken Language*. New York: Worth Publishers

Diamond, M. C.(1988).*Enriching heredity: The impact of the environment on the anatomy of the brain*. New York: Free Press

Derwing, T. M. (2010). *Utopian Goals for Pronunciation*. In J. Levis & K. LeVelle (Eds.), *Proceedings of the 1st Pronunciation in Second Language Learning and Teaching Conference,* Iowa State University, Sept. 2009. (pp. 24-37), Ames, IA: Iowa State University.

Derwing, T.M., &Rossiter M.J. (2001) *ESL learners' perceptions of their pronunciation needs and strategies.* System 30(2), 155-166.

Derwing, T. M., Munro, M. J., &Carbonaro, M. (2000). Does popular speech recognition software work with ESL speech? *TESOL Quarterly*, 34(3), 592-603.

De Saussure, F. (1971) *Cours de Linguistique Générale. 5eme Ed. Bally et Schehaye. Payot. Paris*

Dörnyei, Z. &Skehan, P. (2003).Individual Differences in Second Language Learning. In *the handbook of second language acquisition*, edited by Catherine Doughty and Michael Long, M. (Eds), 589-630. Oxford: Blackwell.

Duez, D. (1982). Silent and non-silent pauses in three speech styles. p.13) *Language and Speech*

Ehsani, F., &Knodt, E. (1998) Speech technology in computer aided language learning: Strengths and limitations of a new CALL paradigm. *Language Learning & Technology July 1998, Volume 2, Number 1.*

Eimas, P (1971) Speech perception in early infancy. *Science, New series. Volume 171, No; 3968. 304-306.*

El-Dib, M. A. B. (2004). Language learning strategies in Kuwait: Links to gender, language level, and culture in hybrid context. *Foreign Language Annals*, 37, 85-95.

Encyclopedia Britannica inc, (1994) William Benton Publisher.

Eskenazi M. (1999) *Using automatic Speech Processing for Foreign Language Pronunciation Tutoring: Some issues and Prototype, Language Learning & Technology,* 2/2: 62-67

Fan, M. Y. (2003). Frequency of use, perceived usefulness, and actual usefulness of second language vocabulary strategies: A study of Hong Kong learners. *Modern Language Journal*, 87(2), 222-241.

Fant, G. (1973). *Speech sounds and features*. Cambridge, MA: MIT Press.

Flege, J. E. (1981). The phonological basis of foreign accents: A hypothesis. *TESOL Quarterly*, 15(4), 443-455.

Florez, M. C. (1998). Improving adult ESL learners' pronunciation skills. *National Center for ESL Literacy Education*. Retrieved July 11, 2005 from http://www.cal.org/caela/digests/Pronun.htm

Fraser, H. (1999) *ESL Pronunciation Teaching: could it be more effective?* Presented at the ALAA Conference, Perth.

Gardner, R. C., & Lambert, W. E. (1972). *Attitudes and motivation in second language learning*. Rowley, MA: Newbury House

Gattegno, C. (1972).*Teaching foreign languages in schools: The Silent Way*. New York: Educational Solutions, 1972)

Gattegno, C. (1976). *The common sense of teaching foreign languages.* New York: Educational Solutions.

Gattegno, C. (1985a). *The learning and teaching of foreign languages*. New York: Educational Solutions.

Gay, L. R. (1992). *Educational research* (4th Ed.). New York: Merrill. (p.298).

Gilbert, Judi. B. (2008). *Teaching Pronunciation: Using the Prosody Pyramid*. Cambridge University Press.

Graham, R. C. (1985). *Beyond integrative motivation: The development and influence of assimilative motivation.* In P. Larson, E. L. Judd, & D.S. Messerschmidt (Eds.), *On TESOL'84* (pp.75-87). Alexandria, VA: TESOL.

Guiora A. Z. (1972), Construct validity and transpositional research: Toward an empirical study of psychoanalytic concepts. *Comprehensive Psychiatry* 13(2), 139-150

Guiora, A. Z., Beit-Hallami, B., Brannon, R. C., Dull, C. Y. and Scovel, T. (1972). *The effects of experimentally induced changes in ego stages on pronunciation ability in a second language: an exploratory study. Comprehensive Psychiatry 13, 421-428.*

Gussenhoven, C. (1999) Discreteness and Gradience in Intonation Contrasts. *Language and speech.* 1999.42 (2-3), 286

Gussenhoven, C. (2002). *Intonation and interpretation: Phonetics and Phonology*. Speech Prosody 2002: *Proceedings of the First International Conference on Speech Prosody*

Gussenhoven, C. (2004). *The Phonology of tone and intonation*. Research Surveys in Linguistics.Cambridge University Press.

Harmer, J. (2001).The practice of English language teaching.England:Longman.

Hardison, D. (2004). Contextualized computer-based L2 prosody training: Evaluating the effects of discourse context and video input. *CALICO Journal*, 22(2), 175-190

Hinofotis, F., & Bailey, K. (1980).American undergraduates' reactions to the communication skills of foreign teaching assistants. In: J. C. Fisher, M. A. Clarke, & J. Schachter (Eds.), *On TESOL '80* (pp. 120-133). Washington, DC: Teachers of English to Speakers of other languages.

Hirata, Y. (2004). Computer-assisted pronunciation training for native English speakers learning Japanese pitch and duration contrasts. *Computer Assisted Language Learning*, 17 (3–4), 357–376.

How to read a spectrogram? https://home.cc.umanitoba.ca/~robh/howto.html

Howatt, A.P. R. (1984) *A history of English language teaching*. Oxford. Oxford University Press.

Honig, F. Batliner, A,Weilhammer, K. Noeth, E.(2009) *Employing word accent information for pronunciation quality assessment of English L2 learners.* SLaTE 2009, Birmingham. England.

Jacobson, R. (1931) Phonemic Notes on Standard Slovak. *Selected Writings. I. Phonological Studies. 3rd Edition. Mouton. The Hague.*

Jacobson, R. (1938) *Observations sur le Classement phonologique des consonnes.* Proceedings of the 3rd International Congress of Phonetic Sciences, ed. by Edgar Blancquert and William Pee. Laboratory of Phonetics of the Universit.

Jacobson, R. book C. Gunnar, M. Font. M, Halle (1952) "*Preliminaries to Speech Analysis*: *The distinctive Features and their correlates.* Massachusetts Institute of Technology Cambridge, Massachusetts

Jakobson, R.& Morris, H. (1956). *Fundamentals of Language* (The Hague,); Germ, trans, by G. F. Meier Grundlagen der Sprache (Berlin, 1962) (1938/1962, p.274)

Jakobson, R. (1941). Child language, aphasia, and phonological universals. A. R. Keiler,trans. The Hague: Mouton.

Jakobson, R. (1949) *On the Identification of phonemic Entitie.* Kopenhagen : Nordisk Sprog- og Kulturforlag.

Jakobson, R.(1952) *Notes on the French Phonemic Pattern.* Selected Writings. *I. 1962 by R. Jakobson.* The Hague. Mouton.

Jacobs, B. (1988). Neurobiological differentiation of primary and secondary language acquisition. Studies in Second Language Acquisition, 10,303-337.

Jenkins, J. (2000).*The Phonology of English as an international language*. Oxford University Press.

Jenkins, J. (1998).Which pronunciation norms and models for English as an international language.*ELT Journal*, 52/2.

Jones, D.(1972). An outline of English phonetics, *9th Ed.* Heffer. Cambridge

Jones, D. (1957). *An outline of English phonetics*. Cambridge: Heffer& Sons, 8Edition.

Jones, D. (1991). *English pronouncing dictionary*, 14thed. Cambridge: Cambridge Univeristy Press.

Judith Weidman's class at the American Language Center of UCLA Extension.

Kawai, G. & Hirose, K. (2000).Teaching the pronunciation of Japanese double-mora phonemes using speech recognition technology. *Speech Communication*, 30, 131–143.

Kelly, L.G. (1969).*25 centuries of language teaching*. Rowley, MA: Newbury House. Oxford University Press

Kenworthy, J. (1987).*Teaching English Pronunciation*.Longman Handbooks for Language Teachers.Longman London and New York.

Krashen, S. D. (1973). Lateralization, language learning, and the critical period: Some new evidence. *Language learning*, 23, 63-74.

Krashen, S. D. (1982). *Principles and practice in second language acquisition*. Oxford: Pergamon Institute.

Krashen, S. D. & Terrell, T.D. (1983) *The Natural Approach*. Hayward, CA: Alemany Press.

Kolb, D. (1984). *Experiential learning: Experience as the source of learning and development*. Englewood Cliffs, New Jersy: Prentice-Hall. K

Komissarchik, E.et al (2000). *Application of Knowledge- based Speech Analysis to Suprasegmental Pronunciation*, Proc. AVIOS.2000:243-248

Kul, P. K.,& J.D. Miller (1978) Speech Perception by the chinchilla: Identification functions for synthetic VOT stimuli. *Journal of the Acoustical society of America 63, 905-917*

Ladefoged, P. (1967). *Three areas of experimental phonetics*. Oxford University Press.

Ladefoged, P. (2001). *Vowels and consonants. An introduction to the sounds of languages*. Oxford: Blackwell (chapters 4-6).

Ladefoged, P., K. Johnson (2006). *A course in Phonetics*. 7th Ed. Product Director: Monica Eckman Product Manager: Kate Derrick. Library of Congress. USA. Cengage Learning WCN: 02-200-208.

Ladd, D. R. (1980).*The structure of intonational meaning*. Bloomington: Indiana University Linguistics Club. (p.26)

Lambacher, S. (1999).A CALL tool for improving second language acquisition of English consonants by Japanese learners.*Computer Assisted Language Learning*, 12 (2), 137–156.

Larsen-Freeman, D. (1986) *Techniques and Principles in Language Teaching*. Oxford University Press

Larsen-Freeman, D. and Long, M. H. (1991), an *Introduction to Second Language Acquisition Research*. London and New York: Longman.

Lass, N. J. (1996) The Acoustic Characteristics of American English. In N. J. Lass (ed.), *Principles of Experimental Phonetics*, pp. 185-225.

Lasnik, Howard, Daniel N. Osherson. (1990) *An Invitation to Cognitive Science* Vol 1: Language. Cambridge. MIT Press. 1st Edition 1990. *Speech Perception*. Joanne L. Miller.

Leather, J., & James, A. (1991).The acquisition of second language speech. *Studies in Second Language Acquisition*, 13, 305-341.

Lieberman, P. & Blumstein, S. E. (1988). Speech physiology, speech perception, and acoustic phonetics. Cambridge: Cambridge University Press.

Liberman, A. M. & Mattingly, I.G. (1985) The Motor Theory of Speech Perception Revised". Cognition, 21 (1985) 1-36. Haskins Laboratories, University of Connecticut

Lehiste, I. (1970). *Suprasegmentals*. Cambridge, Mass: M.I.T. Press.

Lehiste, I., & Peterson, G, E. (1961). Transitions, glides, and diphthongs. Journal of the Acoustical Society of America, 33(3), 268-272.

Lehiste, I., & Peterson, G. (1959).Vowel amplitude and phonemic stress in American English. *Journal of the Acoustical Society of America*. 31

Lennberg, E. (1967). *Biological foundations of language*. New York: Wiley

Levis, J. (2007) '*Computer Technology in Teaching and Researching Pronunciation'*, Annual Review of Applied Linguistics, 27, pp. 184–202. doi: 10.1017/S0267190508070098.

Levis, J., (2005). Changing contexts and shifting paradigms in pronunciation teaching, 369-378. Volume 39, Issue 3. *TESOL Quarterly*

Liberman, A. M. & Mattingly, I.G. (1985) *The Motor Theory of Speech Perception Revised*". Cognition, 21 (1985) 1-36. Haskins Laboratories, University of Connecticut 378

Lieberman, P. & Blumstein, S. E. (1988).*Speech physiology, speech perception, and acoustic phonetics*. Cambridge: Cambridge University Press.

Levow, G. A. (2009). *Investigating Pitch Accent Recognition in non-native speech, ACL 2009, Singapore, 2009*

Loos et al (1997) What is a phone? LingualLinks: *Glossary of linguistic terms. SIL International. Retrieved 16 December 2013.*

Lukmani, Y. (1972). Motivation to learn and learning proficiency. *Language Learning*, 22, 261-273.

Macdonald, D.Yule, G& Powers (1994).The different effects of pronunciation teaching. *International Review of Applied Linguistics in Language Teaching* (IRAL), 33(4), 345-350.

Macdonald, D.; Yule, G. & Powers, M. (1994). Attempts to improve English L2 pronunciation: The variable effects of different types of instruction. *Language Learning*, 44(1), 75-100.

MacDonald, S. (2002). Pronunciation–views and practices of reluctant teachers. Prospect, 17(3). Retrieved April 10, 2007, from http://nceltr.edu.au/ prospect/17/pros17_3smac.asp

Martin, P. (2008).*Phonétique acoustique: Introduction à l'analyse acoustique de la parole*. Armand Colin.

McGurk, H and J. MacDonald (1976) Hearing lips and seeing voices. *Nature 264, 746-748*

Massaro, D. W. (1987).Speech perception by eye and ear: A paradigm for psychological inquiry: *Program in Experimental Psychology*, University of California, Santa Cruz

Massaro, D.W & M.M. Cohen, (1983) Evaluation and integration of visual and auditory information in speech perception. *Journal of experimental psychology: Human perception and performance 9, 753-771.*

Mateescu, Dan (2003) *English phonetics and phonological theory*. 20 th Century Approaches. Editura Universitatii din Bucuresti. McNerney, M & D. Mendelsohn (1992). Suprasegmentals in the pronunciation class: Setting priorities. In: Avery, Peter & Susan Ehrlich (Eds.). *Teaching American English pronunciation* Oxford: Oxford University Press, (pp185–196).

McNerney, M & D. Mendelsohn (1992). Suprasegmentals in the pronunciation class: Setting priorities. In: Avery, Peter & Susan Ehrlich (Eds.). *Teaching American English pronunciation* Oxford: Oxford University Press, (pp185–196).

Molholt, G. (1988). Computer assisted instruction in pronunciation for Chinese speakers of American English. *TESOL Quarterly*, 22 (1), 91-113.

Morley, J. (Ed).(1987). *Current perspectives on pronunciation*. Alexandria, VA: TESOL.

Morley, J. (1991). The pronunciation component in teaching English to speakers of other languages. *TESOL Quarterly* 25(3), 481–520.

Morely, J. (1994a).Pronunciation pedagogy and theory: new views, new dimensions. Alexandria, VA: TESOL.

Munro, M. & Derwing, T. (1999).*Foreign accent, comprehensibility, and intelligibility in the speech of second language learners*. Language Learning, 285-310

Murphy, J. (1997). Phonology courses offered by MATESOL programs in the United States. *TESOL Quarterly*, 31 (4), 741–764.

Neri, A., Cucchiarini, C., &Strik, H. (2002). Feedback in computer assisted pronunciation training: When technology meets pedagogy. *Proceeding from CALL Professionals and the Future of CALL Research.*

Neri, A., Cucchiarini, C., &Strik, H. (2003). Automatic speech recognition for second language learning: How and why it actually works?

Neri , A. Cucchiarini, C. , Strik, H. &Boves, L. (2002). The pedagogy-technology interface in computer assisted pronunciation training. Computer Assisted Language Learning, 15 (5), 441–467.

Neisser, U. (1967).*Cognitive Psychology*. New York: Appleton-Century-Crofts.

Nilsen, D.L.F., &Nilsen, A.P. (1973) *Pronunciation contrasts in English*. New York: Regents.

Neumyer,L., Franco, H., Digalakis, V., Weintraub, M (2000). Automatic scoring of pronunciation quality.*Speech communication* 30 (2000), 83-93.

Neumeyer, L. Franco ,H. Digalakis ,. V. & Ronen, O. (2000). Combination of machine scores for automatic grading of pronunciation quality. Speech Communication, 30, 121–130.

Norton, B., &Toohey, K. (2001).Changing perspectives on good language learners.*TESOL Quarterly*, 35(2), 307-322.

O'Connor, J.D.(1980). Better English Pronunciation. Cambridge English Language Learning. Cambridge University Press.

Otlowski, M., Fraser, H. (1999) ESL pronunciation teaching: Could it be more effective? *Australian Language Matters*.

Oxford English Dictionary. (2003) 2nd Ed. Oxford University Press.

Penfield, W., & Roberts, L. (1959). *Speech and brain mechanisms*. Princeton, NJ: Princeton University Press.

Pennington, M.C. (1994). *Recent research in L2 phonology: implications for practice*. In J. Morley (Ed). Pronunciation pedagogy and theory: New views, new dimensions (pp. 92-108). Alexandria, VA: *TESOL.*

Pennington, M.C. (1999) Computer –aided pronunciation pedagogy: promise, limitations and directions. *Computer assisted language learning*, vol 12, pp427 -440

Pennington, M. & Richards, J. (1986) Pronunciation Revisited. *TESOL Quarterly* 20(2): pp 207-225

Peter, B. D., Pinson, E.N. (1993).The Speech Chain: The Physics and Biology of Spoken Language. Worth Publishers

Peterson, S.S. (2000). Pronunciation learning strategies: A first look. Unpublished research report. (*ERIC Document Reproduction Service* ED 450 599; FL 026. 618).

Pierrehumbert, J. B. (1980) *The Phonetics and Phonology of English Intonation*, PhD dissertation. MIT. Published by Garland Press, New York, 1990.

Pica, T. (1984).Pronunciation activities with an accent on communication. *English Teaching Forum*, 22(3), 2-6.

Picket, J.(1980) *The sounds of speech communication*. Baltimore, MD: University Park Press.

Postovsky, V. (1974).The effects of delay in oral practice at the beginning of second language learning. *Modern Language Journal*, 58 (3), 229-239.

Prater, C. H., (1991).The cornerstones of method. In M. Celce-Murcia (Ed.), *Teaching English as a second or foreign language*, 2nd ed. (pp11-22). Boston: Heinle&Heinle.

Price, P. (1998). "How can Speech Technology Replicate and Complement Good Language Teachers to Help People Learn Language?" STILL, 1998: 81-90.

ProteaTextware. (2007). Retrieved July. http://www.proteatextware.com/

Raux, A., & Kawahara, T. (2002).*Automatic intelligibility assessment and diagnosis of critical pronunciation errors for computer-assisted pronunciation learning. In Proc. ICSLP.*

Ranalli, M. J. (2002).*Discourse Intonation: To teach or not to teach?* University of Birmingham

Riney, T. J., &Flege, J. E. (1998).Changes over time in global foreign accent and liquid identifiability and accuracy. *Studies in Second Language Acquisition*, 20,213-243.

Riney, T. J., Takada, M., & Ota, M. (2000).Segmentals and global foreign accent: The Japanese flap in EFL. *TESOL Quarterly*, 34(4), 711-737.

Roach, P., (1983).*English Phonetics and Phonology*. Cambridge University Press.

Roach, P., (1991) *English Phonetics and Phonology*: a practical course (2nd ed). Cambridge University Press.

Roach, P., (2009) *English Phonetics and Phonology*. A practical course. Cambridge: Cambridge University Press

Rogers, C. R. (1951). *Client-centered therapy: Its current practice, implications, and theory*. Boston: Houghton Mifflin. Saladin, Kenneth, S. (2011) Human Anatomy. Fourth Edition. Georgia College & State University.

Setter, J., & Jenkins, J. (2005).State of the art review article: *Pronunciation. Language Teaching*. Vol. 38, N.1.

Schumann, J.H. (1975). Affective factors and the problem of age in second language acquisition. *Language Learning*.25 (2),209-235.

Schumann, J.H. (1986). Research on the acculturation model for second language acquisition. *Journal of Multilingual and Multicultural Development*,7(5),379-392

Scott, M. L., (1989). *Younger and older adult second language learners: A comparison of auditory memory and perception*. PhD. Dissertation, University of California, Los Angeles.

Scovel, T.(1969).Foreign accents: language acquisition, and cerebral dominance. *Language Learning*.Volume 19.Issue 3-4, Pages 245-254. University of Michigan.

Scovel, T. (1988). *A time to speak: a psycholinguistic inquiry into the critical period for human speech.* New York, NY: Newbury House Publishers.

Shadle, C.H. (1990) *Articulatory-Acoustic Relationships in Fricative Consonants*. In W.J. Hardcastle and A. Marchal (eds.), *Speech Production and Speech Modelling*, pp. 187-209.

Singh, S. & Singh, K,. (1976) *Phonetics*: Principles and practices. Baltimore, MD; University Park Press.

Skehan, P., (1989). *Individual differences in second language learning*. Edited by: Peter Robinson. John Benjamins Publishing Company. Amsterdam/ Philadelphia

Skehan, P. (1998). *A cognitive Approach to Language Learning*. Oxford University Press.

Snow, M. A., &Shapira, R. G. (1985).The role of social-psychological factors in second language learning. In M. Celce-Murcia (Ed.), *Beyond basics: Issues and research in TESOL* pp. 3-15). Rowley, Massachusetts: Newbury House Publishers.

Stern, H. H. (1992). *Issues and options in language teaching*. Oxford University Press.

Stern, H. H. (1983). *Fundamental concepts of language teaching*. Oxford University Press.

Stetson, R.H (1951) *Motor phonetics. A study of speech movements in action*. University of Michigan.

Stevick, E.W. (1976).M*emory, meaning and method: Some psychological perspectives on language learning*. Rowley, MA Newbury House

Stevick, E. (1980) *Teaching Languages: A way and ways.* Rowley MA. Newbury house.

Steven, K.N. (1998). Acoustic phonetics, MIT Press. *Speech communication*. Spring 2004.

Stuckless, R. (1994). Developments in real-time speech-to-text communication for people with impaired hearing. In M. Ross (Ed.), *Communication access for people with hearing loss* (pp.197-226). Baltimore, MD: York Press

Trubetzkoy, N (1938) *Grundzüge Der Phonologie. English.* University of California Press, 1969.

Trubetzkoy, N. (1939) *Principles of Phonology*. *Grundzüge Der Phonologie.* Translated By Christiane A.M. Baltaxe. Berkeley and Los Angeles. University of California Press, 1969 Los Angeles.

Truong, K., Neri, A., de Wet, F., Cucchiarini, C., & Strik, H. 2005. *Automatic detection of frequent pronunciation errors made by L2 learners*. Proceedings of Inter Speech (pp. 1345–1348). Lisbon, Portugal.

Van Compernolle, D. (2001) Recognizing speech of goats, wolves, sheep and …non-natives. *Speech communication*, 35, 71-79

Vaissière. J. (2007). *La phonétique* .Que sais- je ? Presses Universitaires de France.

Vardanian, R., Marie, (1964). *Teaching English Intonation through Oscilloscope* Displays. University of California.1964

Vitanova, G., & Miller, A. (2002). Reflective practice in pronunciation learning. *The Internet TESL Journal*, 8(1). Retrieved July 8, 2005, from: http://iteslj.org/

Universal Declaration of Human Rights. Article one

Wang, X. & M. Munro (2004). Computer-based training for learning English vowel contrasts. *System*, 32, 539–552.

Widdowson, H. (1978). *Teaching language as communication*. Oxford: Oxford University Press.

Wilson, l. (2005).Using Praat and Moodle for teaching segmental and suprasegmental pronunciation.*TESOL Quarterly, 3.*

Wong, R.(1987a). *Teaching Pronunciation: Focus on English Rhythm and Intonation***.** Center for Applied Linguistics, 1987

Wong, R. (1987b). Learner variables and prepronunciation considerations in teaching pronunciation.In J. Morley (Ed), *Current perspectives on pronunciation* (pp. 13-28). Alexandria, VA: TESOL.

Webster's Collegiate Dictionary (1975) by Merriam Webster Editors.

Yip, M. (2002) *Tone*. Cambridge & New York: Cambridge University Press.

Yule, G. (2006). *The study of language*. Fourth edition. Cambridge University Press.

Webography:

http://projects.iq.harvard.edu/cb45/pages/diagram-organs-speech-replacement

https://www.languagesoftware.net/tellmemore-review.html

http://www.canadavisa.com/rosetta-stone-language-learning-program.html

http://www-01.sil.org/computing/sa/index.htm

https://www.englishclub.com/pronunciation/terms.htm#vowel

Appendices

1.Questionnaire

Dear students,

We are currently conducting an investigation on innovations in practical pronunciation teaching using some software, which displays automatic and visual feedbacks.

The main purpose is to obtain your attitudes towards pronunciation as a skill and examine the way it is taught in our department.

Thank you for completing this questionnaire. Your answers will be strictly confidential and will be used only for the purpose of academic research.

General Instruction:

- Please, tick the appropriate box :(√)
- Where extended response is required, please write up in the space provided.

1. Sex: Male ☐

Female ☐

2. What is your current age? ☐

3.Are you satisfied with your English Pronunciation?

Yes ☐

No ☐

Why?

....

....

....

4. Which pronunciation do you prefer?

American ☐

British ☐

Why?

....

....

....

5. How important is pronunciation to you?

Very Important	Important	Less important	Not important

Why?

....

....

....

6. Do you practice your English outside the class?

Yes ☐

No ☐

Why?

....

....

....

7. What do you do to improve your pronunciation?

....

....

....

....

....

8. When you speak English, do you use what you learned in Phonetics lectures?

Yes ☐

No ☐

If yes, how often?

....

....

....

If no, why not?

....

....

....

9. Do you like to have pronunciation teaching classes?

Like	Undecided	Dislike

10. Do you agree that pronunciation teaching should be given more importance and must be included as an independent part of preparation programs at our universities?

Agree	Undecided	Disagree

11. Do you encounter difficulties when you speak or listen in English because of pronunciation problems?

Yes ☐

No ☐

12. In which of the following situations do you encounter pronunciation difficulties?

A. When I watch movies ☐

B. When I chat with native speakers in English ☐

C. When I listen to foreign music in English ☐

D. When I listen to the radio and television broadcast ☐

E. When I watch the different news channels ☐

13. Have you ever heard about a spectrogram before your teacher introduces it?

Yes ☐

No ☐

14. While dealing with the thePraat Software in classroom and explanation of its use as a new tool in teaching an accurate pronunciation. How do you evaluate it?

Very important	Important	less important	Not important

15. After having been introduced to the thePraat software, do you keep exercising at home?

Yes ☐

No ☐

If yes, how often?

Every day	Once a week	When I have time	Before I have phonetics lectures

16. Given more time to practice, do you think this program would help you improve your pronunciation?

Yes ☐

No ☐

17. After practicing several times in classroom and at home -if at all- are you able to be aware of the correctness or incorrectness of your own pronunciation?

Yes ☐

No ☐

18. When learning English pronunciation, I keep working until I reach the goals that I make for myself.

Strongly agree	Agree	Partly Agree	Slightly disagree	Disagree strongly	Disagree

19. When you are conversing with English speakers, do you ask for feedback on your English pronunciation?

Yes ☐

No ☐

20. When you aren't sure of your English pronunciation, what additional methods do you use to correct yourself?

……………………………………………………………………………………………

……………………………………………………………………………………………

……………………………………………………………………………………………

……………………………………………………………………………………………

21. When you do not know how to pronounce a word in English, do you ask your teachers for help?

Yes ☐

No ☐

Thank you very much for your participation in this study.

Lynda Badache

2.Teachers' Interview

a. Pronunciation

- Do you assess your students' speaking skills?
- What exactly do you assess and how?
- What do you use in classroom as activities for speaking?
- When students 'speaking is very difficult to be grasped, what other ways used as a remedy?

b**. Teaching Content**

While planning the speaking activities, can you alone, as a teacher, decide the content ofteaching to your students or have other considerations?

How much time is devoted to pronunciation?

c. Teachers'Perspective, Feelings, Opinions and Practices

How important is pronunciation?

Do you teach pronunciation? Why? Why not?

If you teach pronunciation, which aspects do you focus on?

Do you feel confident when you teach it?

Which pronunciation do you use? The American or the British?Or the mixture of both and why?

Which pronunciation materials are you aware of or familiar with?

Which do you use? Why?

Have you heard about the thePraat software before?

Have you an idea about what a spectrogram is?

a) Do you have a prior knowledge in the acoustic phonetics?

b) According to you, why phonetics represents a ghost for teachers?

c) What are your suggestions for teachers who teach oral expression and phonetics?

e.The Interviews Transcripts

Mr. Tahar. K

The researcher (Question) **Do you assess your students' speaking skills?**

The Interviewee (Answer) Yes, I do. Because usually the question which is asked to students at the University is: do you speak English and that the speaking skill that is considered as a measure to know whether a student masters English or not.

The Researcher (Question) **What and how do you assess?**

The Interviewee (Answer) Mum, well I do assess them without following a real certain criteria but I try to be available when and where necessary by providing a feedback and usually the feedback must be done in an encouraging discrete way.

The Researcher (Question) **How do you report /note down that assessment?**

The Interviewee (Answer) I have a very easy way to record assessment sometimes by an asterisk, by plus sometimes by a letter which I know what does it stand for.

The Researcher (Question) **What activities do you use in class for speaking?**

The Interviewee (Answer) Well, not real activities but sometimes a broadcast that will help them get into discussion which will enable them to express themselves in situations that would look like daily life situations.

The Researcher (Question) **What do you do with students whose speaking is very difficult to be understood?**

The Interviewee (Answer) Well I do not have a real remedy a magic one, but I try to make them believe that it's something that it could be changed, it is something that they can better provide that they work hard.

The Researcher (Question) **Are there any external considerations when planning the speaking component of the course for your students or can you alone, as a teacher, decide what to teach to your group of students? (What are the external considerations, if any?)**

The Interviewee (Answer) No, for me I do not have external considerations but usually the teacher is in a position and he is, I guess the only one who knows what fits his students or his pupils and his choice is going to be that choice which will not be beyond that reach.

The Researcher (Question) In teaching/learning situations with prescribed curricula that you have some knowledge of, or experience in, to what extent is pronunciation a feature?

The Interviewee (Answer) Mum pronunciation is vital in learning any foreign language especially in our case which is English. For me I try always to correct, to help students figure out some phonetic rules that will help them pronounce better.

The Researcher (Question) How much time within these lectures is devoted to pronunciation?

The Interviewee (Answer) Well in my case normally all that lesson is going to be divided for pronunciation but sometimes you cannot do this because the overcrowded groups and not students will welcome being corrected all the time because this will interrupt the flow of the course.

The Researcher (Question) What emphasis do you feel should be placed on pronunciation? (How important is pronunciation?)

The Interviewee (Answer) Well as I said pronunciation is vital and it should be given enough attention not just by the teacher of oral expression but teachers of phonetics as well.

The Researcher (Question) Do you teach pronunciation?

The Interviewee (Answer) Well I do not teach pronunciation as such but I do never miss an occasion to correct my students' mispronunciation

The Researcher (Question) So, why don't you teach pronunciation?

Because I guessed that another teacher is doing the job and as a teacher of oral expression I try to see if those problems that need to be dealt with.

The Researcher (Question) **Which aspects do you focus on when it comes to you to correct students' mispronunciation?**

The Interviewee (Answer) I do not have a special area that I would focus on but I said earlier

The Researcher (Question) **I mean do you focus on the segmental or suprasegmental aspect of the language?**

The Interviewee (Answer) Well it depends on those situations where you as a teacher you must be available for your pupils and do your best to get them rid of their troubles related to pronunciation**.**

The Researcher (Question) **Some different approaches to pronunciation are listed below, which of these do you use? Describe when, how often and why you choose each/some/a particular method?**

- Whole class tasks/activity with pronunciation as the main focus
- Whole class tasks/activity with pronunciation as component.
- One to one situation
- Small groups of students
- Referral to a specialist teacher where available
- Other approaches?

The Interviewee (Answer) Well I think that I myself would be more comfortable when involving the whole class in tasks in which the principal activity is pronunciation because the class acts as a one little opportunity will be given to students in order to express themselves freely and here the teacher is going to act as a controller and when he is in control students will be denied, the opportunity to use their past experiences.

The Researcher (Question) **Which pronunciation materials are you aware of/familiar with?**

The Interviewee (Answer) In my case I did download lot of broadcast on BBC which I use as authentic material which I guess will help students have better pronunciation.

The Researcher (Question) **Why do you think that downloading these BBC broadcasting is really vital?**

The Interviewee (Answer) Like students will benefit better from being exposed to a formal people rather than being familiar with just the teacher and the more exposure they will have the better English their will have.

The Researcher (Question) **Have you heard about the thePraatSoftware?**

The Interviewee (Answer) No,

The Researcher (Question) **Do you have any knowledge in acoustic phonetics?**

The Interviewee (Answer) Just a definition

The Researcher (Question)According to you, why are many teachers afraid of phonetics?

The Interviewee (Answer) Well, may be because it is more technical.

The Researcher (Question) **What are your suggestions for teachers who teach oral expression and phonetics?**

The Interviewee (Answer) To have a continuous coordination between both teachers.

Ms.Radhia G

The Researcher (Question) **Do you assess your students' speaking skills?**

The Interviewee (Answer) I do,

The Researcher (Question) **How do you assess your students speaking Skill**?

The Interviewee (Answer) Evaluating my students 'speaking skill I take into consideration four parameters or four elements, the first one is fluency, accuracy vocabulary and pronunciation as well.

The Researcher (Question) **Whatactivities do you use in class for speaking?**

The Interviewee (Answer) Actually there are different types of activities, they could be labelled as interactive and communicative at the same time, I tend to vary from listening tasks to reading tasks and even the speaking tasks which is based on discussion, based on dialogues, based on for example role playing, and formation gap, jigsaw activities and other various activities mainly just to motivate them to speak.

The Researcher (Question) **Whatdo you do with students whose speaking is very difficult to be understood?**

The Interviewee (Answer) I do pay attention to those kinds of students and for me what matters for me is just to motivate them to push them to speak when they have any difficulty in speaking I mean what kind of speaking I want to focus, is it the pronunciation factor or is it the vocabulary factor, so I do provide students with tips, with hints for example some ways to improve for example I would tell them for example would you like to overcome your shyness while speaking so you should do this for example I would provide students with tips with regard to this if they have difficulty in pronouncing the word I show them how to pronounce the word correctly and fluently and practice is the best thing for me I give more opportunities for students to speak during the session.

The Researcher (Question) Whichpronunciation do you use? The American or the British?And why?

The Interviewee (Answer) Mypronunciation? Are you asking about my pronunciation? Actually if I can use the word “eclectic” in other words to say I do vary from American to British, wherever the words, for example if I feel myself comfortable when pronouncing the word in American way I would pronounce it such the American pronunciation but I feel myself much more comfortable pronouncing that word, different word in a British accent I would use the British accent so I do vary from American to British but I do tell my students that this word is pronounced like this in the American accent and this word is pronounced like this in the British accent.

The Researcher (Question) Whichpronunciation materials are you aware of or familiar with?

The Interviewee (Answer) I’m using... I base.. My work is based on the Cambridge online on my pronunciation system I mean I have software taken from Cambridge on my dictionary I work on that software so I have the two types of pronunciations the American and the British and I always listen to those pronunciations whenever I have time or whenever I work on any different words

The Researcher (Question) Have you heard about the thePraat software before?

The Interviewee (Answer) I've heard about it from you actually in one of our informal discussion

The Researcher (Question) Have you an idea what a spectrogram is?

The Interviewee (Answer) May be if I may say it's a software if I may call it , that would draw a line, the wave length or how a word is pronounced, and it would have its physical resonance, it's a drawing the curve of the sound how it is pronounced the sentence how it is pronounced etc

The Researcher (Question) Do you have a prior knowledge in the acoustic phonetics?

The Interviewee (Answer) Honestly I have not studied; if I am talking about studying it academically or I have no idea, I but just from my own knowledge about a prior knowledge, but not academically studying it.

The Researcher (Question) According to you, why phonetics represents a ghost for teachers in our department?

The Interviewee (Answer) For me I think the main reason, is because our teachers, we as teachers most of us were not exposed to native speakers environment, we do have a lack in getting in touch with, I mean how can we as foreign language teachers teach phonetics I mean academic pronunciation either the IPA pronunciation the British or the American without being exposed to this environment or without even travelling or living to that environment for at least one month or two months, we as teachers, we are not exposed, we did not have an opportunity to be a part of that environment, that authentic genuine environment, so they do lack, I mean thanks god nowadays we've got the web and we can expose our ears to that types of pronunciations but still the teachers cannot I mean are afraid of teaching phonetics because they do not know , they do not master we say it's better because they know but they do not master the appropriate rules of pronouncing correctly or appropriately and this is due to the lack of opportunities given to them to live in an authentic environment, to spend at least a good time in an authentic environment may be.

The Researcher (Question) **Do you relate phonetics as a module to pronunciation?**

The Interviewee (Answer) For me yeah it is part of it but it is not all of that, because when we say phonetics it has to do with even the anatomy if we may use the word, the anatomy of the speaking skill for example how the word is spoken we take into account the different organs and it is much more into we can even embark into physics in phonetics, so all these things we don't pay attention to this for us we focus in the teaching of phonetics on just pronunciation, there are some other factors that we do neglect here in our department such things the acoustic side or physics of this and so on and so forth.

The Researcher (Question) **What are your suggestions for teachers who teach oral expression and phonetics?**

The Interviewee (Answer) The first thing to suggest is to provide the appropriate environment for teachers and for students, the first thing is the equipments, the tools, the instruments; we do have teachers who do have capacities, students are willing are motivated to work but looking at the environment, we want to work with digital tools, with ICT tools, but those tools we don't find them so the environment is not somehow is not just encouraging us as teachers or even some students feel disappointed for not having that opportunity for them we do have computers we do have those ICT but we don't use them, we need in this department to provide the required environment by providing tools and instruments for both teachers and students.

The Researcher (Question) **Would like to teach phonetics as a module one day?**

The Interviewee (Answer) Yes why not, I would love to teach it, why not, because I am the kind of people who love, I mean we were supposed to be experts of the language and the side of pronunciation here is one of the sides that I am fond of I really appreciate talking a lot about pronunciation because you're dealing with a neural stuff you appreciate the rhythm, appreciate the way it is spoken you appreciate the melody, the music of the sound, and it is one of the things that motivate the teacher to work on.

Mrs. Sara H

The Researcher (Question) **Do you assess your students' speaking skills?**

The Interviewee (Answer) I certainly do

The Researcher (Question) **What and how do you assess?**

The Interviewee (Answer) I assess the speaking skill right after the student finishes his or her answer

The Researcher (Question) **How do you report /note down that assessment?**

The Interviewee (Answer) Could you repeat the question..I give my students an opportunity to narrate oral stories or fables so on through which I to evaluate the way they tell the stories, the mastery of the language and the speaking abilities.

The Researcher (Question) **What activities do you use in class for speaking?**

The Interviewee (Answer) The ones I just talked about them

The Researcher (Question) **What do you do with students whose speaking is very difficult to be understood?**

The Interviewee (Answer) I merely ask them to repeat what they have said

The Researcher (Question) **Are there any external considerations when planning the speaking component of the course for your students or can you alone, as a teacher, decide what to teach to your group of students? (What are the external considerations, if any?)**

The Interviewee (Answer) I alone decide what to teach

The Researcher (Question) **Inteaching / learning situations with prescribed curricula that you have some knowledge of, or experience in, to what extent is pronunciation a feature?**

The Interviewee (Answer) Frankly don't give it too much room, but I generally focus on common mispronunciation of words like the word "jan and the word liked because I have noticed that my students not only my students but in general almost all the students they pronounce the word jan like /..janer.../ and the word liked like /....../

The Researcher (Question) **Howmuch time within these lectures is devoted to pronunciation?**

The Interviewee (Answer) No more than five minutes

The Researcher (Question) **Whatemphasis do you feel should be placed on pronunciation? (How important is pronunciation?)**

The Interviewee (Answer) It's so important in fact mainly when it comes to a foreign language like English.

The Researcher (Question) **Do you teach pronunciation?**

The Interviewee (Answer) I don't teach it.

The Researcher (Question) **Why not?**

The Interviewee (Answer) BecauseI see that it's more acquired than learned besides my interest in the world of literature it wins all my focus.

The Researcher (Question) **Somedifferent approaches to pronunciation are listed below, which of these do you use? Describe when, how often and why you choose each/some/a particular method?**

- **Whole class tasks/activity with pronunciation as the main focus**
- **Whole class tasks/activity with pronunciation as component.**
- **One to one situation**
- **Small groups of students**
- **Referral to a specialist teacher where available**
- **Other approaches?**

The Interviewee (Answer) I think that it depends on the situations we are in; in general I focus on one to one situation not all the time but in general.

The Researcher (Question) **Which pronunciation materials are you aware of/familiar with?**

The Interviewee (Answer) Certainly I am but I ignore the types of these devices you are talking about.

The Researcher (Question) **Have you heard about the thePraat software before?**

The Interviewee (Answer) Ah never, I have never heard about this before.

The Researcher (Question) **Have you an idea about what a spectrogram is?**

The Interviewee (Answer) No, never. I have never heard about such devices.

The Researcher (Question) **According to you, why are many teachers afraid of phonetics?**

The Interviewee (Answer) Frankly, don'tlike it. I prefer literature

The Researcher (Question) **What are your suggestions for teachers who teach oral expression and phonetics?**

The Interviewee (Answer) Well, to try to work together.Provide the appropriate environment for teachers and for students.

Ms.Houda. B

The Researcher (Question) **Do you assess your students' speaking skills?**

The Interviewee (Answer) Yeah of course, sometimes it's important to assess their speaking skill especially when it is a course that has to do basically with the speaking skill itself in the first place like the oral expression of course where students are supposed to say something in English, so then you give more attention to their speaking capacities or the speaking skill.

The Researcher (Question) **What and how do you assess?**

The Interviewee (Answer) Well, for me what it is important in the speaking capacities or the speaking skill is their fluency, this is the first thing that, I mean if they are fluent in English this is a positive sign but they can improve their speaking skill with time. How to assess it's only just by listening to what they say how they say it, when it's really a big mistake whether in terms of pronunciation especially pronunciation or they are not using the appropriate word for the appropriate message, I feel obliged to correct them.

The Researcher (Question) **Whatactivities do you use in class for speaking?**

The Interviewee (Answer) Basicallyit's giving them opening the large space for discussion, I mean encourage them to speak in English by enabling to talk many things in English. As far as the material we do have

The Researcher (Question) **What do you do with students whose speaking is very difficult to be understood?**

The Interviewee (Answer) Whatdo you mean by difficulty first? It looks like a native speaker? It's the opposite? I think that it's my job I do my best to ask them again to say it, in English if they fail why not in the native language to say what they want to say.

The Researcher (Question) **Do you make a reference to the native language?**

The Interviewee (Answer) Whynot; in some situation you feel really that it's the last solution that you have to deal with understanding difficulties.

The Researcher (Question) **Are there any external considerations when planning the speaking component of the course for your students or can you alone, as a teacher, decide what to teach to your group of students? (What are the external considerations, if any?)**

The Interviewee (Answer) Well, concerning this point I mean deciding about the content to teach in terms of the speaking skill, as a teacher I wish we would have or we could have been provided with what to teach as far as our experience in teaching the speaking skill is concerned we have more experienced people who may be of a great help in this domain by providing or by suggesting a specific content to teach to our students, so in this , I mean concerning the speaking skill that it was my wish but in a situation where I find myself short of a such a background or such a content I feel obliged to make a decision of course if I take the example of this year it was in collaboration with my students that we have decided on how to have this opportunity of speaking English in classroom.

The Researcher (Question) **Whatemphasis do you feel should be placed on pronunciation? (How important is pronunciation?)**

The Interviewee (Answer) It is. Pronunciation is very important, I mean if we just look at the demands, the needs of our students, their wishes, they want to be categorised as students speaking the British accent or the American accent they have this desire to belong to one of these two groups, unfortunately in our case we cannot we, as teachers... I am talking about myself of course I cannot identify myself as speaking the British accent or the American accent as far as it is an Algerian accent a combination of all accents.

The Researcher (Question) **What do you mean by an Algerian accent?**

The Interviewee (Answer) Algerian it is a collection of a lot of accents, sometimes you pronounce a word you hear it from TV it has been pronounced in a British accent you learn it, and you like it, otherwise you prefer the American accent, the remaining English it's the way you pronounce it, you yourself it's not something you heard before but it's the way you develop its pronunciation as far as it has to do with the IPA I mean it's an understood language, an understood English everybody understands is capable of getting the message , it has no wrong pronunciation it's Algerian.

The Researcher (Question) **It'swhat we refer to as intelligibility, how much can we be understood by the others.**

The Interviewee (Answer) Yes it is

The Researcher (Question) **How much time within these lectures is devoted to pronunciation? If any of course**

The Interviewee (Answer) Not if any, it's not something really that we calculate in terms of time, I mean when the students are speaking, pronunciation is that sometimes you feel obliged not to stop the student even if he is not pronouncing correctly because of the fear of blocking the students once and forever so you let him express himself if you have the chance to go back to what was said and how it was said and try to correct him because as far as your knowledge is limited it's ok sometimes the mistake is so terrible that you cannot ignore it, so pronunciation is like it should be present through the whole course not at specific moments because as I have said the course that gives importance to speaking as it is taught today we cannot have this perspective of giving specific time devoted to pronunciation for example if we have more sophisticated material we could have given students a text ask them to read to hear their readings and to be corrected to hear , this is good but since this is not the case ...

The Researcher (Question) **Whatemphasis do you feel should be placed on pronunciation? (How important is pronunciation?)**

Well, I said it before. I mean pronunciation is very important if you want to improve our students' speaking capacities, I mean it shouldn't be limited to a specific course that gives a certain instruction in how to develop your speaking skill, pronunciation can be present elsewhere in other courses for instance a course where the students could be asked to read loudly a poem or a text. Correct pronunciation is necessary for a good and correct understanding of a specific text.

The Researcher (Question) **Do you teach pronunciation?**

The Interviewee (Answer) Asa separate element? No

The Researcher (Question) **Why not? not to teach pronunciation**

The Interviewee (Answer) I don't teach pronunciation .It is not a separate thing that I do teach but as I said before pronunciation is present throughout the whole course itself, Ok so I can tell them we should pronounce this word this way and not that way..But to devote a specific aspect or specific part of the course entitled pronunciation and students are aware that this is pronunciation and not something else not intonation, or not something else, No!

The Researcher (Question) **Intonation is a part of the pronunciation!**

The Interviewee (Answer) Yeah pronunciation say not phonetic transcription of words for example, they know it's pronunciation, they have to practice their sounds, their pronunciation of words, their intonation..

The Researcher (Question) Yes, the suprasegmental features.

The Researcher (Question) When/if you teach pronunciation, which aspects do you focus on?

The Interviewee (Answer) Well, correct pronunciation as simple as that, as far as the words are pronounced in correct English whether American or British, as far as it is a correct pronunciation it is good.

The Researcher (Question) Somedifferent approaches to pronunciation are listed below, which of these do you use? Describe when, how often and why you choose each/some/a particular method?

- **Whole class tasks/activity with pronunciation as the main focus**
- **Whole class tasks/activity with pronunciation as component.**
- **One to one situation**
- **Small groups of students**
- **Referral to a specialist teacher where available**
- **Other approaches?**

The Interviewee (Answer) Concerning the teaching of pronunciation, I believe whole class tasks and group tasks could be very efficient, whole class tasks may enable students to discover each others' capacities in terms of pronouncing English they can contribute in correcting each other, and in correcting each other and showing their different capacities.

Group work also is very effective as far as we can have different capacities put together, we may have the weak working with the good student, this can be like a motivating factor to the weak students to improve the speaking or it may just bring the opposite result, so we have to be very careful when deciding about the composition of those groups, I mean in the decision that we make, sometimes students may prefer to work together, just not to have this barrier of communicating with each other and doing activities concerning pronunciations they won't feel ashamed or shy of saying loudly the words.

The Researcher (Question) Whichpronunciation materials are you aware of/familiar with?

The Interviewee (Answer) No, I am not familiar with any.

The Researcher (Question) So, what do you use to teach pronunciation?

The Interviewee (Answer) According to the teacher and the students

The Researcher (Question) Have you heard about the thePraat software?

The Interviewee (Answer) No!Never.

The Researcher (Question) Have you any idea about what a spectrogram is?

The Interviewee (Answer) I think it's not a new term, but knowing exactly what is it No!

The Researcher (Question) According to you, why are many teachers afraid of phonetics?

The Interviewee (Answer) Because we don't have enough knowledge in teaching it.

The Researcher (Question) **What are your suggestions for teachers who teach oral expression and phonetics?**

The Interviewee (Answer) To coordinate both modules. Give both teachers the important devices and materials to teach these two modules adequately.

V. Houda A.

The Researcher (Question) **Good morning Houda. I want first to know what is your experience in teaching? I mean how many years have you been teaching?**

The Interviewee (Answer) I have been teaching for almost 15 years as a part time teacher and for 3 years as a full time teacher.

The Researcher (Question) **Do you assess your student's speaking skill?**

The Interviewee (Answer) Yes, I do

The Researcher (Question) **Whatand how do you assess?**

The Interviewee (Answer) Wellusually I just use my mental capacities to check how they answer and how they talk in class, I don't use any mean of assessment any other one.

The Researcher (Question) **Whatactivities do you use in classroom for speaking?**

The Interviewee (Answer) Sometimes I give free talks, sometimes I give them a text and ask them to read it, and sometimes I ask them to watch a movie and ask them to come back and give an idea and their opinion about that movie

The Researcher (Question) **Do you think all the students' can watch that movie, can afford it?**

The Interviewee (Answer) No, most of my students don't some of them did.

The Researcher (Question) **What do you do with students whose speaking is very difficult to be understood?**

The Interviewee (Answer) Usually I correct them in class I mean directly; after they make a mistake I drive their attention to the mistake when it is a flagrant mistake and I correct it sometimes I use the board to write the word and ask them to repeat that word many times to memorize it.

The Researcher (Question) **Howmuch time within these lectures is devoted to pronunciation?**

The Interviewee (Answer) Usuallythe quarter of my class of the time of my class not very much time.

The Researcher (Question) **What Importance do you give to pronunciation?**

The Interviewee (Answer) I give medium importance to the pronunciation sometimes.

Why?

The Interviewee (Answer) In the past, I used to be very meticulous when it comes to pronunciation but lately I have noticed that I don't have enough time to correct everybody and to teach everybody to pronounce correctly, this is why I ask them to use their personal means to improve their pronunciation by watching documentaries, and TV and using songs and so on.

The Researcher (Question) **How important is pronunciation to you?**

The Interviewee (Answer) Itis very very important because most of the time if you don't pronounce correctly the message might not get transmitted because most of my students use wrong words or wrong pronunciation and I don't get them.

The Researcher (Question) **Whatemphasis do you feel should be placed on pronunciation?**

The Interviewee (Answer) Weshould emphasize on pronunciation more than any other skills in oral expression class just oral expression.

The Researcher (Question) **Doesit happen to you to teach pronunciation?**

The Interviewee (Answer) As a part time teacher when I used to teach intermediate and advanced levels, I used to give it a very big importance and I give it like at the very beginning of the class I used to ask the students to put the new words that they don't know and I pronounce them for them and I ask the whole class to repeat those words for few times.

The Researcher (Question) **Do you feel confident when you teach pronunciation?**

The Interviewee (Answer) Of course I am very confident.

The Researcher (Question) **When you teach pronunciation, what do you do? Which aspects do you focus on?**

The Interviewee (Answer) I mostly focus on the tone how the students use the tone and the correct pronunciation of every letter.

The Researcher (Question) **Which pronunciation do you use? The American or the British**

I use American English at a 100%. I never swapper to the other one never ever, because I had only an American experience and I am very confident that my American English is correct in order to make sure that I am giving the correct pronunciation. I have no experience with the British English I don't want to give wrong message to students.

The Researcher (Question) **Whichpronunciation materials are you aware of/familiar with?**

The Interviewee (Answer) No one

The Researcher (Question) **Whichdo you use?**

The Interviewee (Answer) I use my personal use of English that's all

The Researcher (Question) **Have you heard about the thePraat software?**

The Interviewee (Answer) No, no never

The Researcher (Question) **Have you an idea what a spectrogram is?**

The Interviewee (Answer) Aspectrogram is a device that measures the tone I think of speaking.

The Researcher (Question) **Doyou have a prior knowledge in acoustic phonetics?**

The Interviewee (Answer) No, never

The Researcher (Question) **According to you why phonetics as the module represents a ghost for teachers in our department?**

The Interviewee (Answer) First of all because we were not taught phonetics as it should be. Second we don't have specialists in phonetics, third we don't have the adequate and the important materials for phonetics because phonetics is not a simple subject to be taught it needs physical materials to be taught we cannot teach it with a chalk and a board No!

The Researcher (Question) **What are your suggestions for teachers who teach oral expression and phonetics?**

The Interviewee (Answer) Well, it is not a suggestion for teachers more, it is a suggestion for the administration to give both teachers the important devices and materials to teach these two modules adequately, some teachers are helpless.

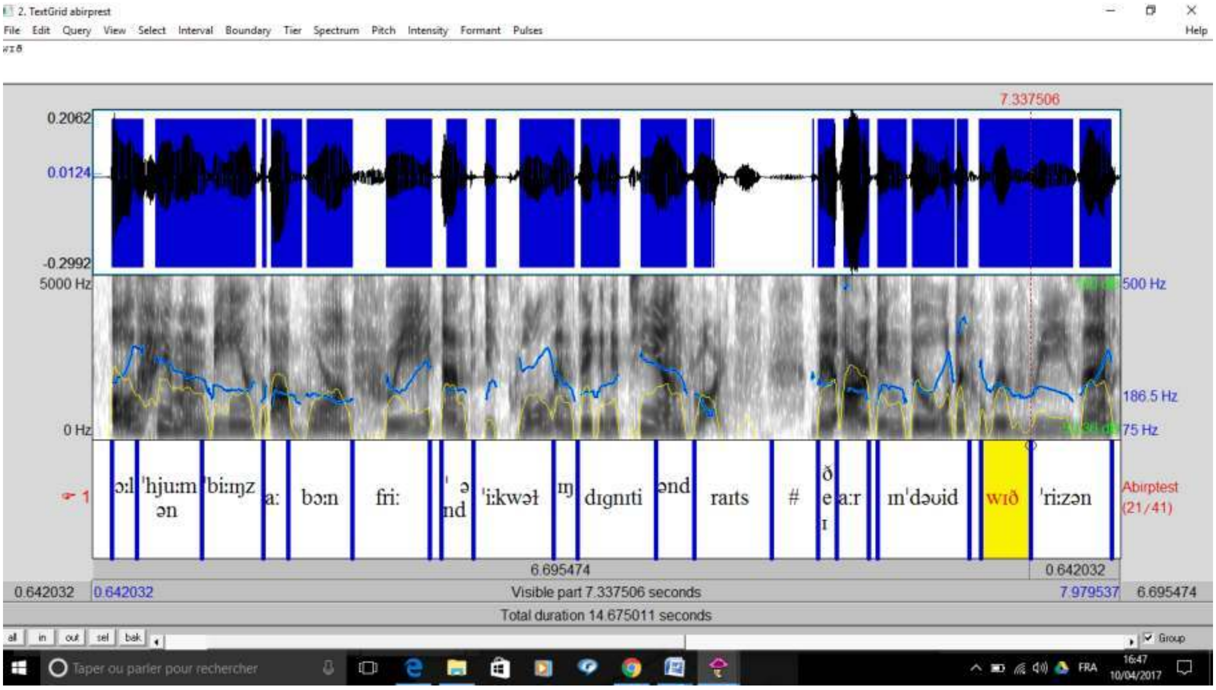
3. Students' Responses

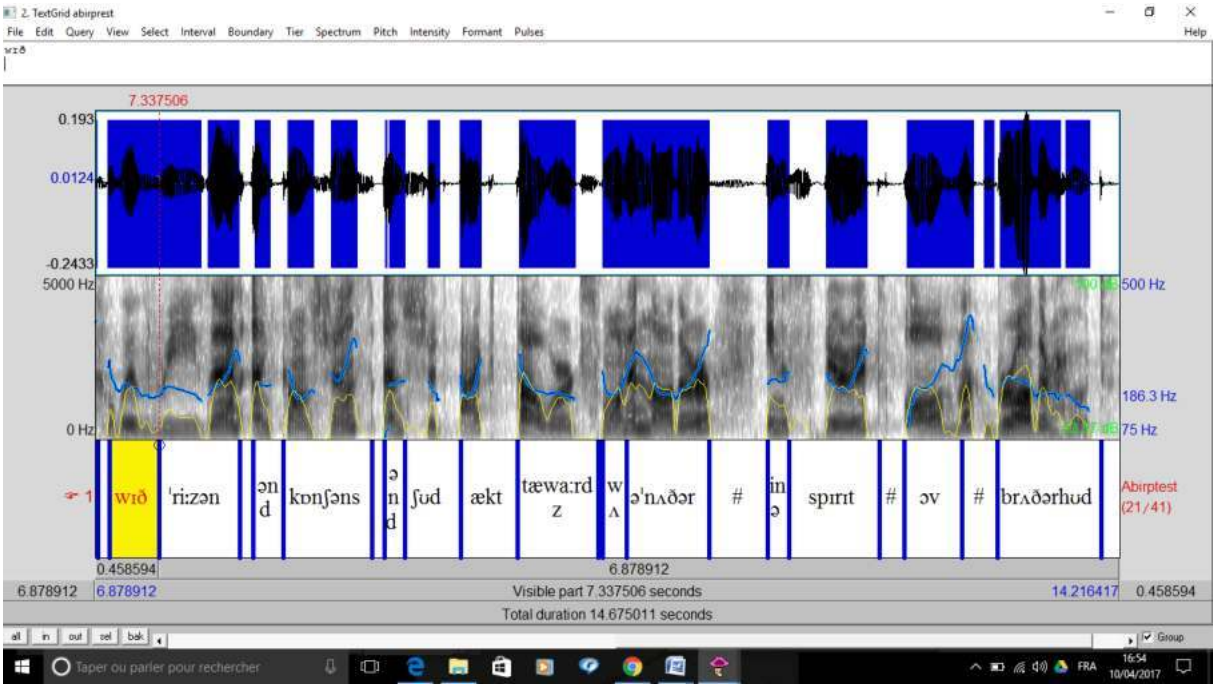
4. Students' Pronunciation

5. Students' Segmentations

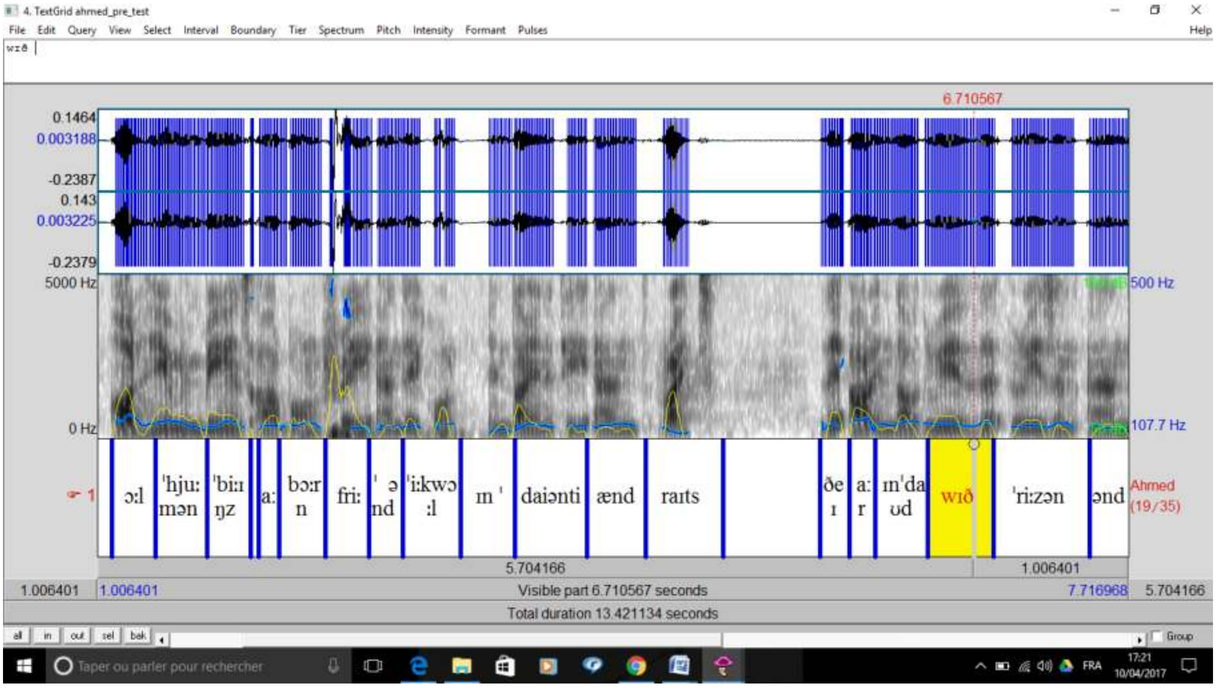
A. Pretest

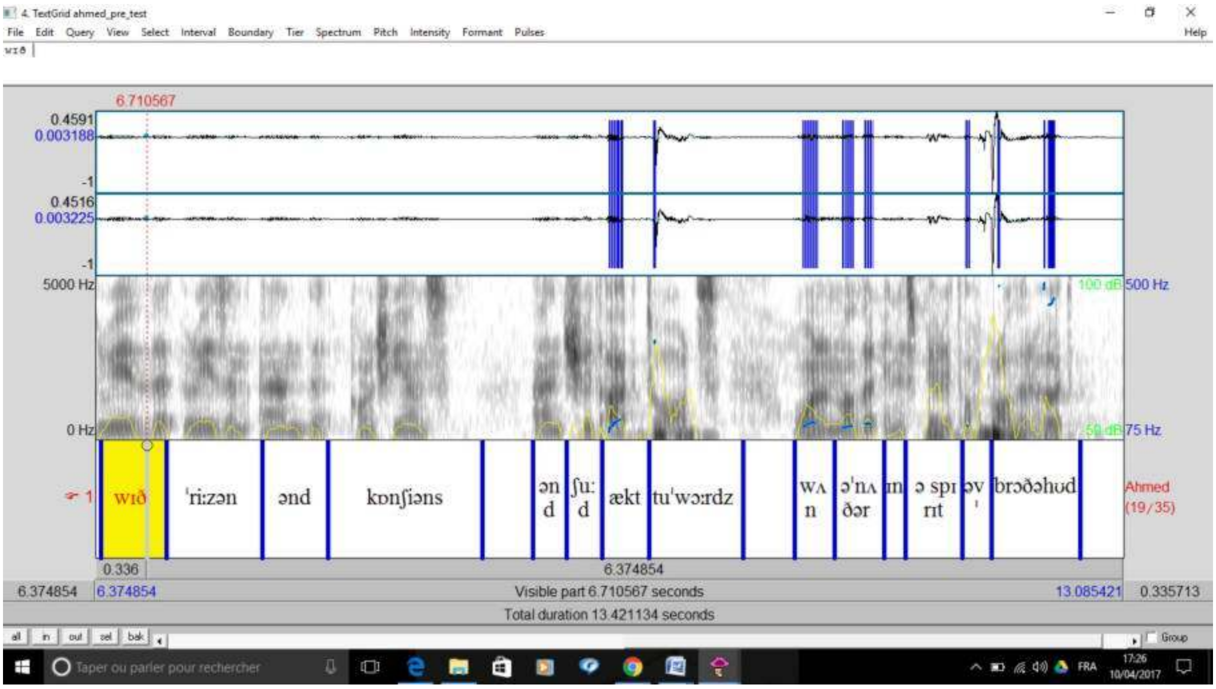
1. Abir. G (Part I & II)



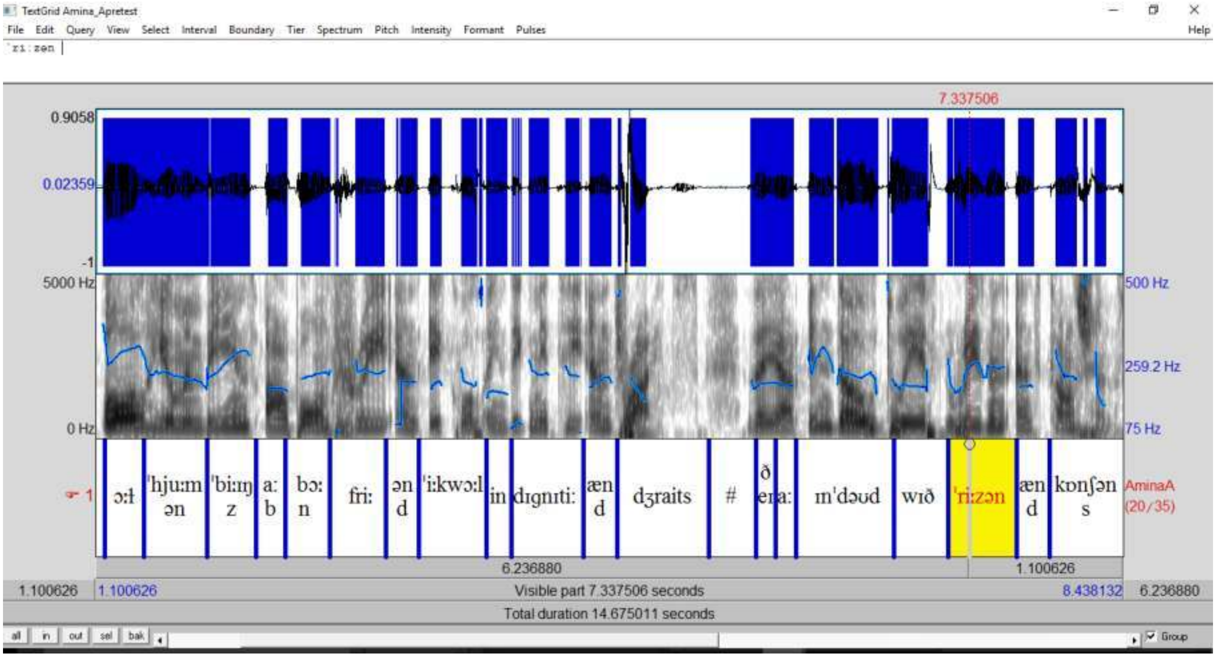


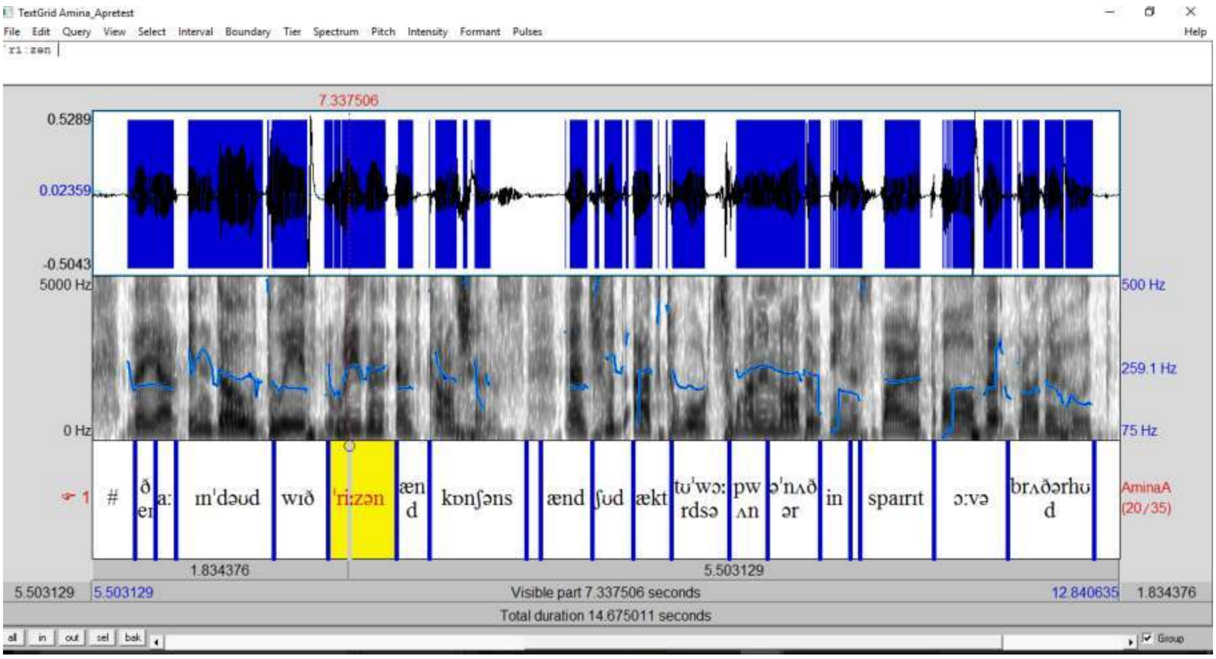
2. Ahmed (Part I &II)



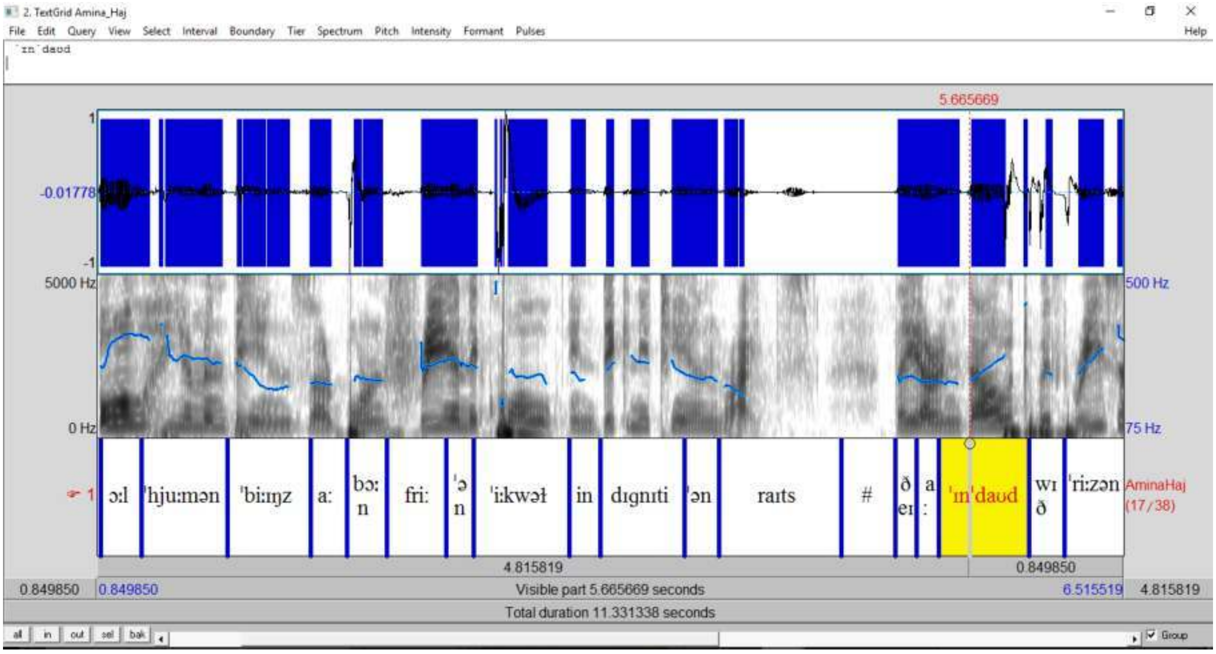


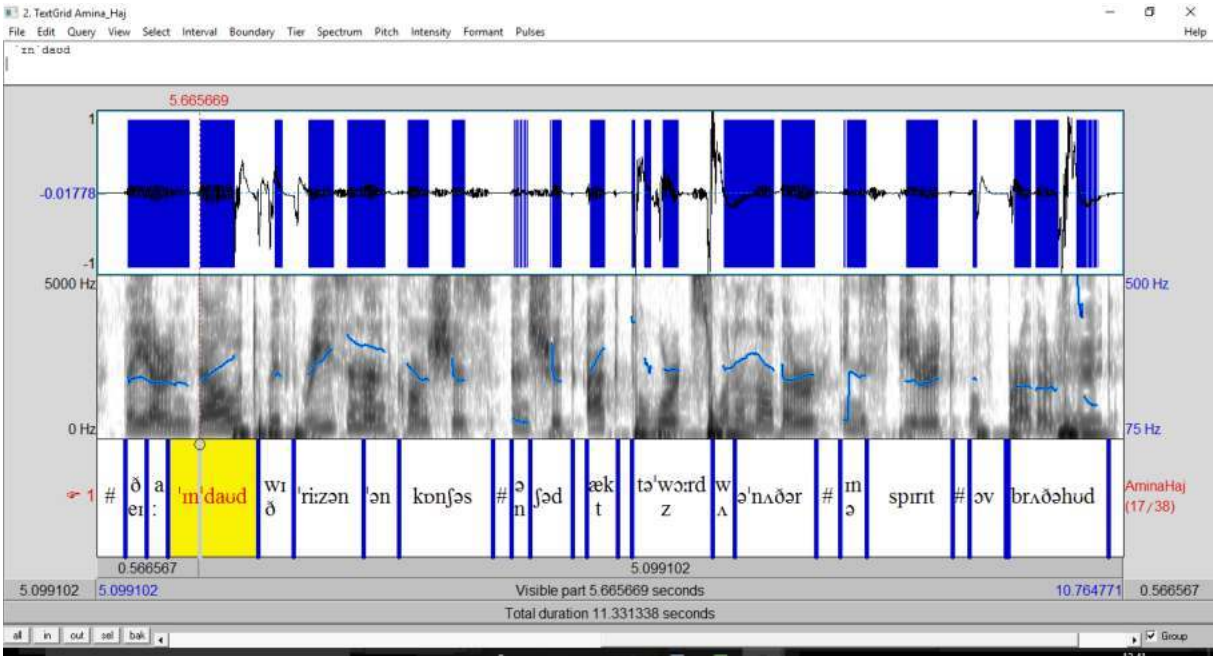
3. Amina A (Part I&II)



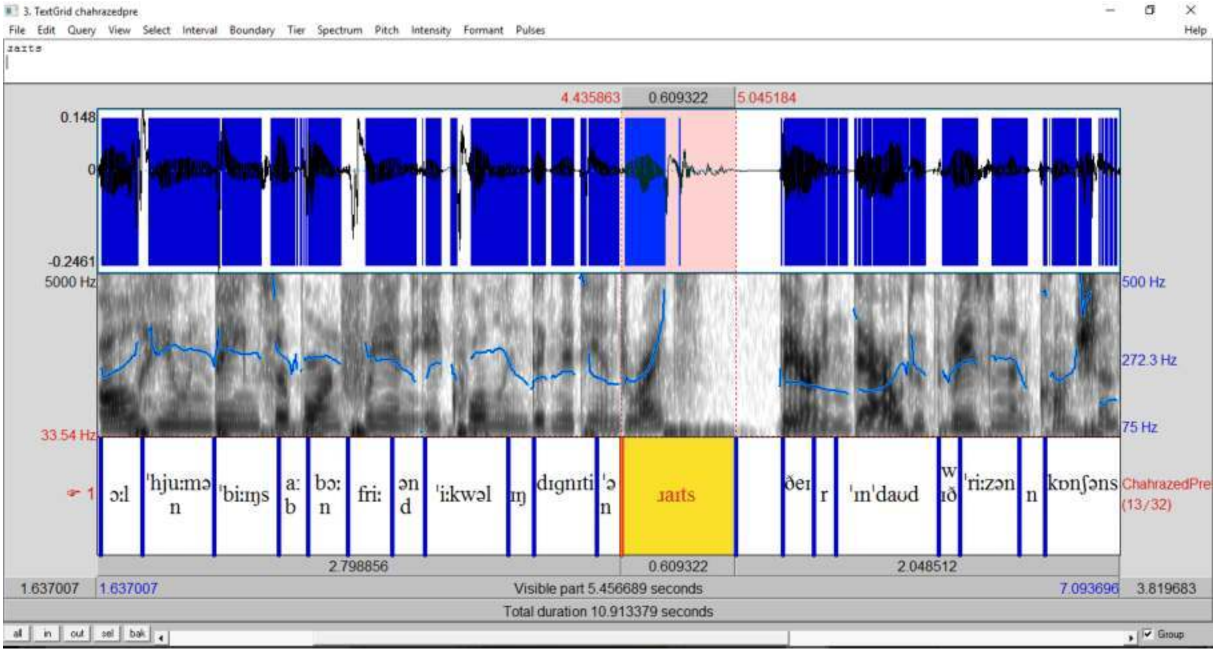


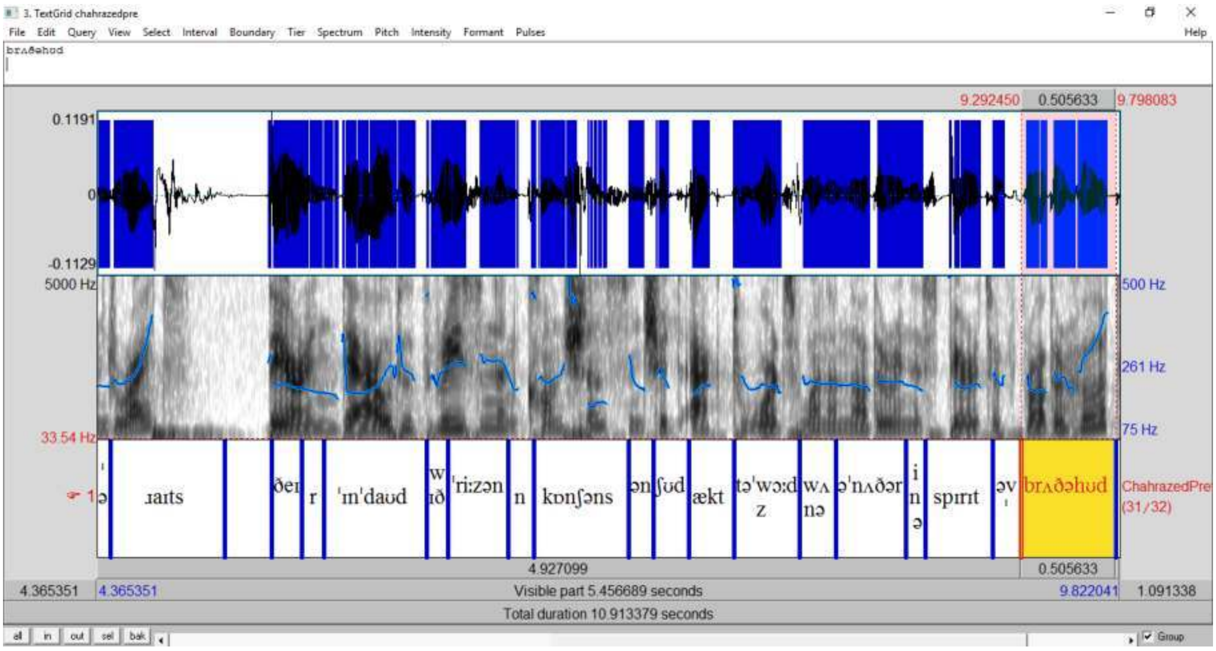
4. Amina H. (Part 1&2)



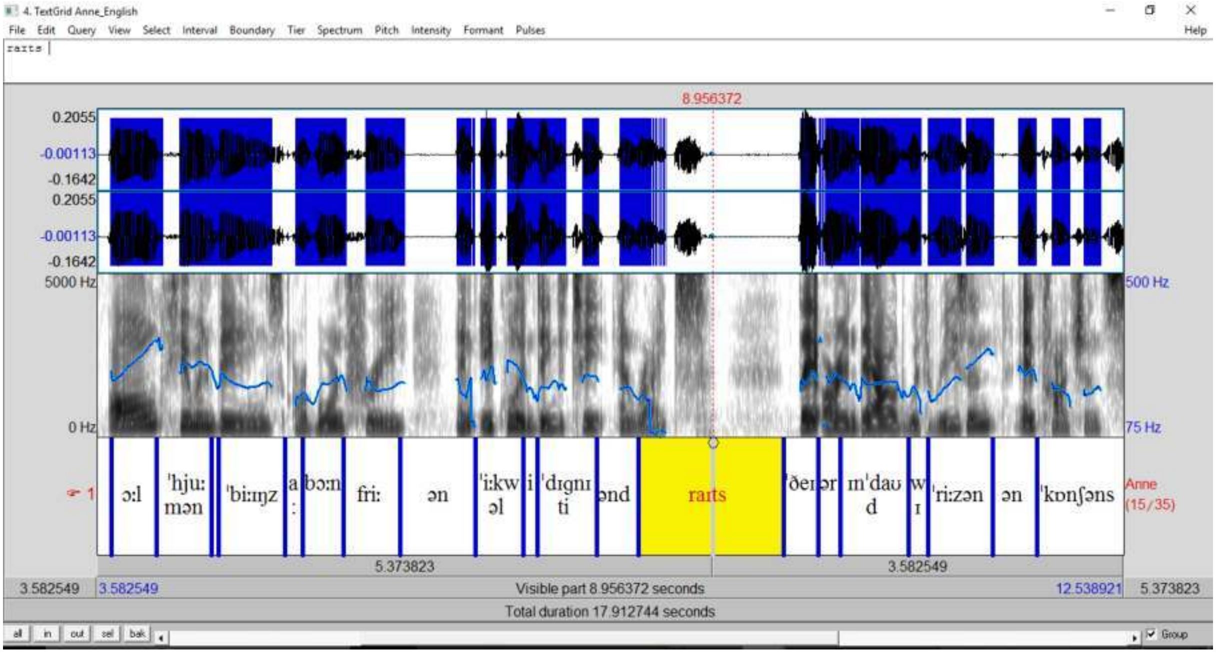


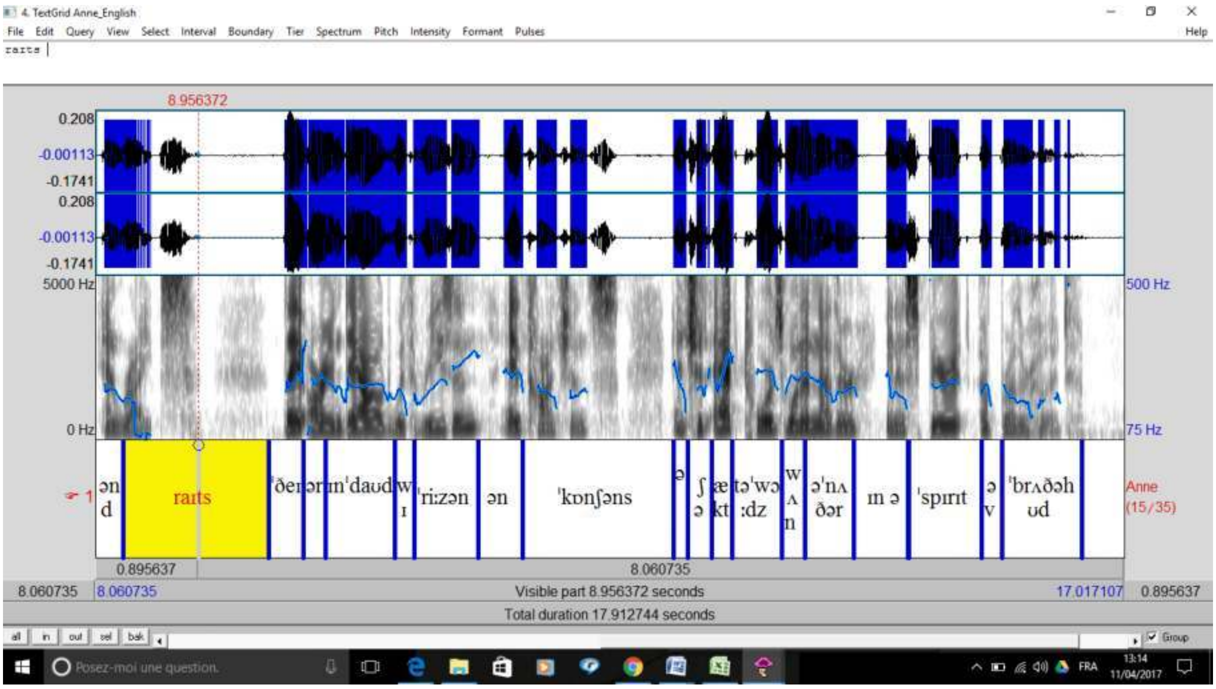
5. Chahrazed (Part 1&2)



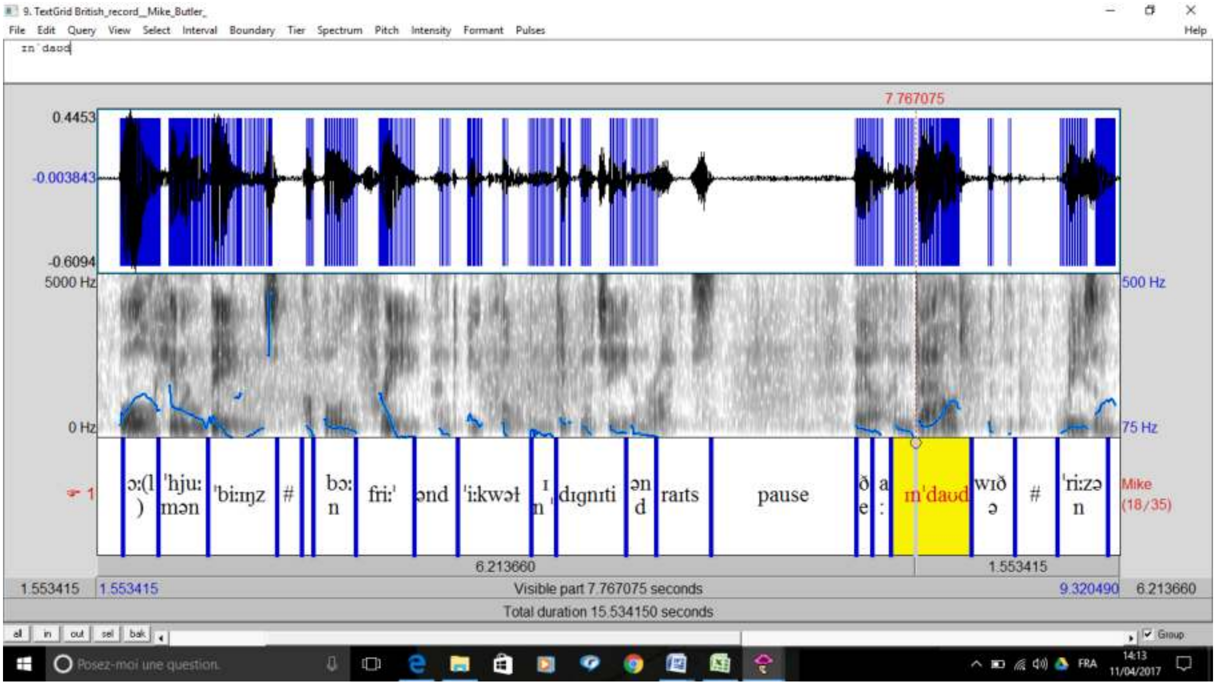


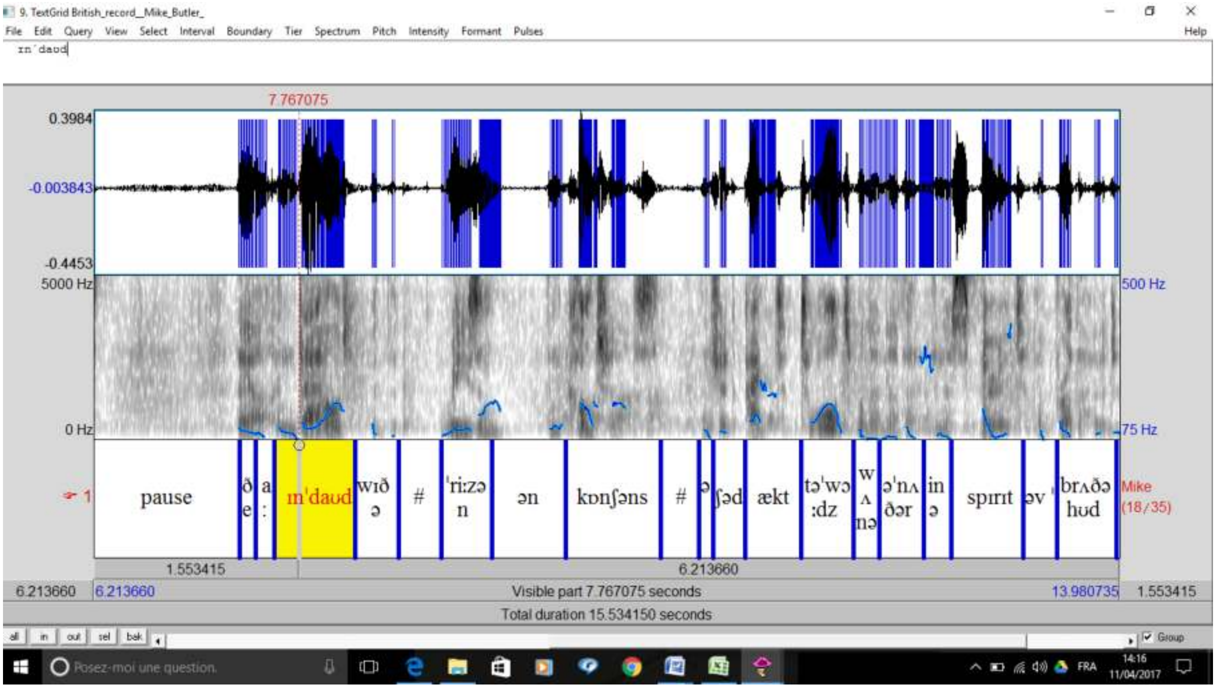
6. Anne Marie (Part 1&2)



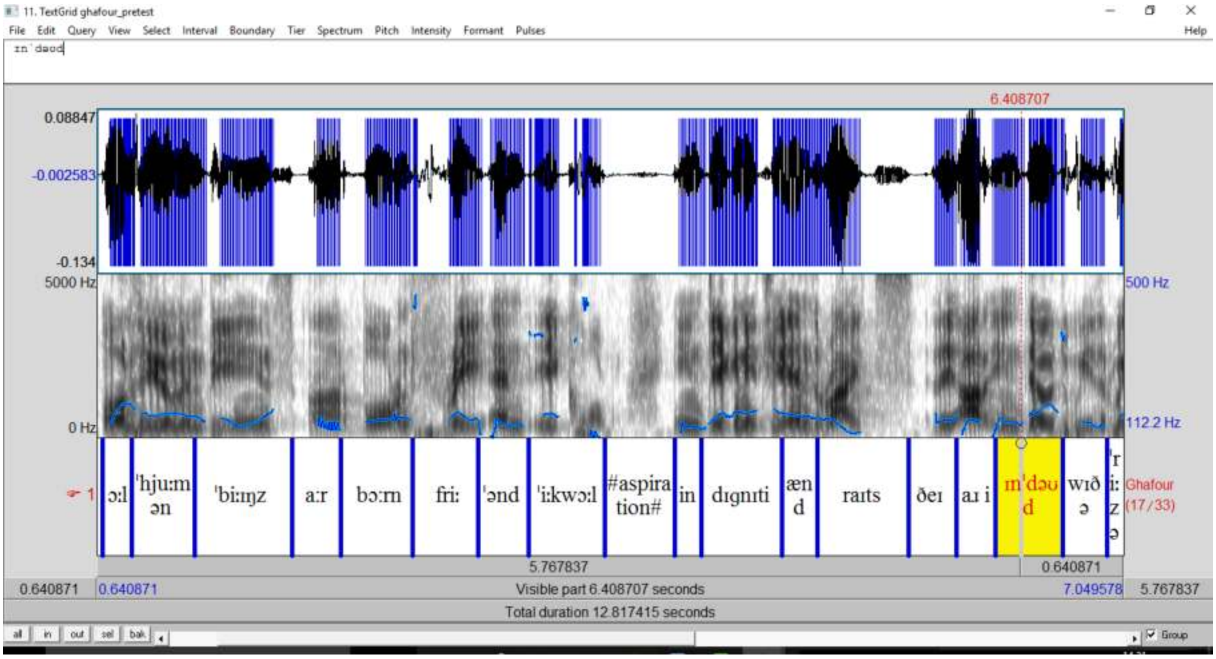


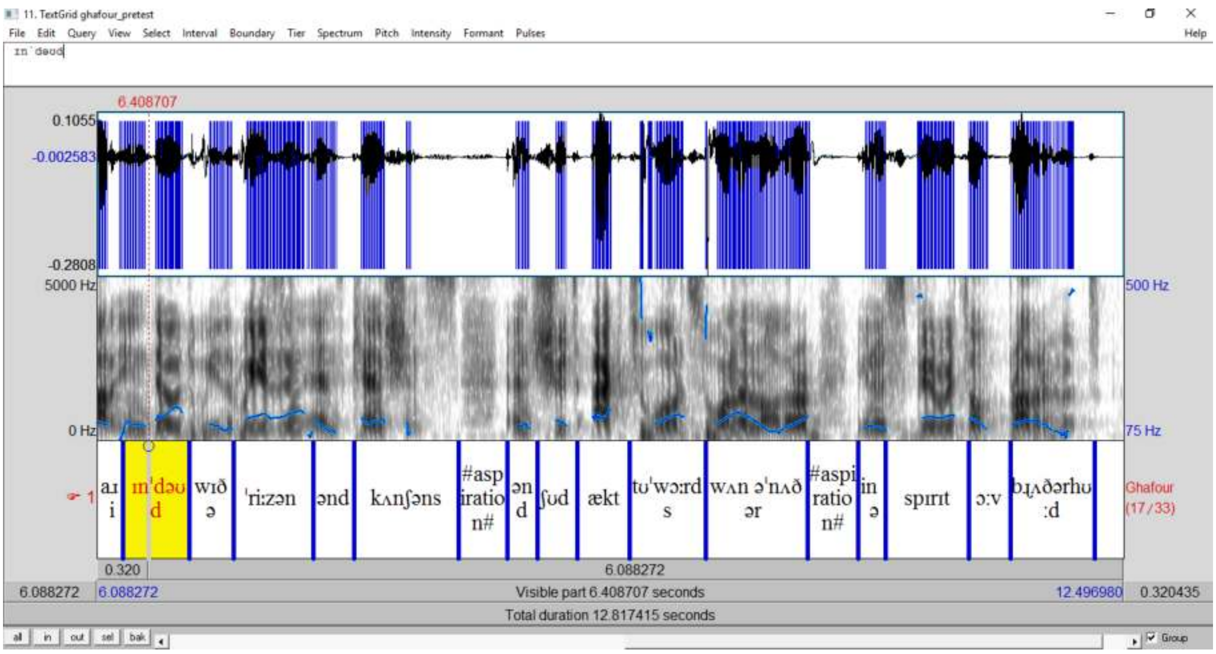
7. Mike Bulter (Part 1&2)



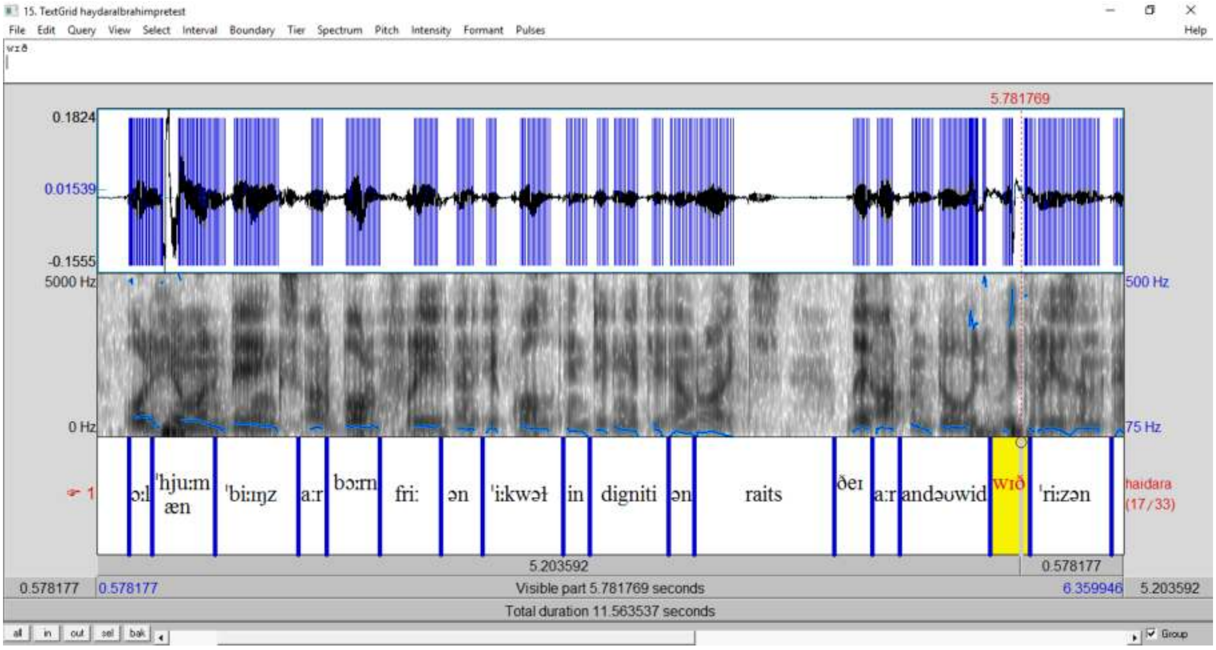


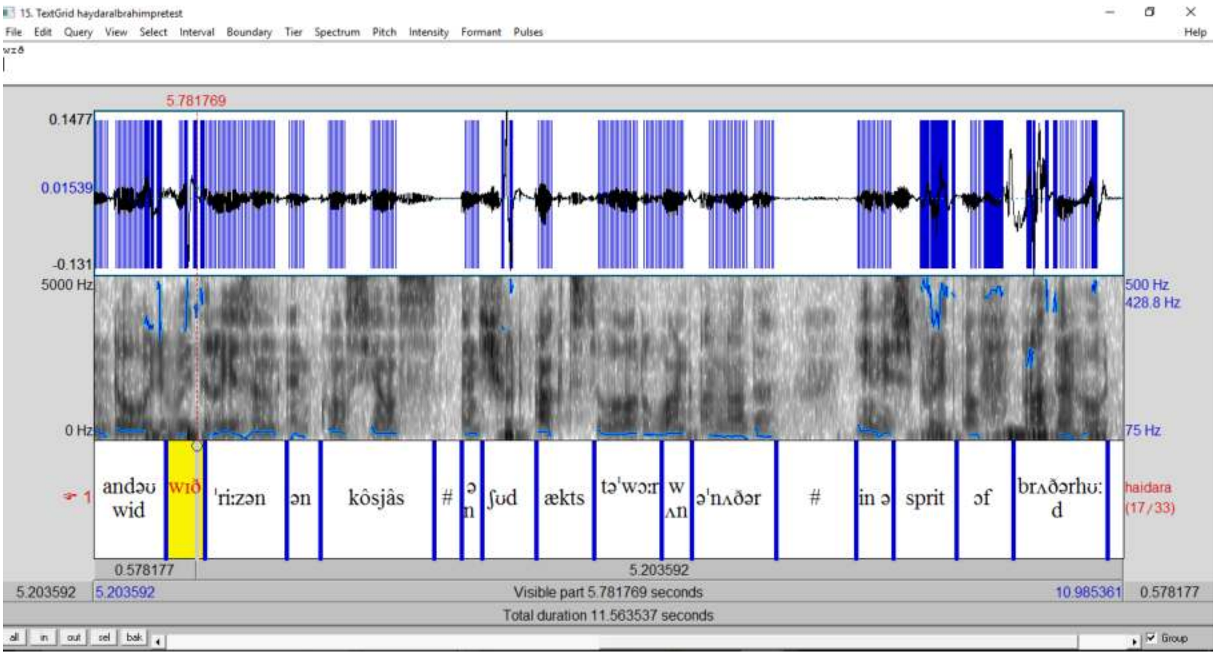
8. Ghafour (Part 1&2)



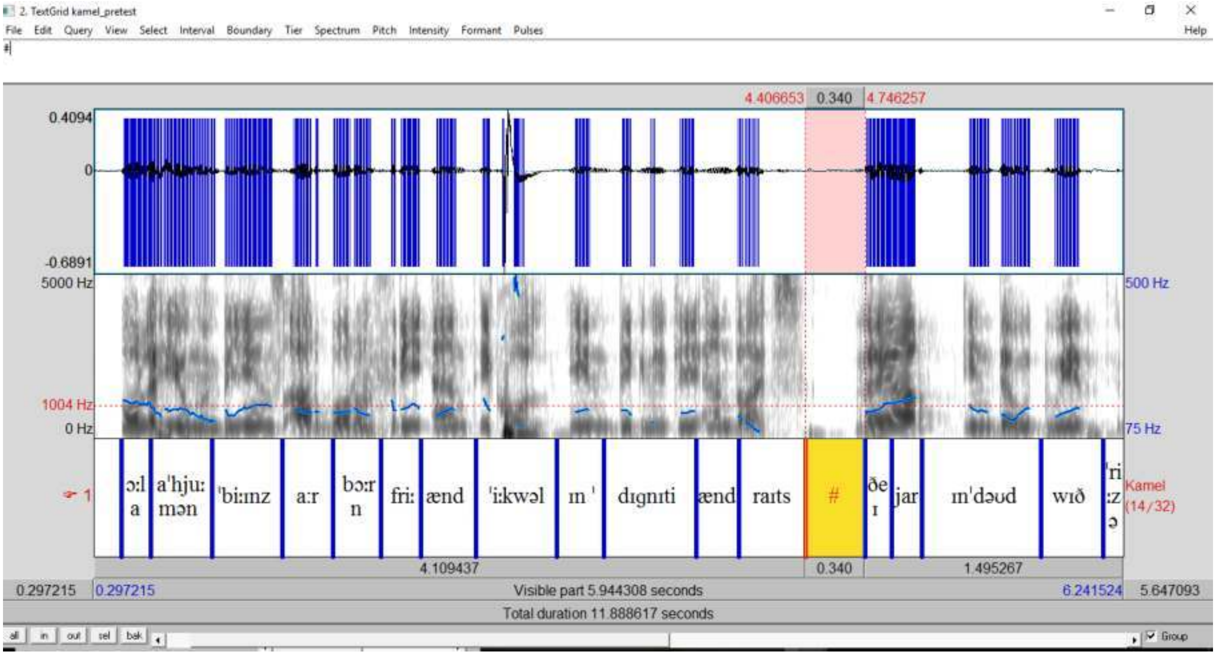


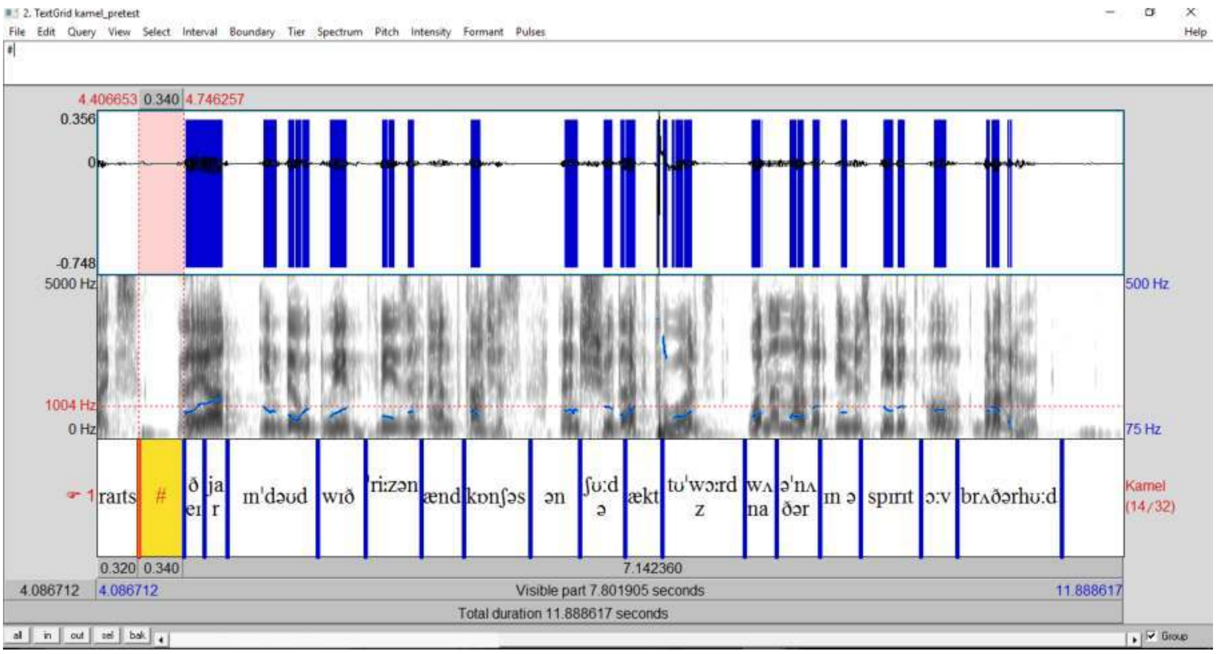
9. Haidara (Part 1&2)



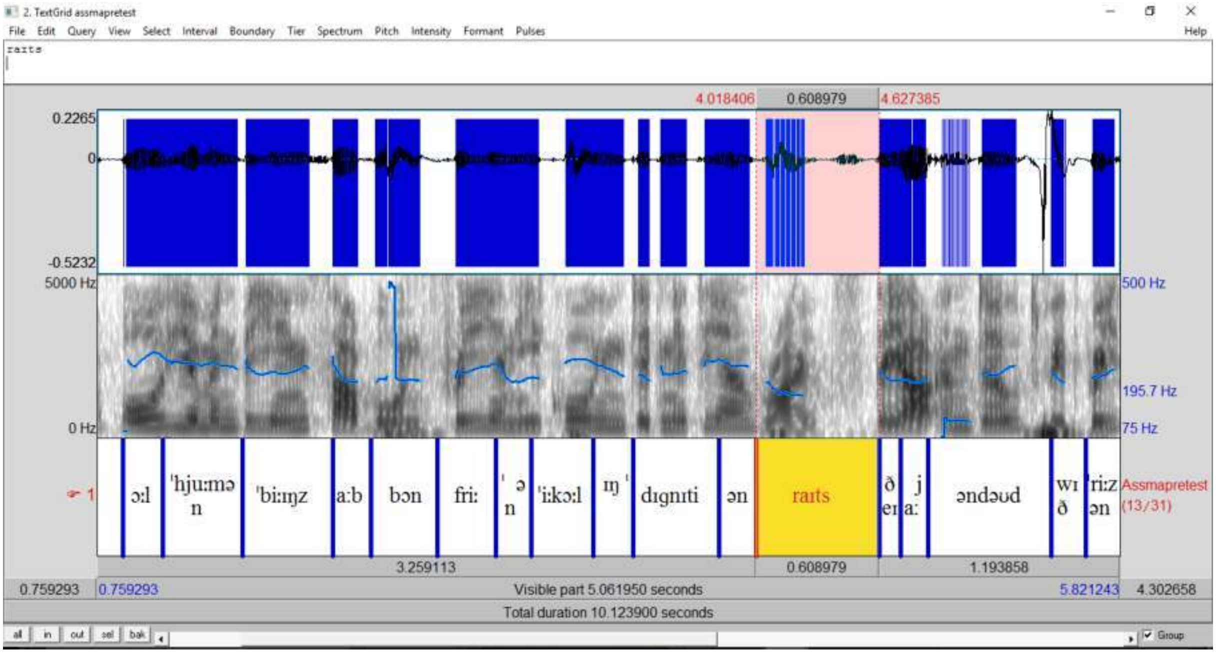


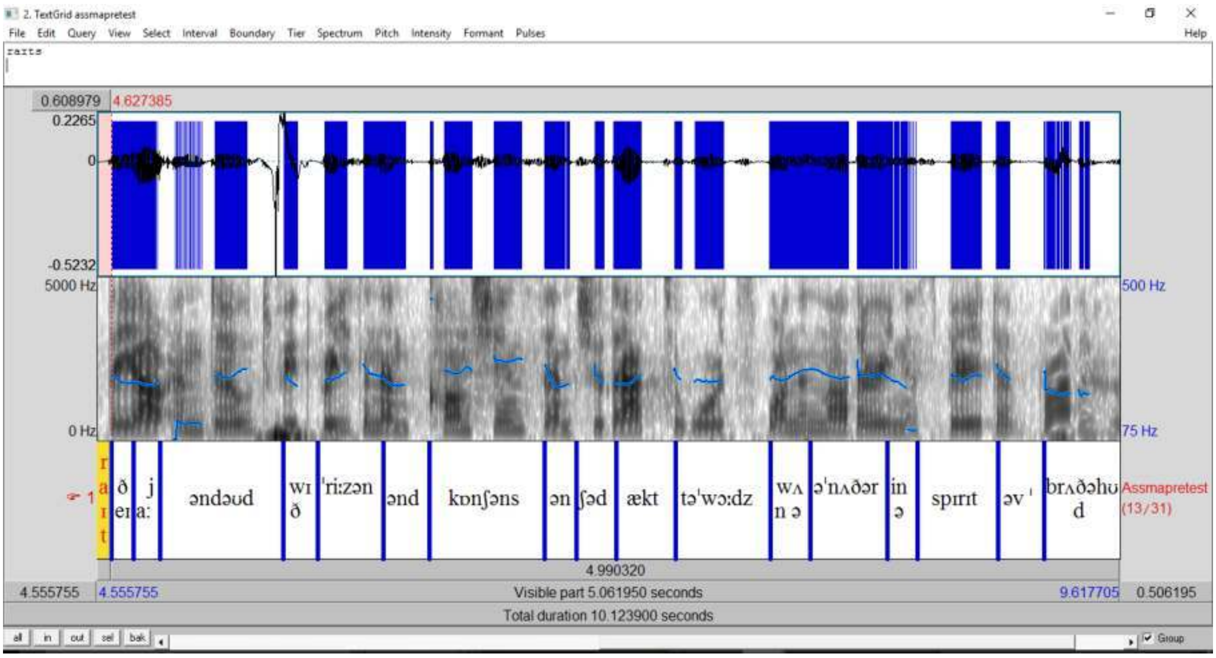
10. Kamel (Part 1&2)



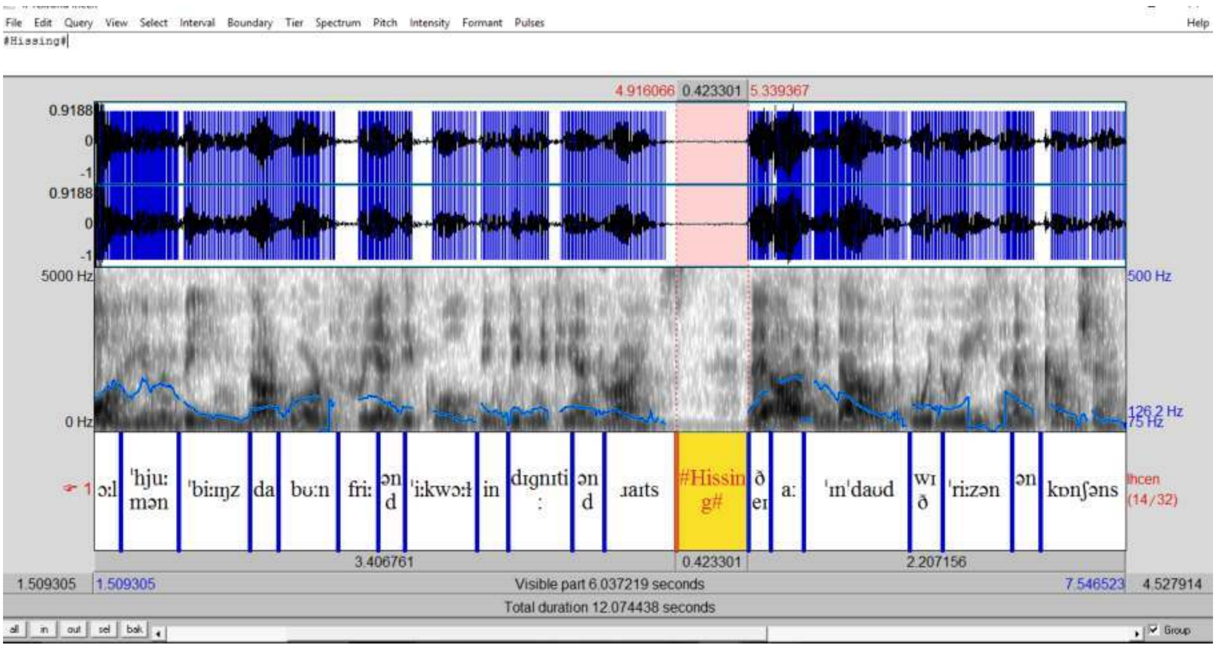


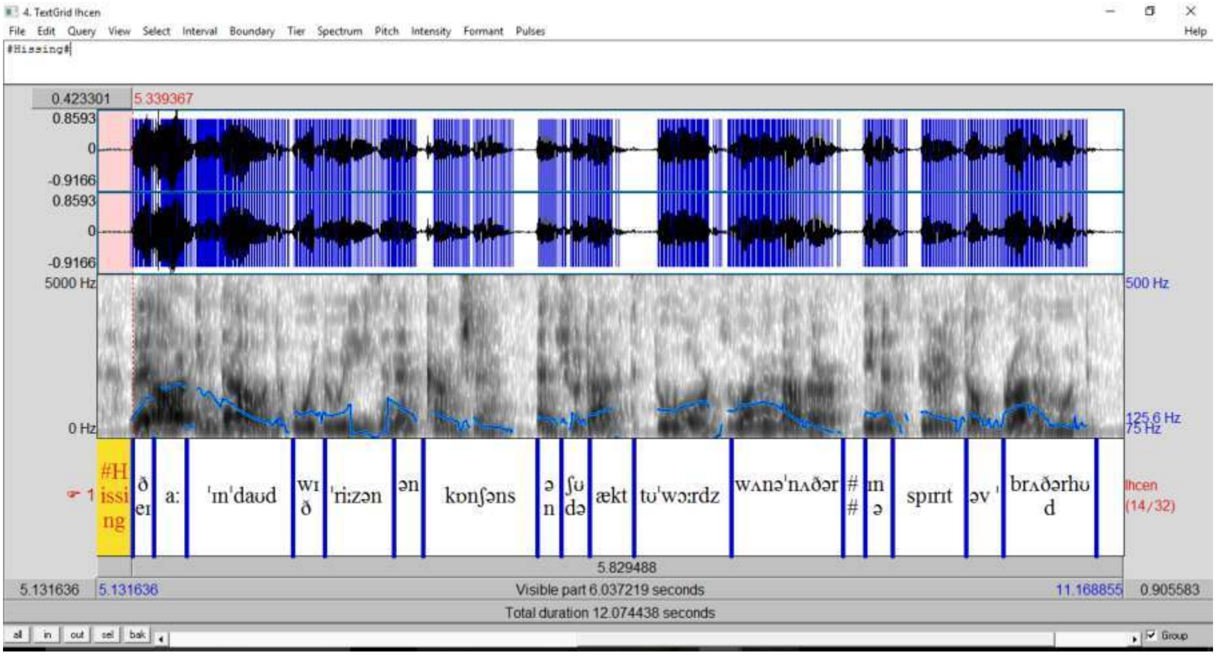
11. KhadidjaAssma(Part 1&2)



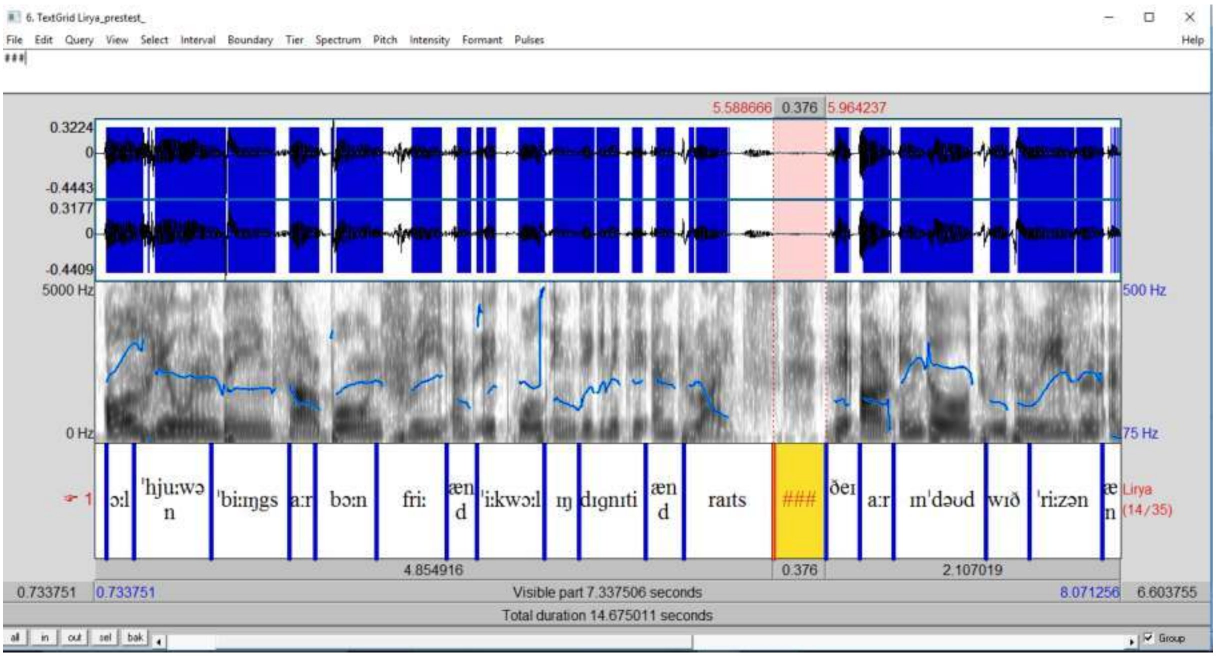


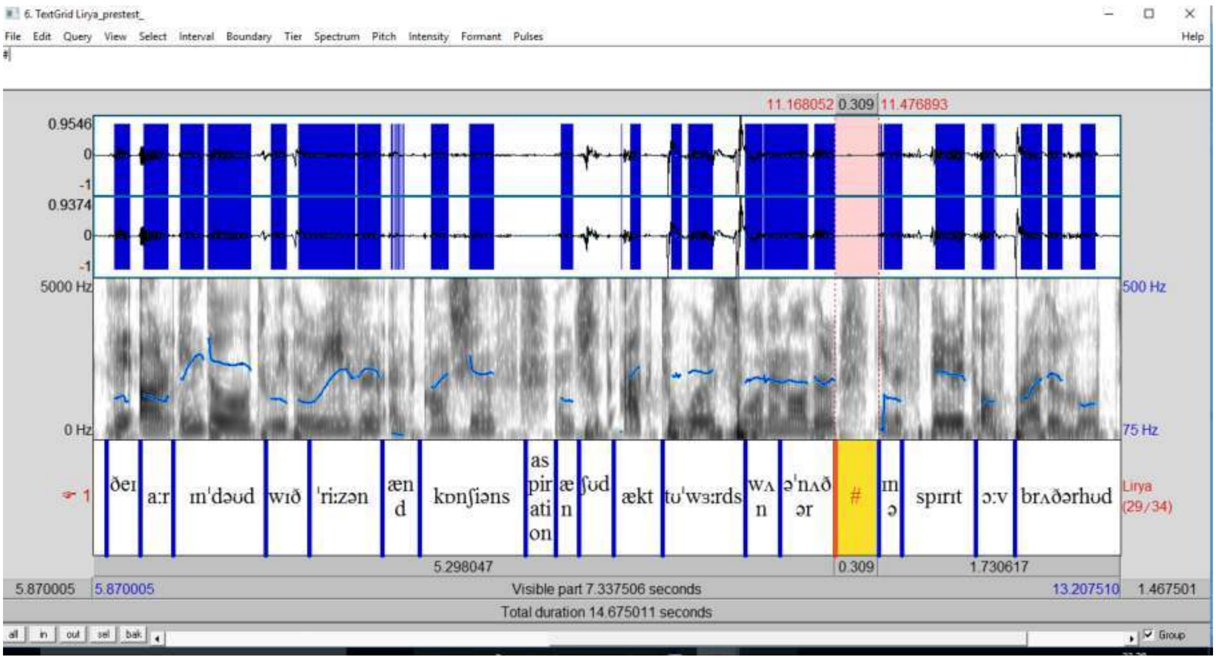
12. Ihcen(Part I &II)



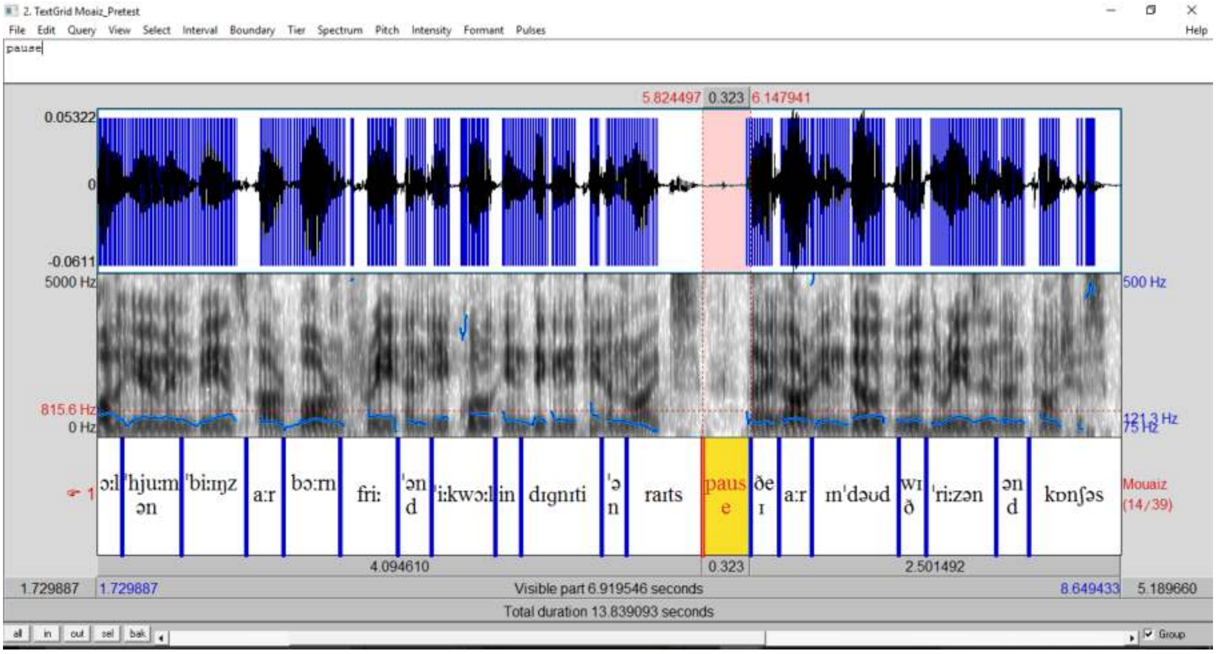


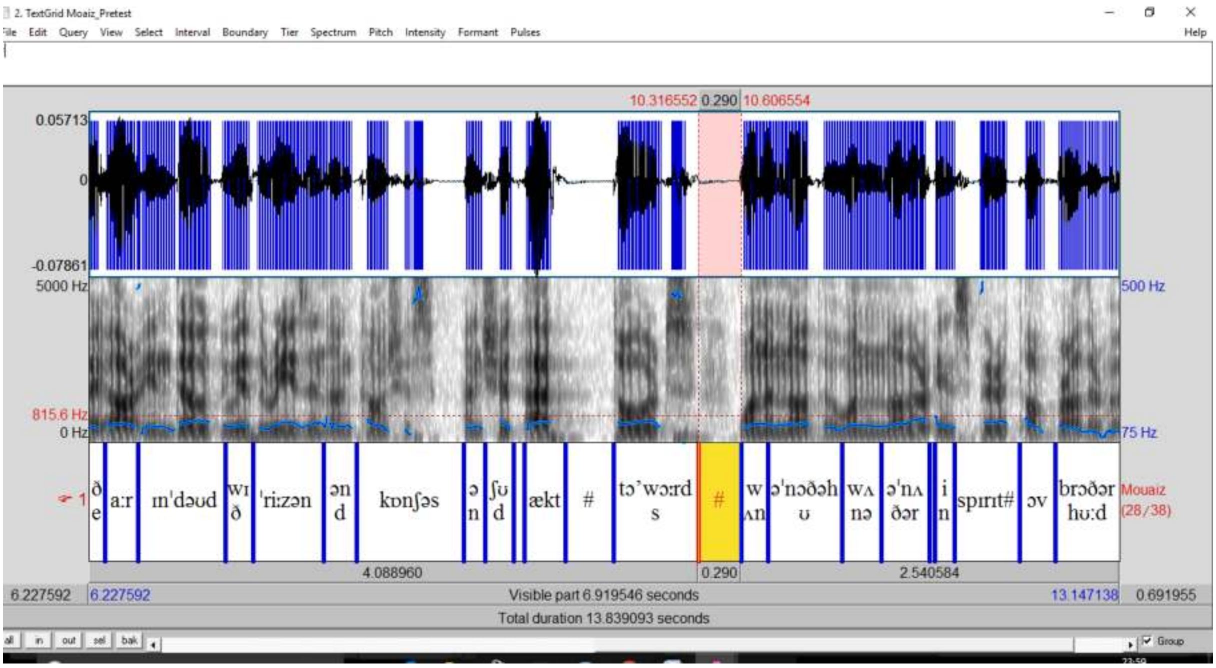
13. Lirya (Part I&II)



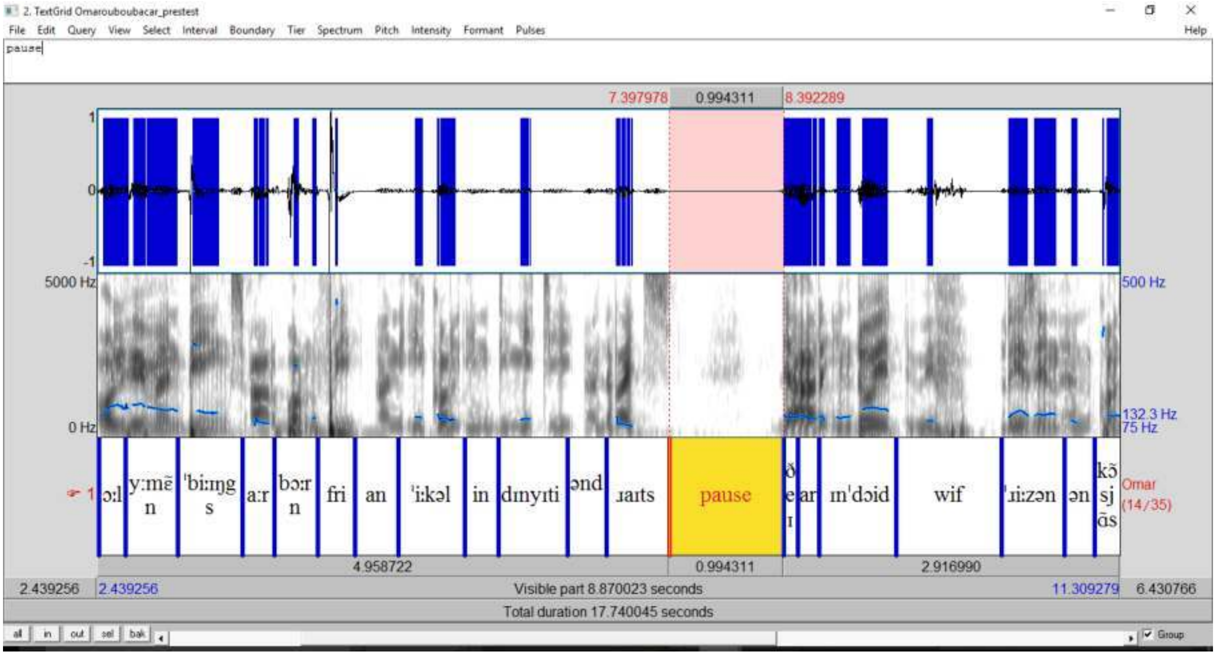


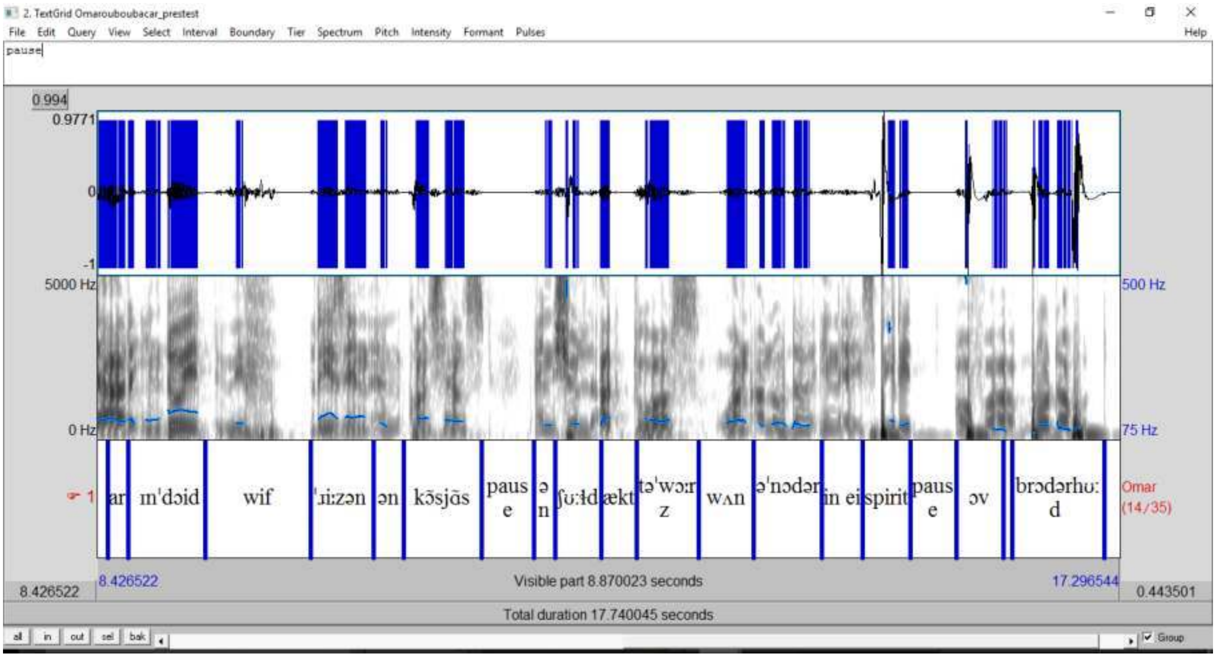
14. Moaiz (Part I & II)



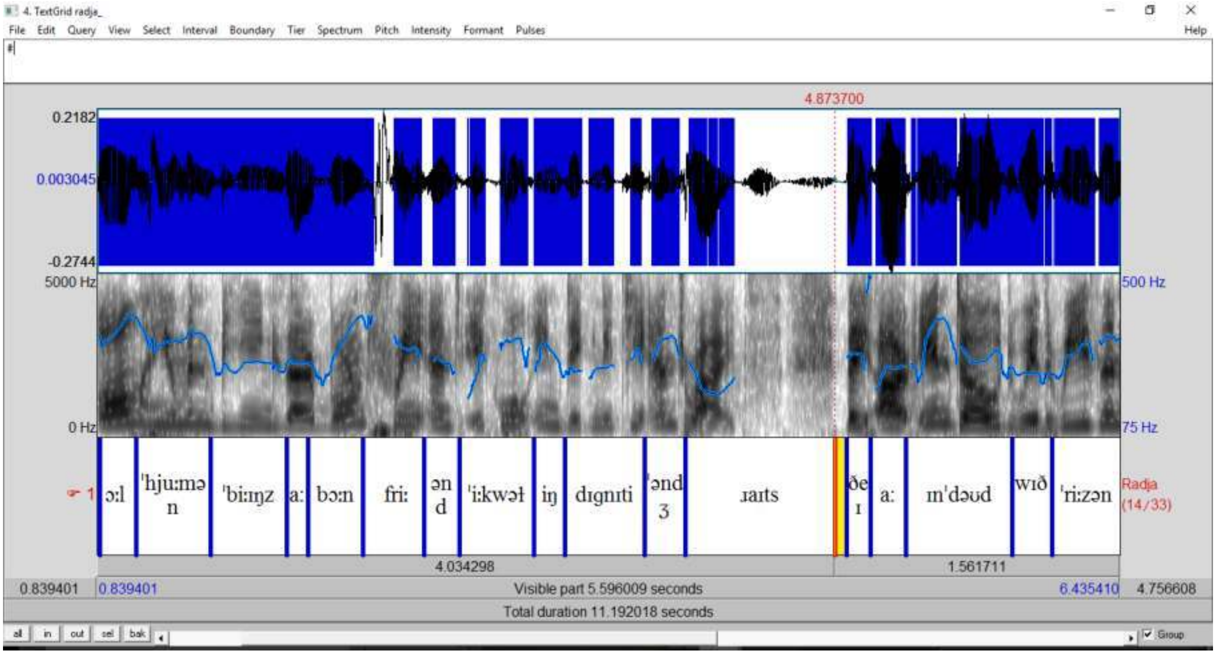


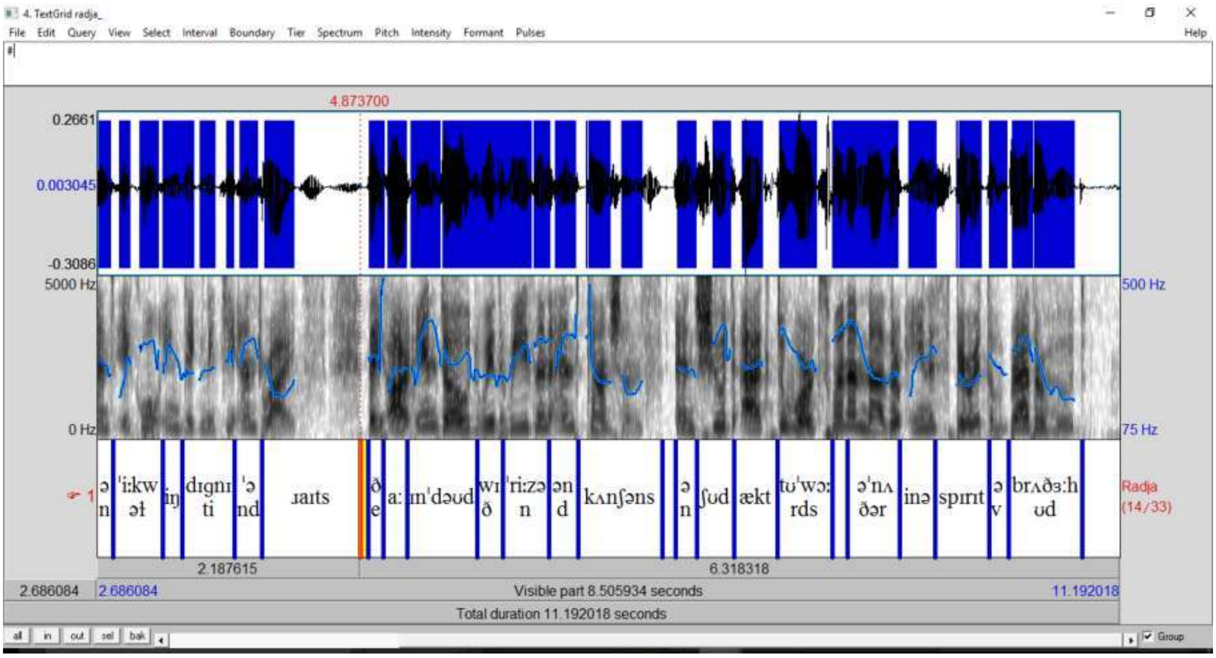
15. Omar Boubakar (Part1 & 2)



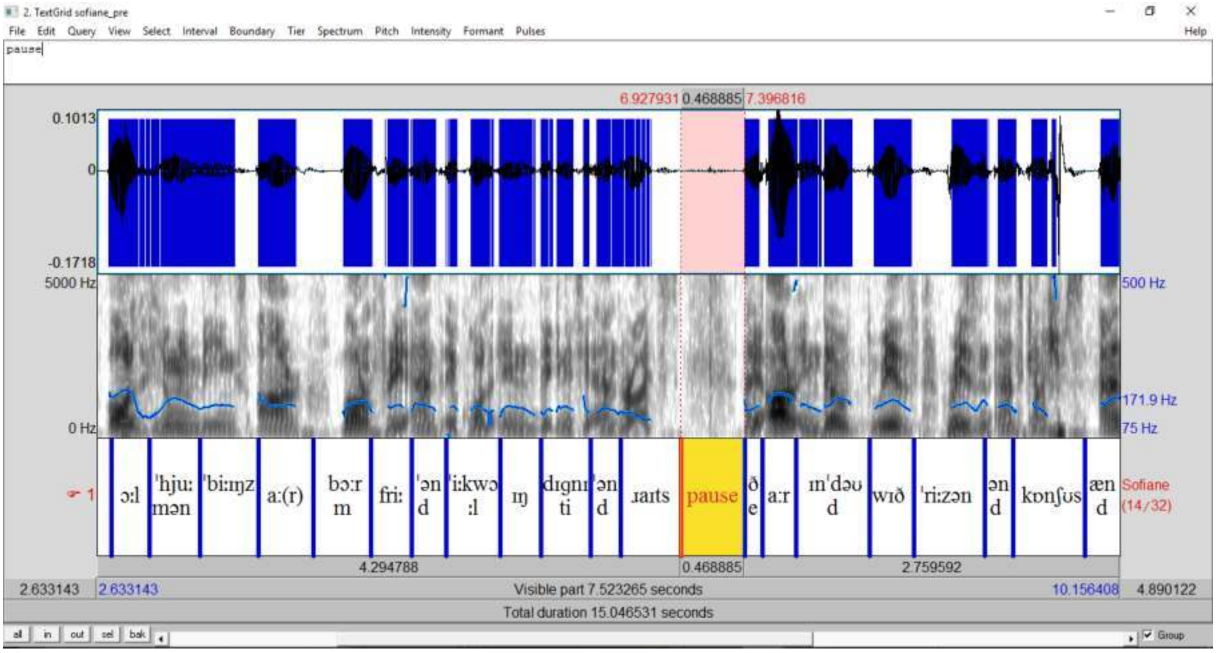


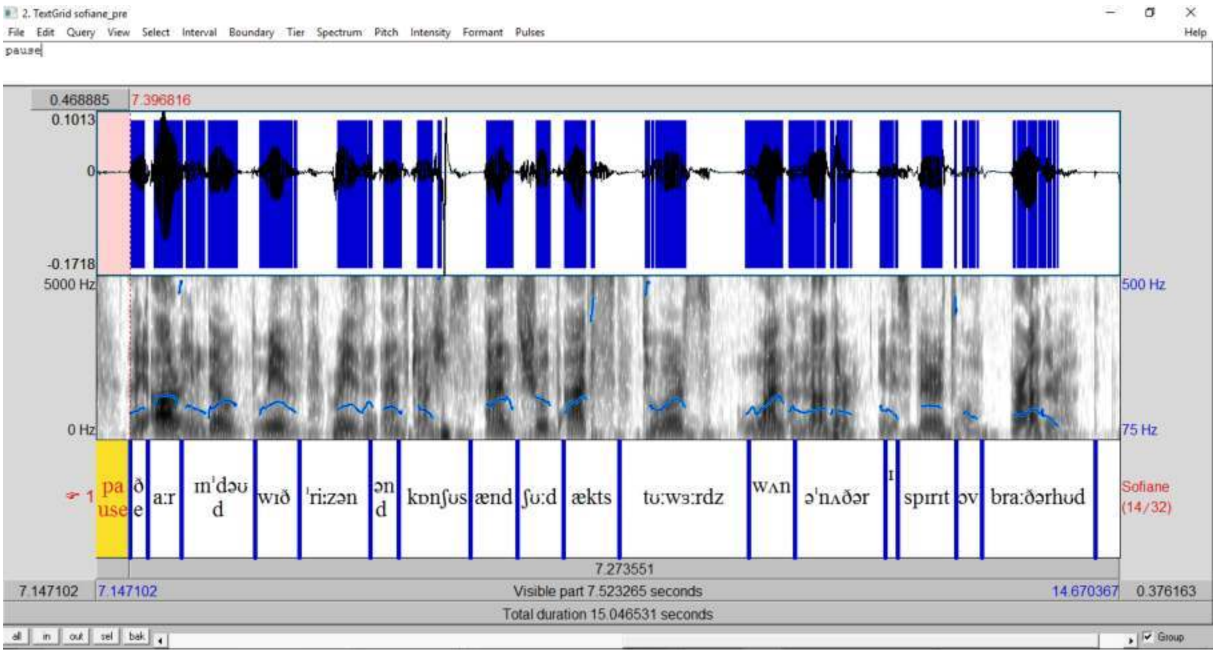
16. Radja (Part1&2)





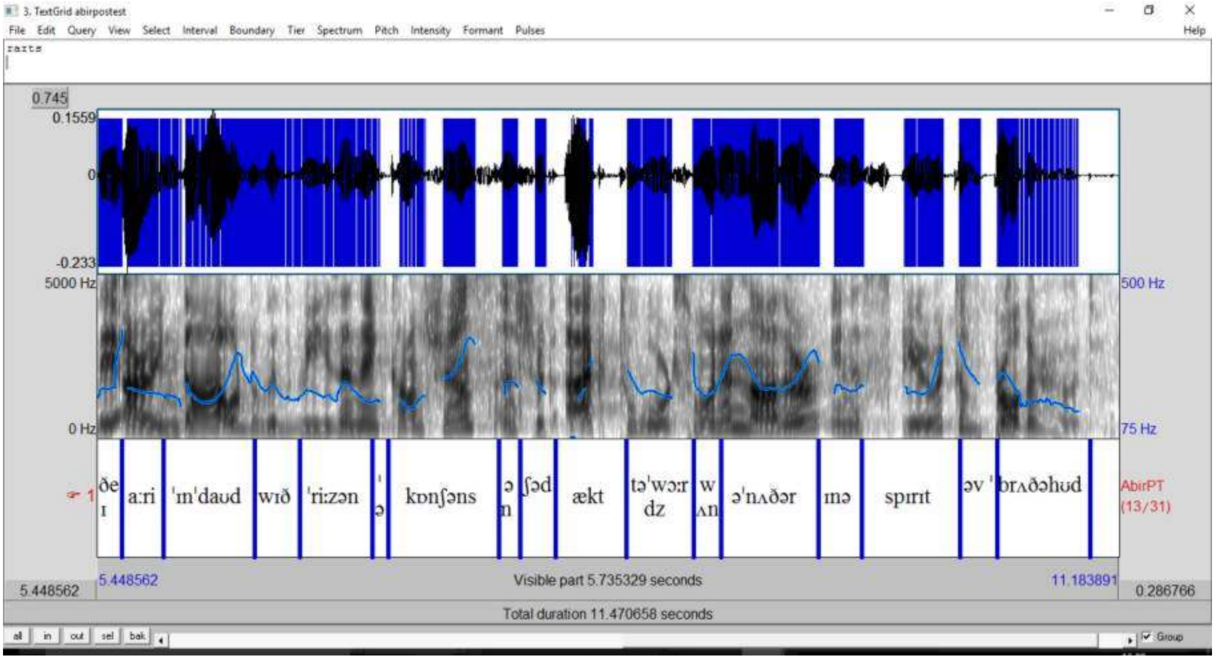
17. Sofiane (Part 1&2)

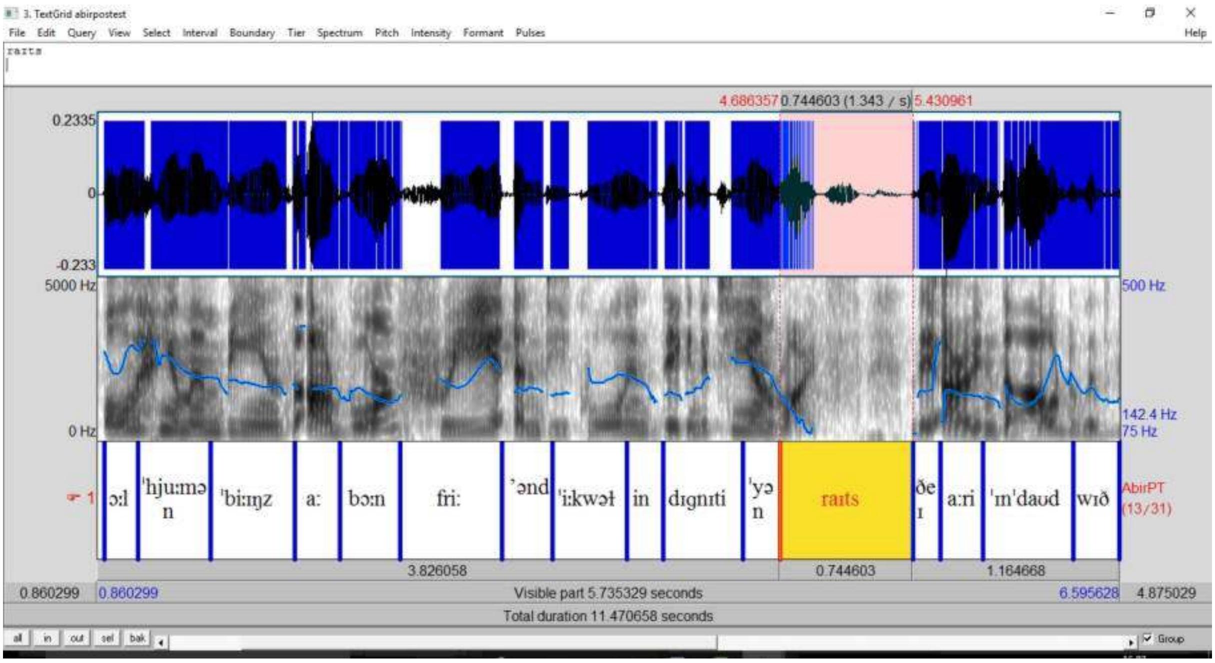

2. TextGrid sofiane_pre
File Edit Query View Select Interval Boundary Tier Spectrum Pitch Intensity Formant Pulses
Help
pause
6.927931 0.468885 7.396816
0.1013
0
-0.1718
5000 Hz
0 Hz
500 Hz
171.9 Hz
75 Hz
ɔ:l ˈhju:mən ˈbi:ɪŋz a:(r) bɔ:rm fri: ˈənd ˈi:kwɔ:l ɪŋ dɪgnɪti ˈənd ɹaɪts pause ðe a:r ɪnˈdəʊd wɪð ˈri:zən ənd kɒnʃʊs ænd
Sofiane (14/32)
4.294788 0.468885 2.759592
2.633143 2.633143 Visible part 7.523265 seconds 10.156408 4.890122
Total duration 15.046531 seconds


2. TextGrid sofiane_pre
File Edit Query View Select Interval Boundary Tier Spectrum Pitch Intensity Formant Pulses
Help
pause
0.468885 7.396816
0.1013
0
-0.1718
5000 Hz
0 Hz
500 Hz
75 Hz
pause ðe a:r ɪnˈdəʊd wɪð ˈri:zən ənd kɒnʃʊs ænd ʃʊ:d ækts tʊ:wɜ:rdz wʌn əˈnʌðər ɪ spɪrɪt ɒv bra:ðərhʊd
Sofiane (14/32)
7.273551
7.147102 7.147102 Visible part 7.523265 seconds 14.670367 0.376163
Total duration 15.046531 seconds

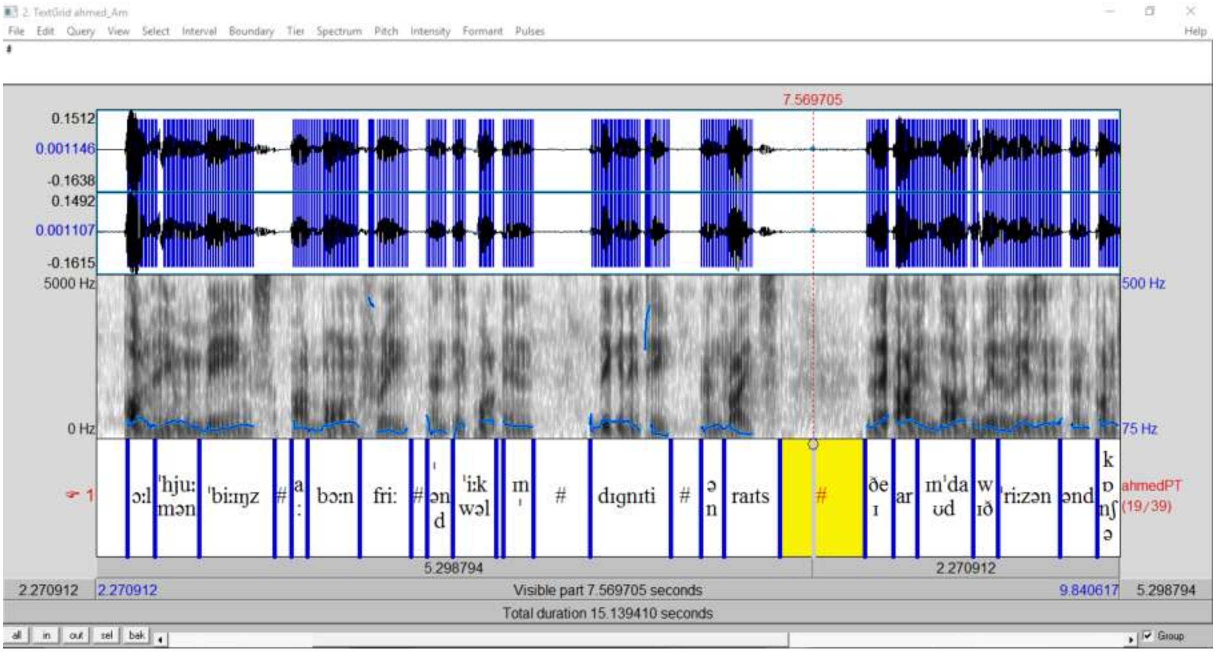
B. Post-Test

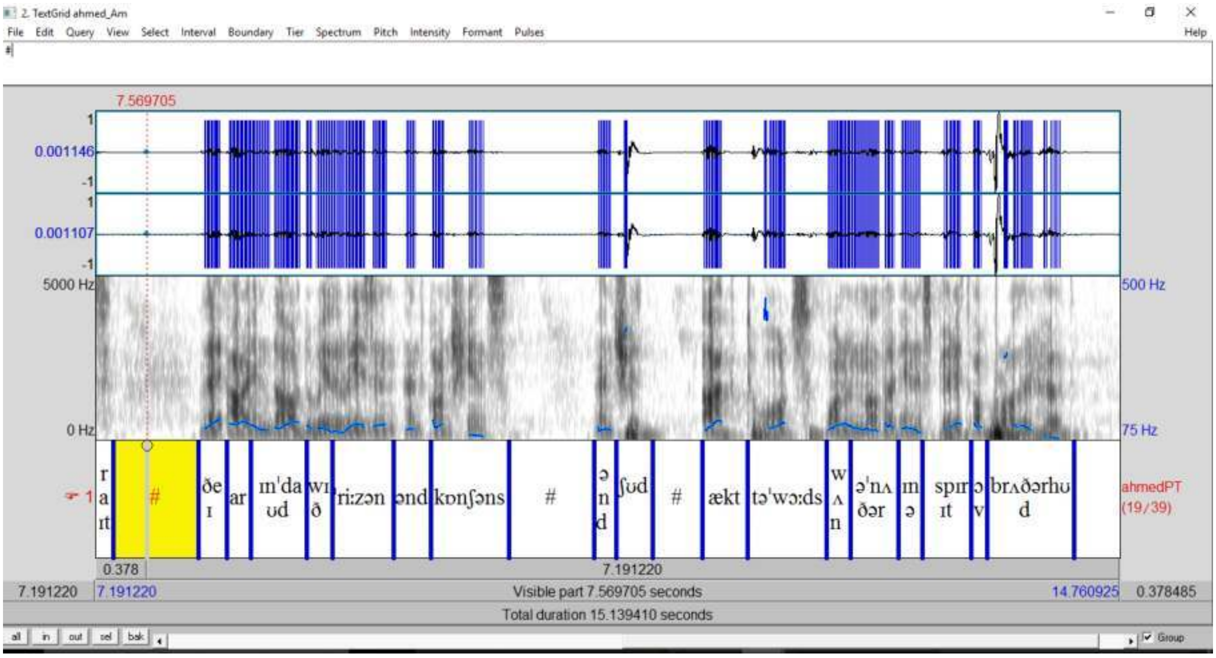
1. Abir (Part 1 &2)



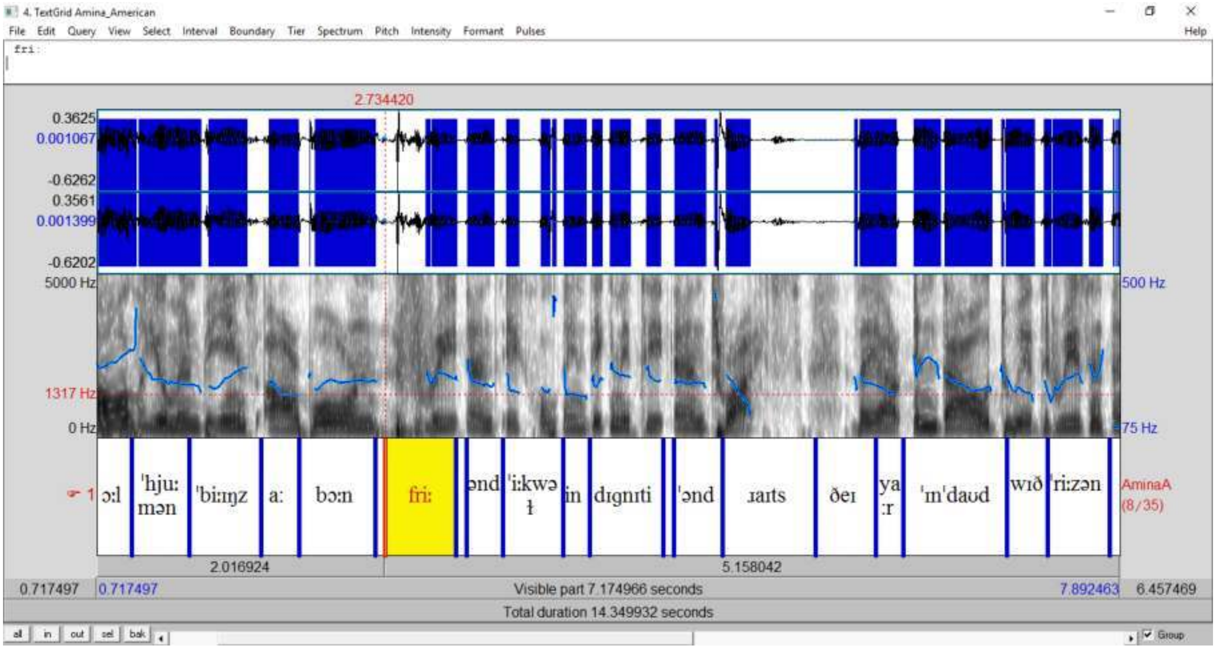


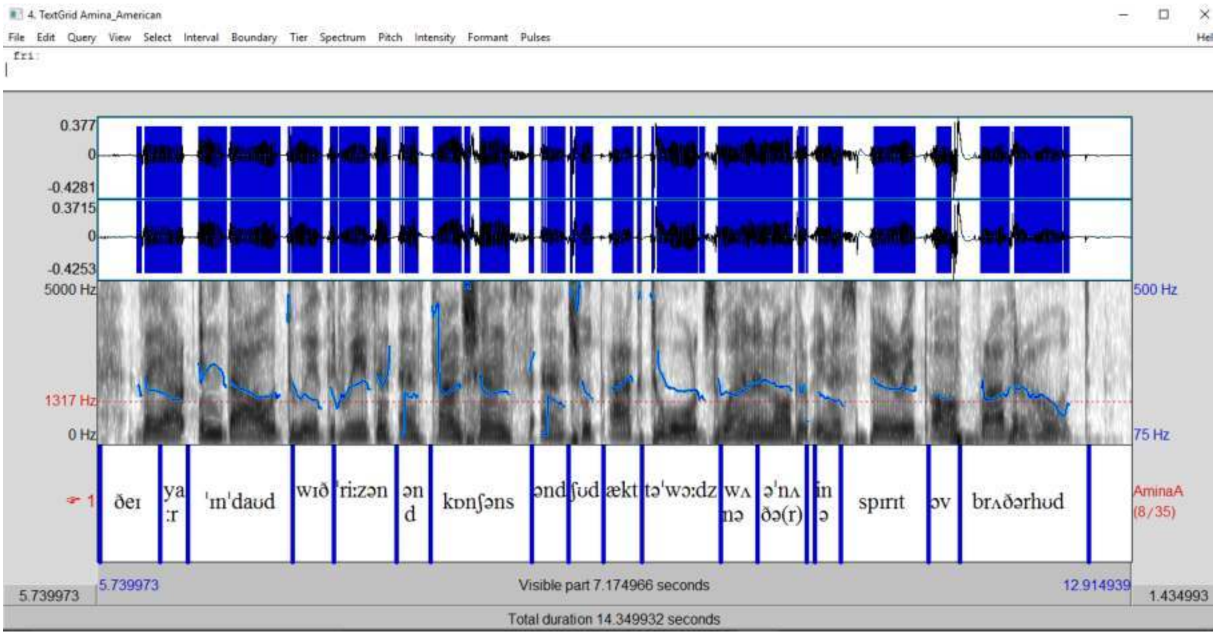
2. Ahmed (Part 1&2)



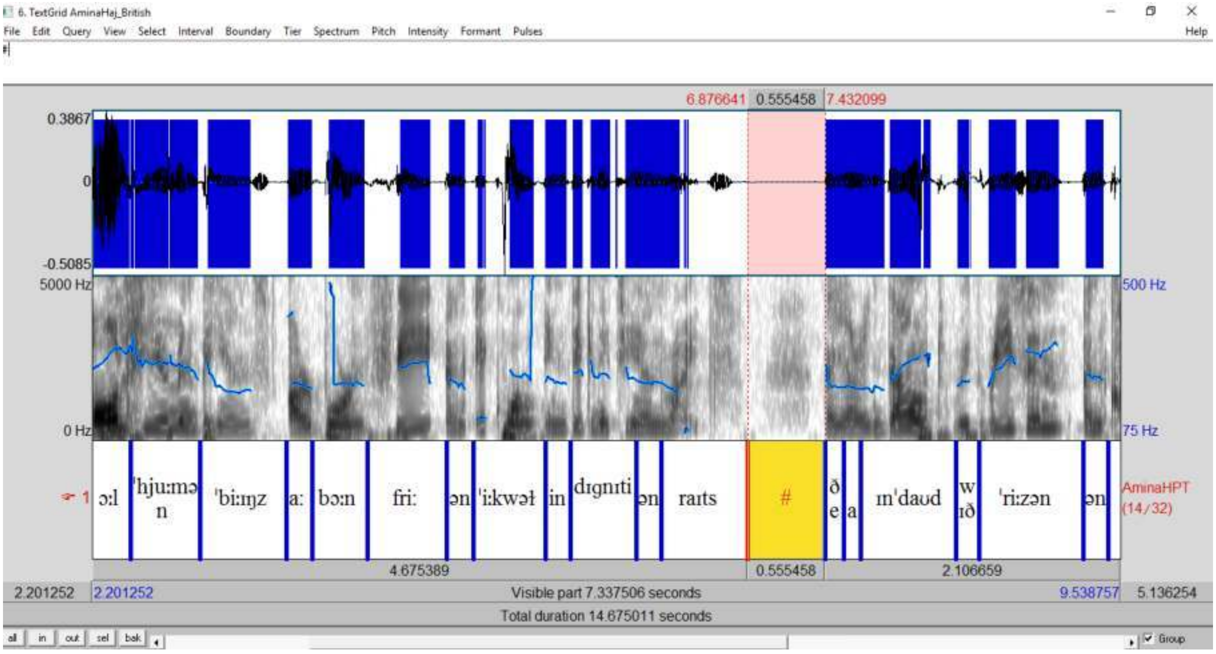


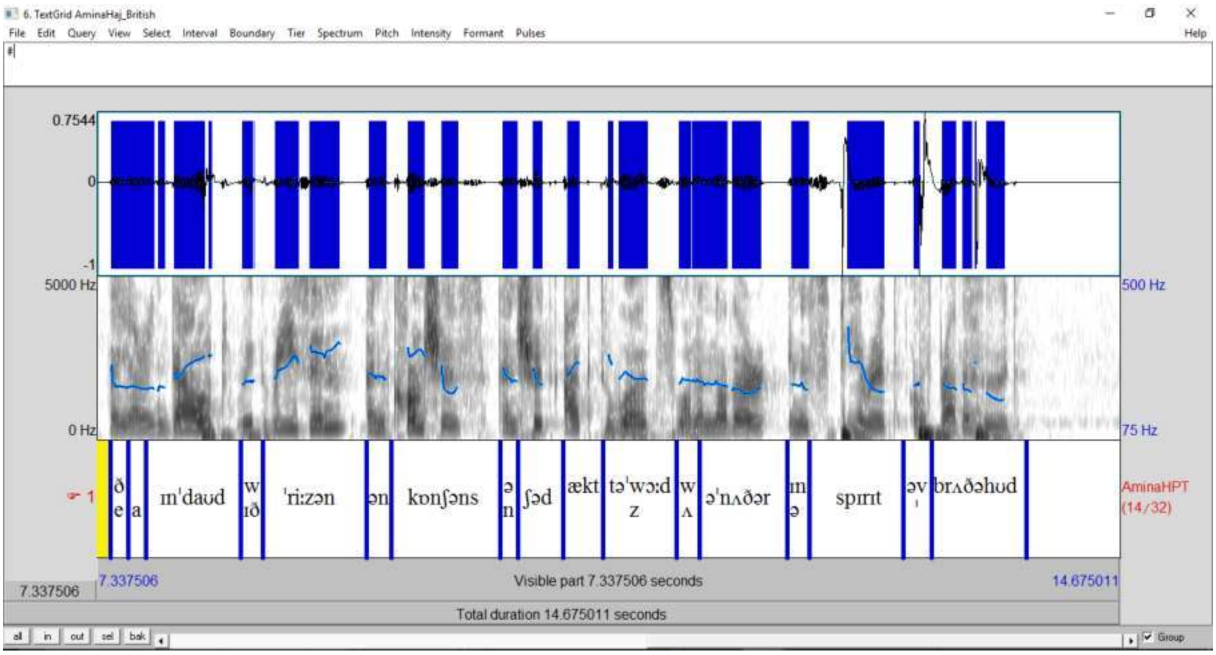
3. Amina A (Part 1&2)



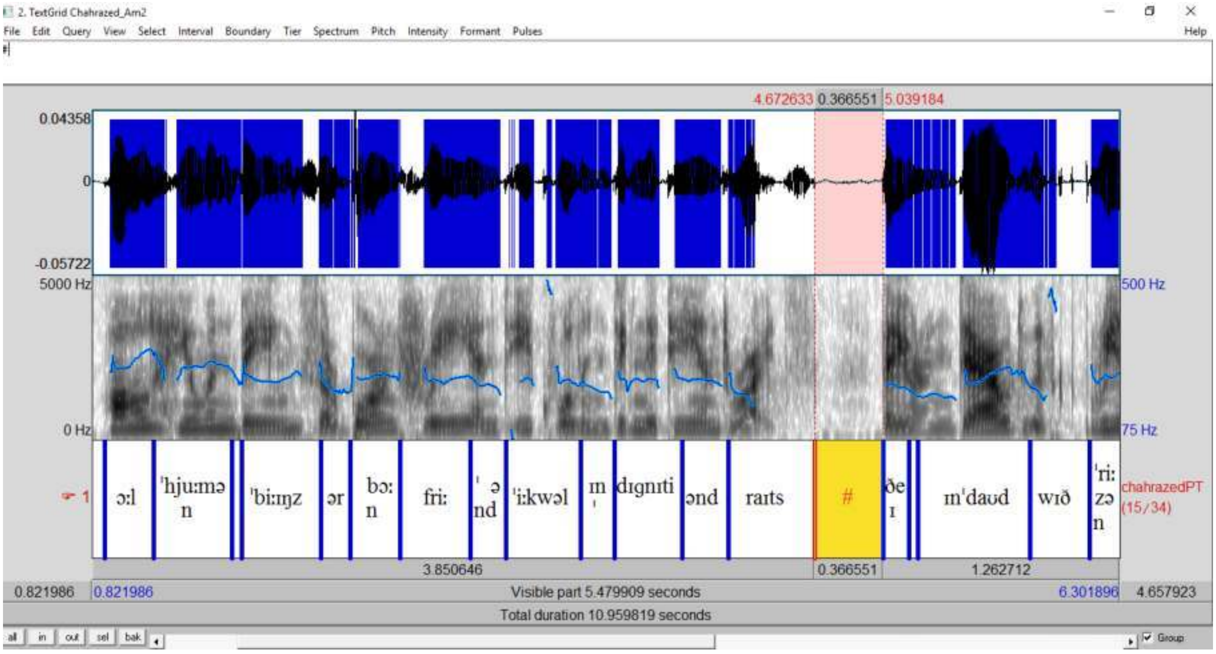


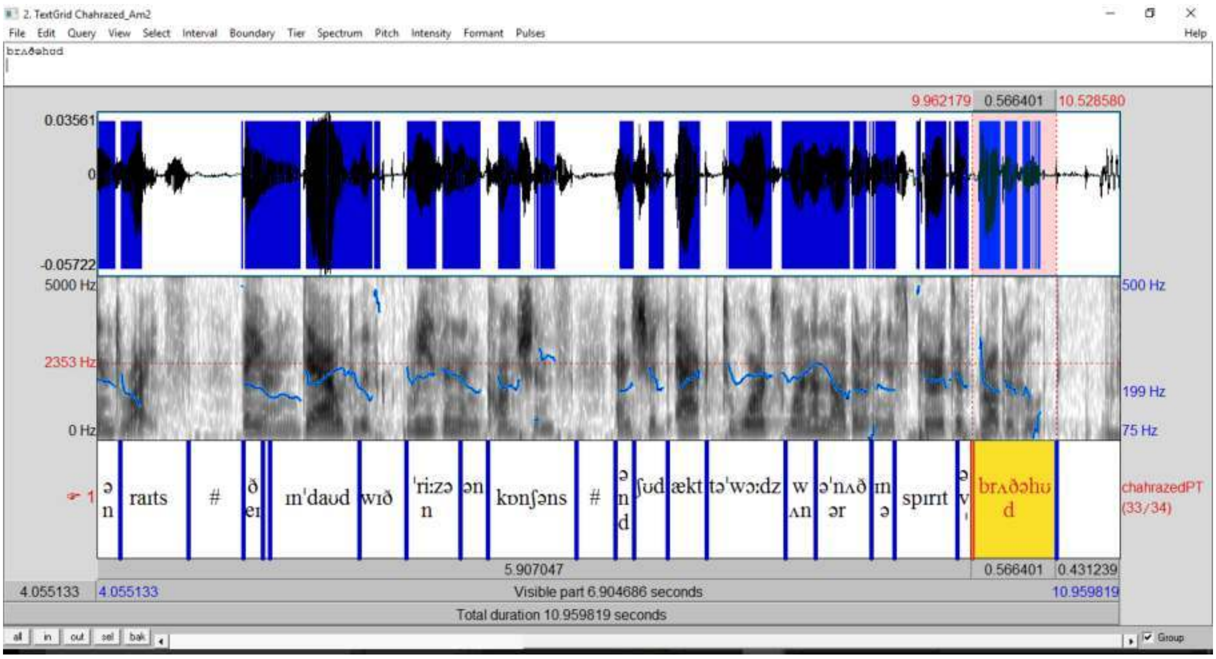
4. Amina H (Part 1&2)



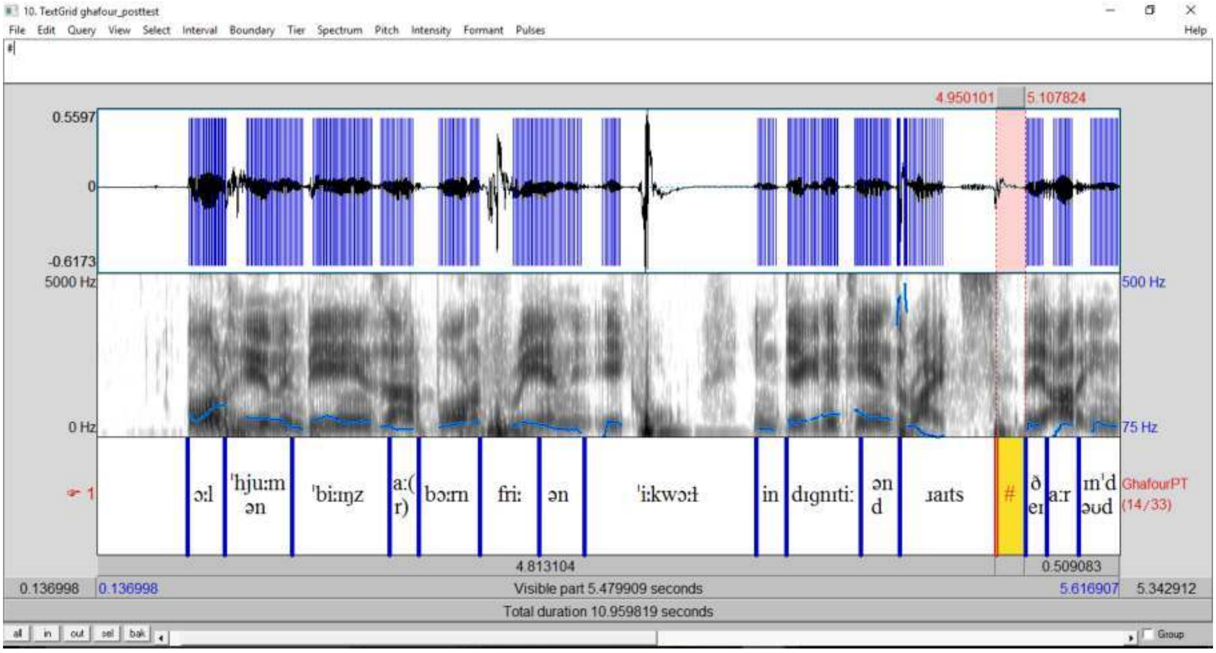


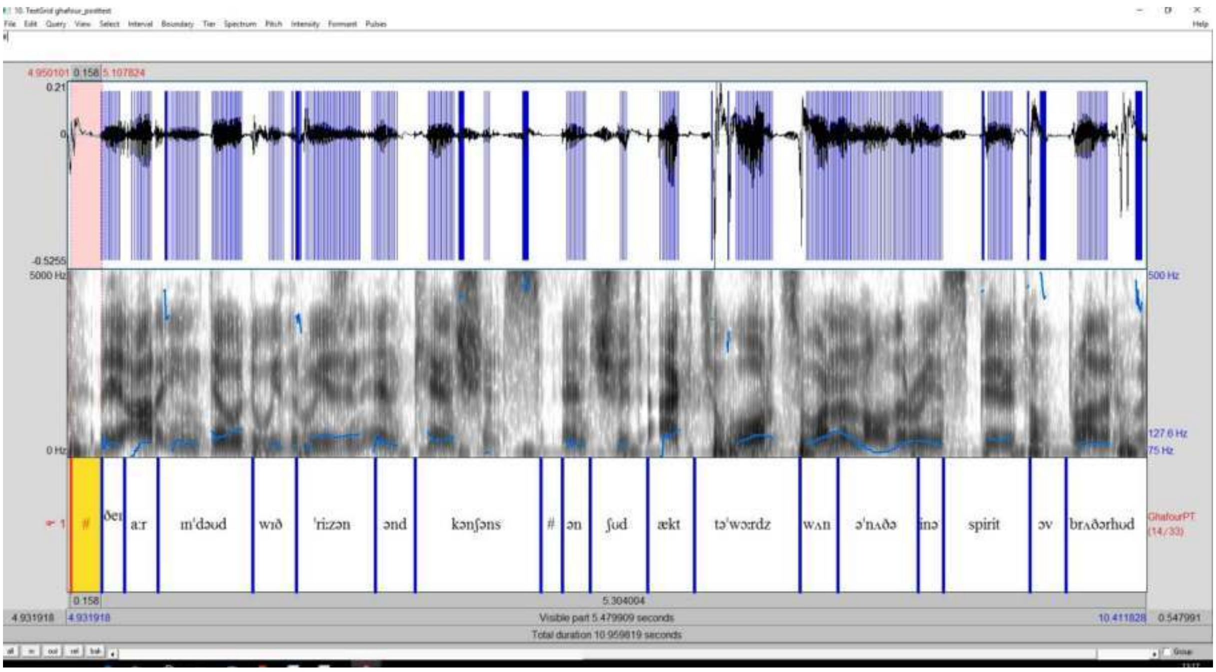
5. **Chahrazed (Part1&2)**



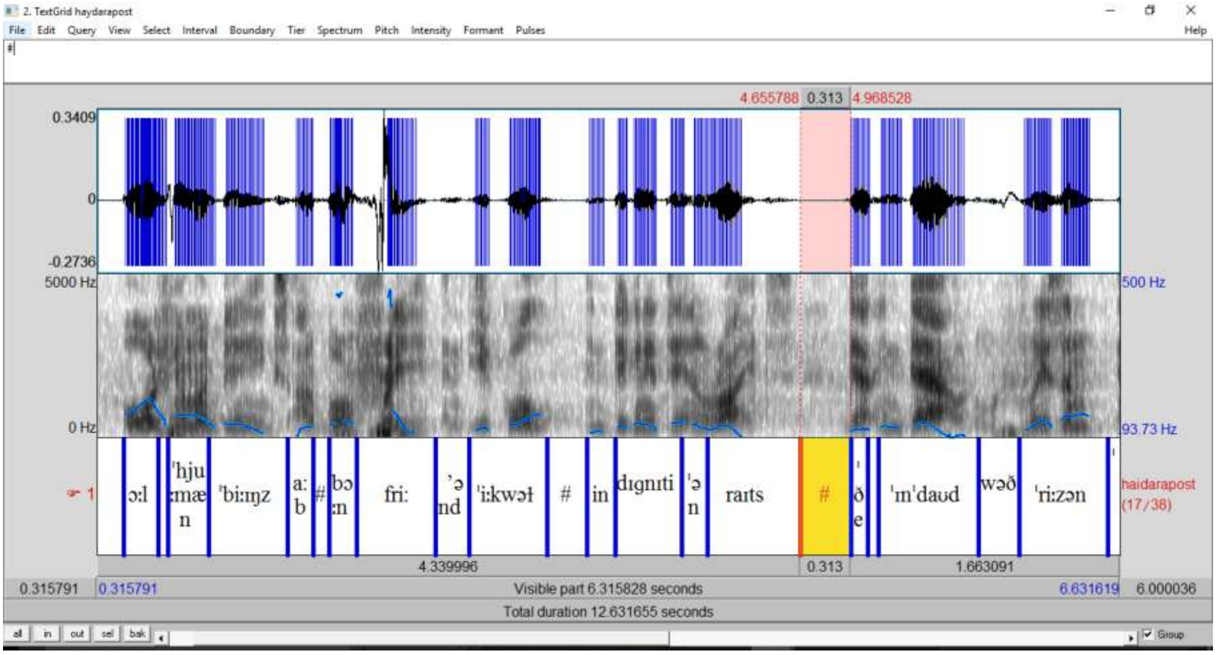


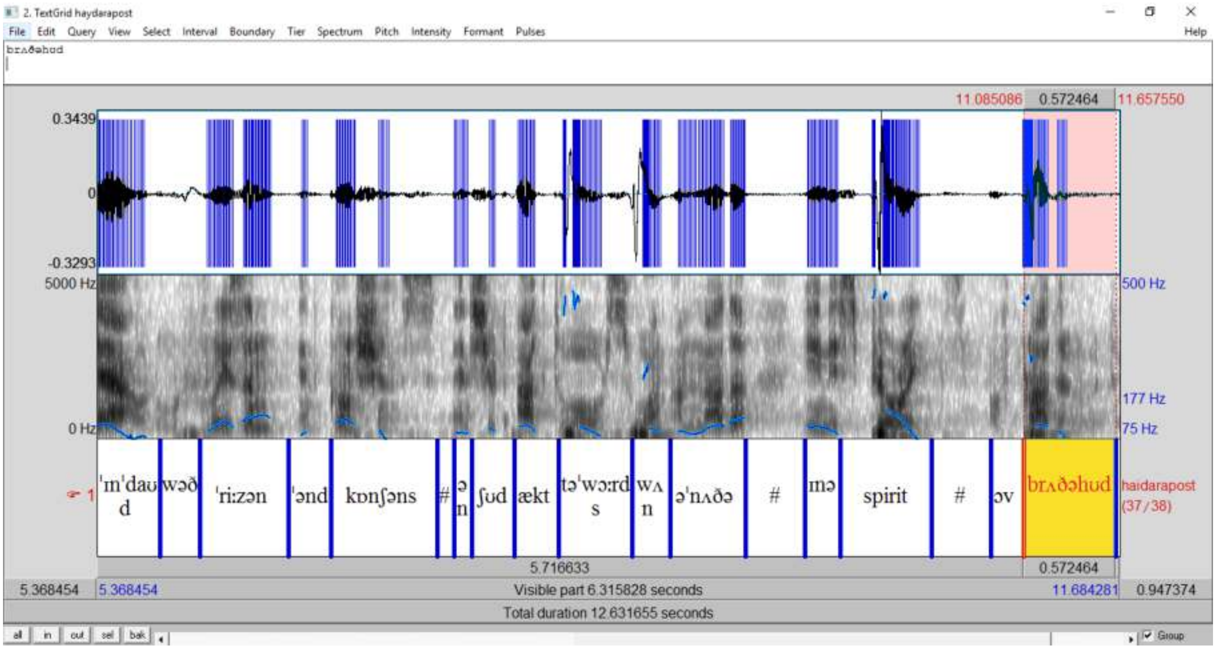
7. Ghafour (Part 1&2)



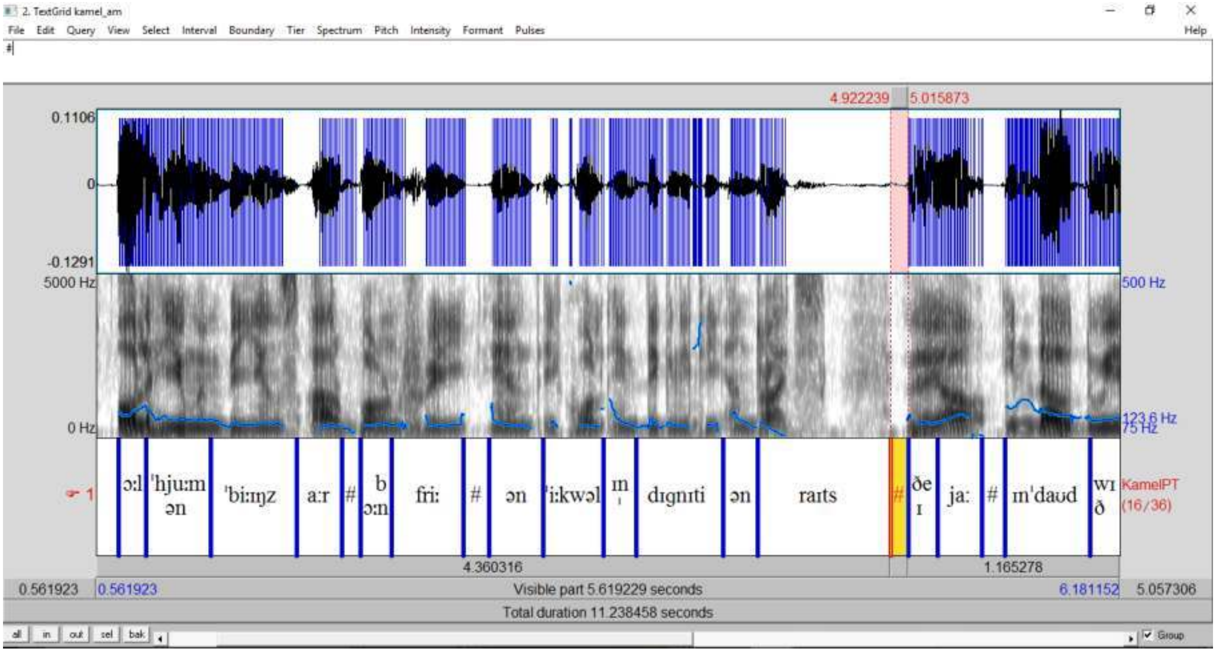


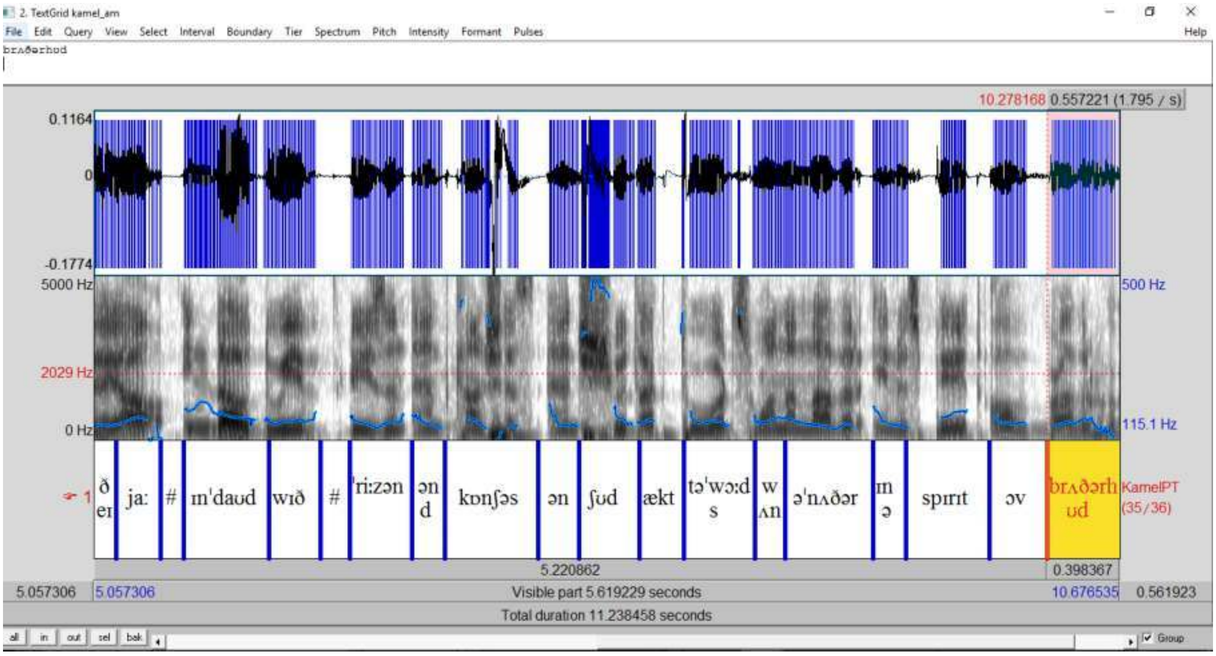
8. Haidara (Part 1&2)



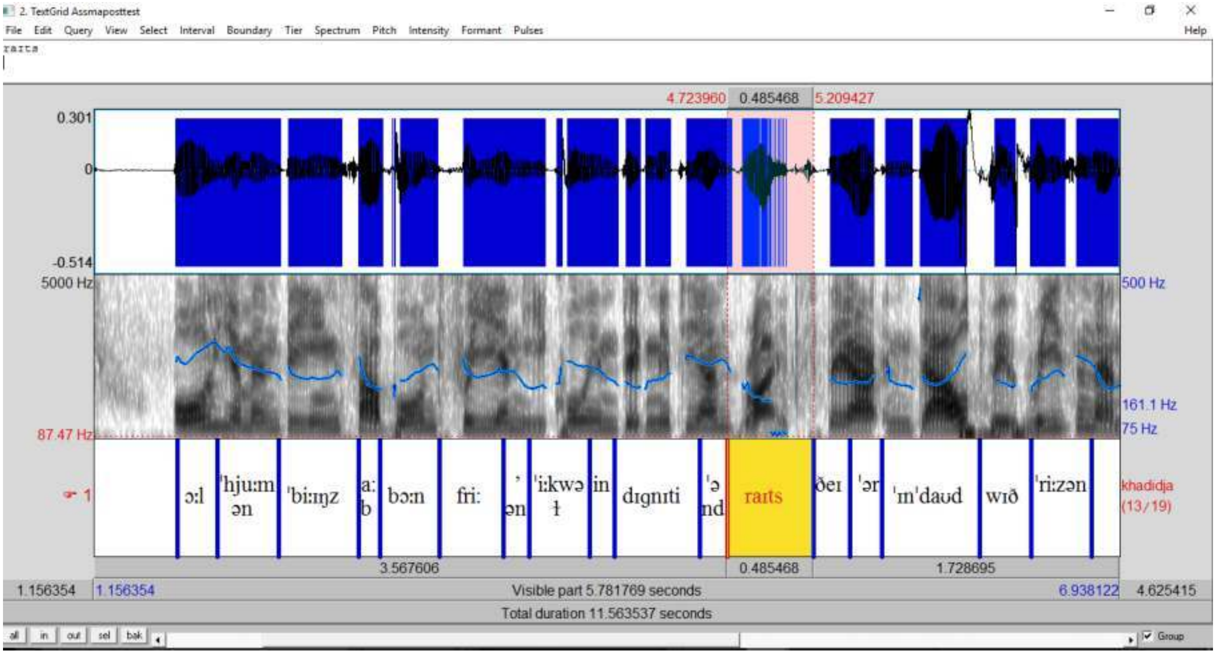


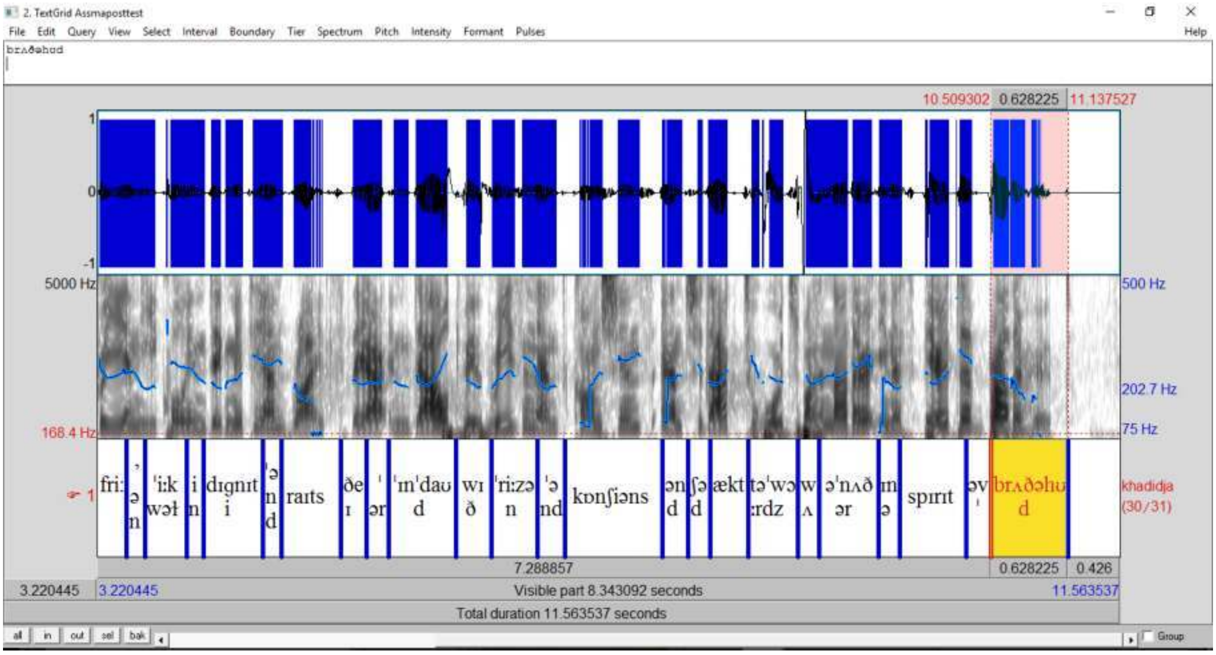
9. Kamel (Part 1&2)



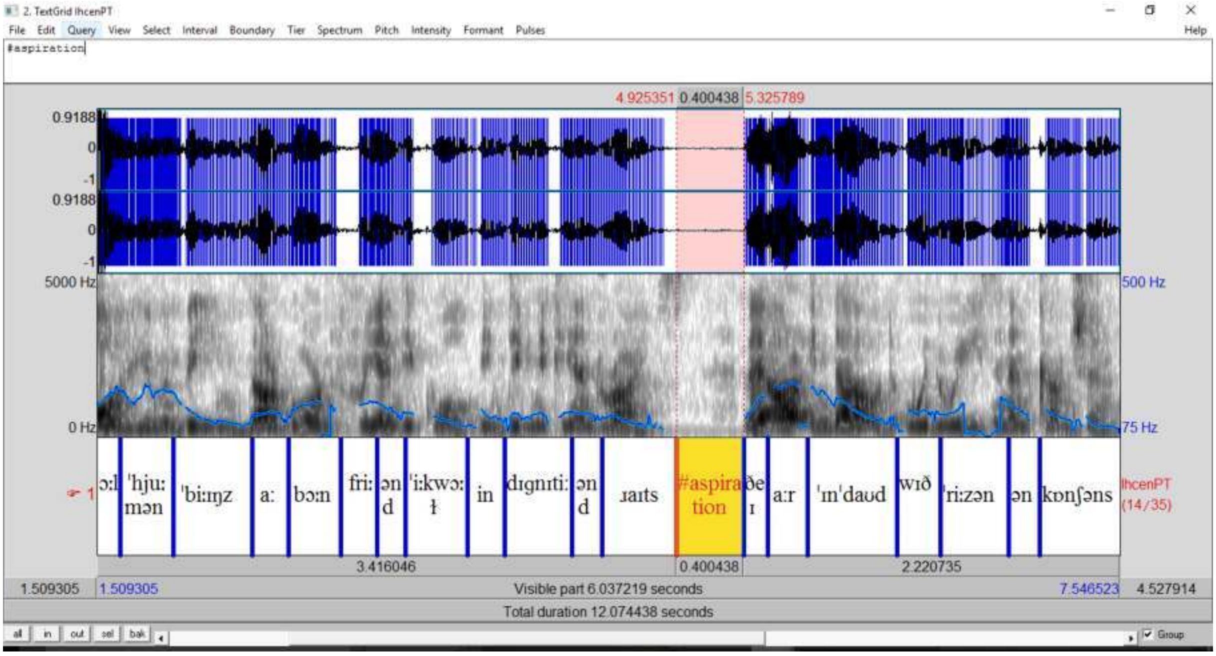


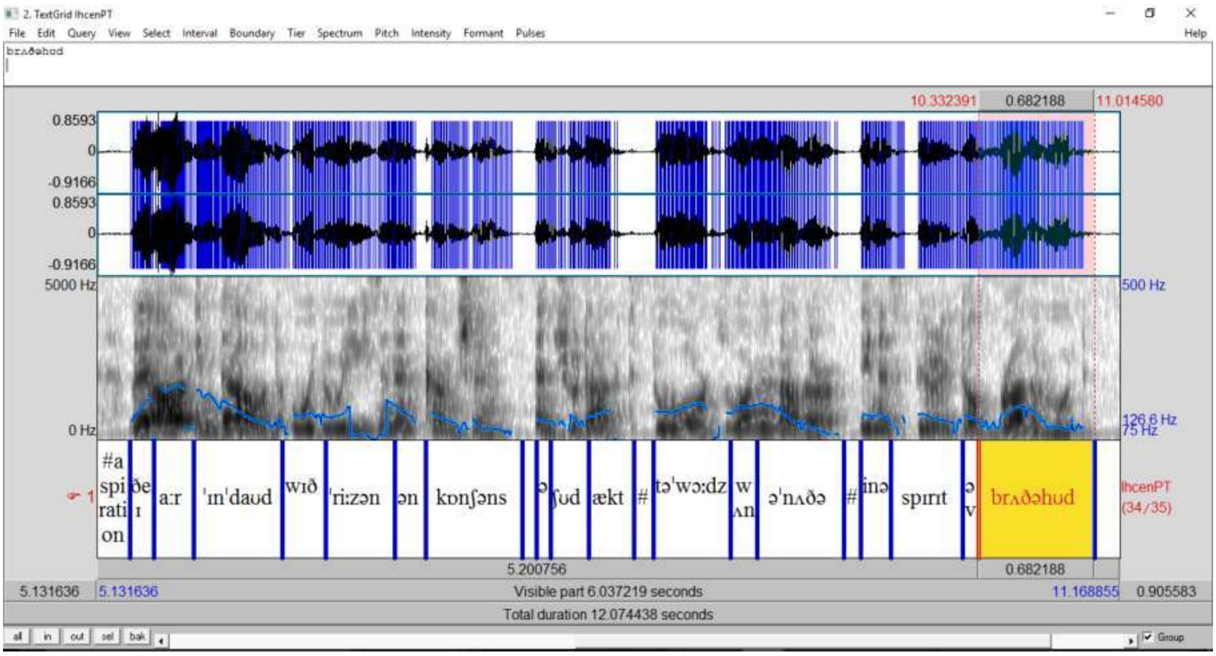
10. KhadidjaAssma (Part 1&2)



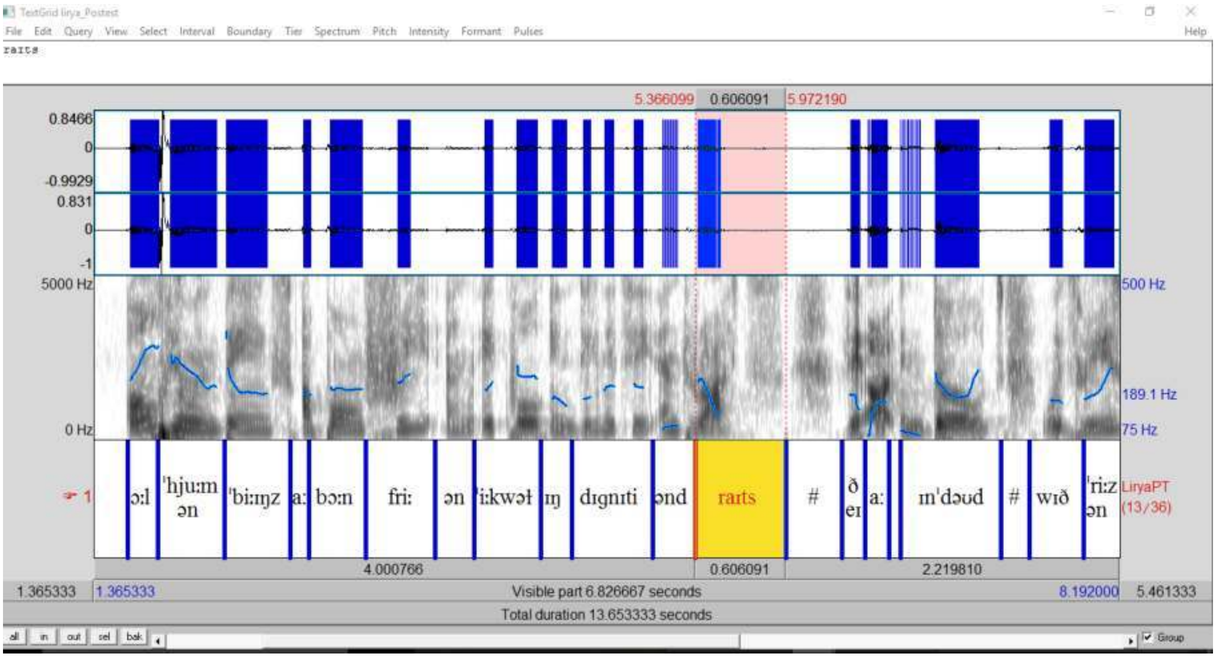


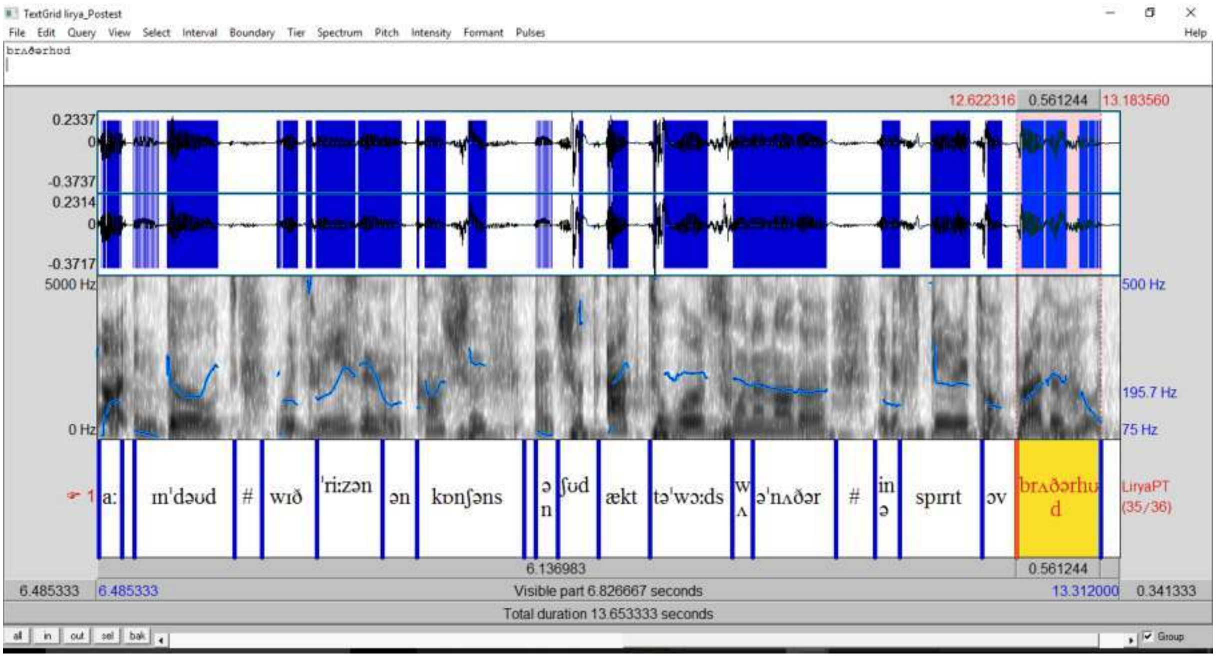
11. Ihcen (Part1&2)



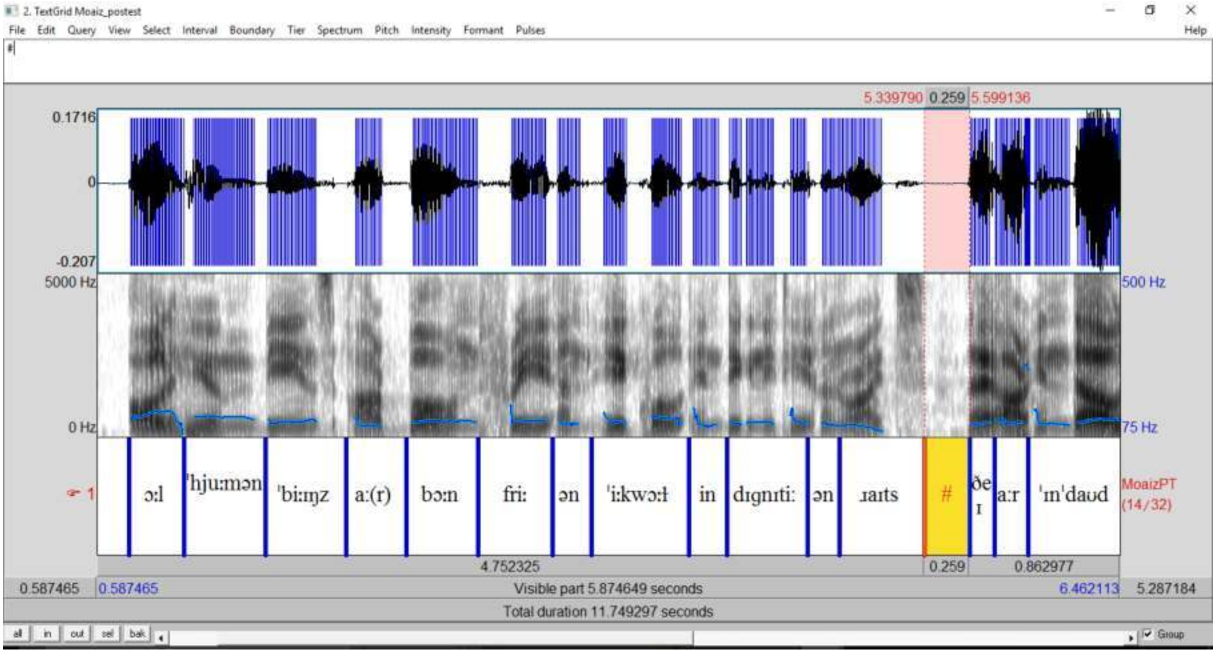


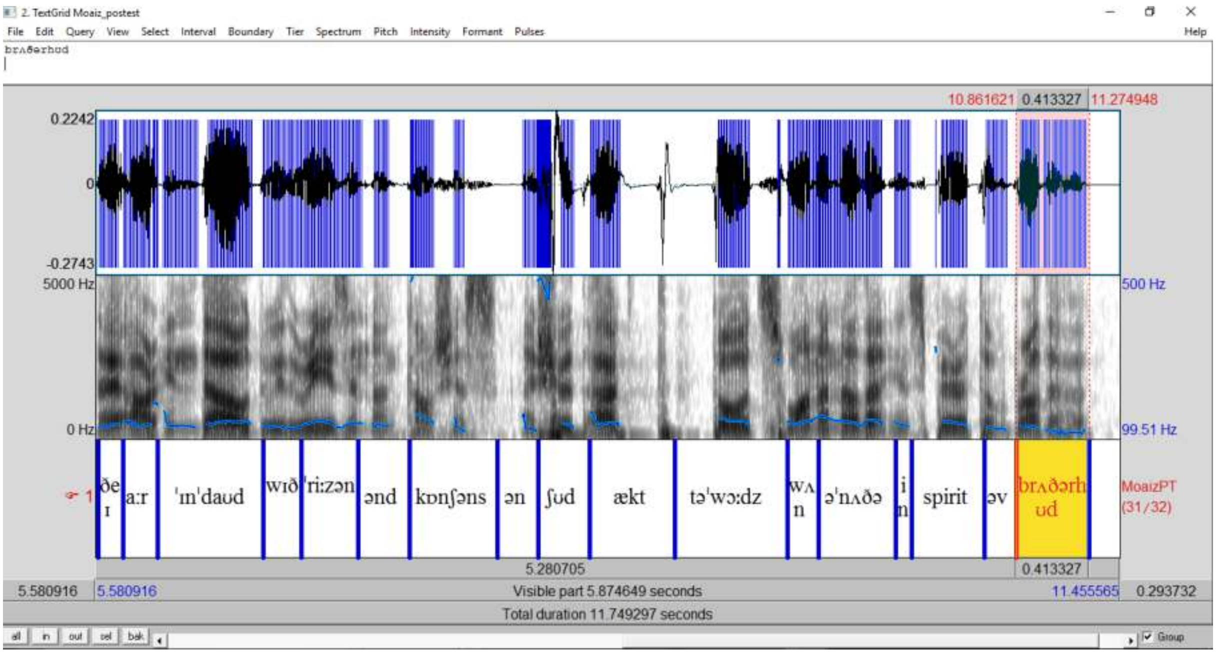
12. Liyria (Part1&2)



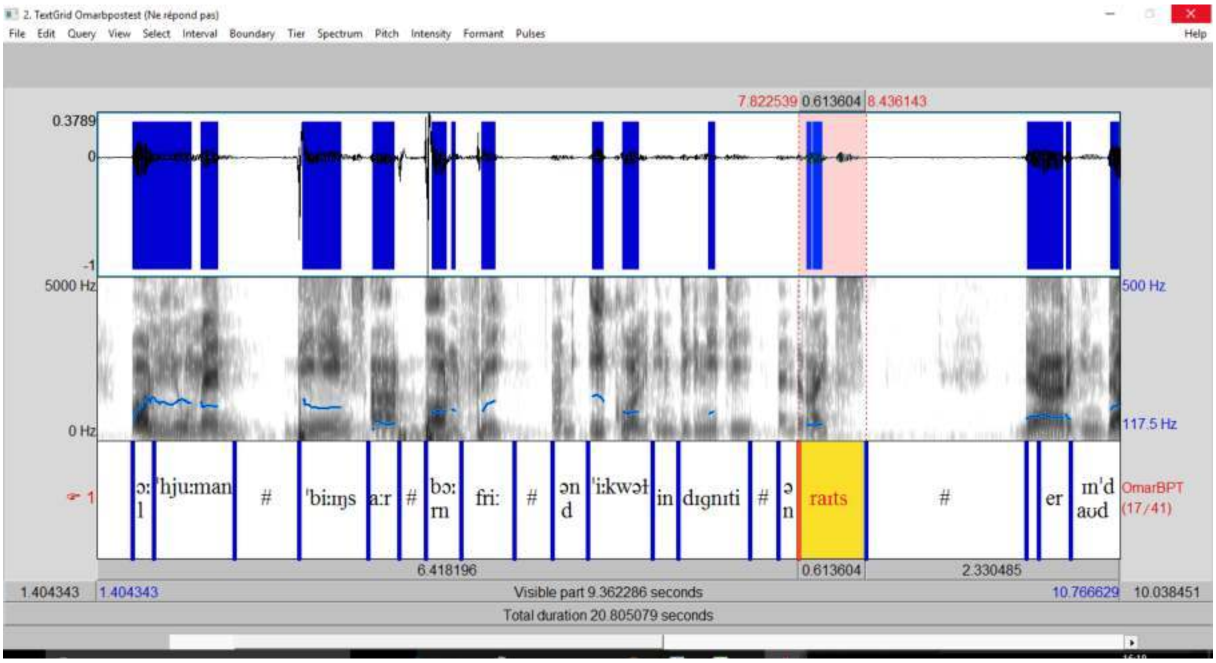


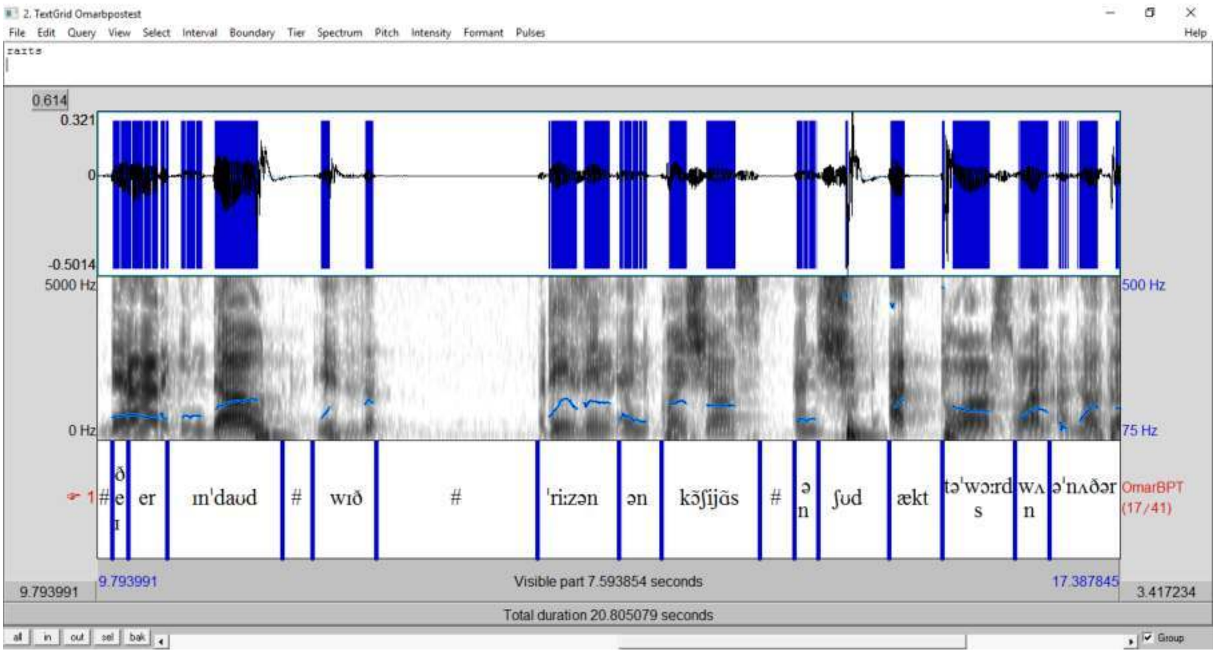
13. Moaiz (Part 1&2)

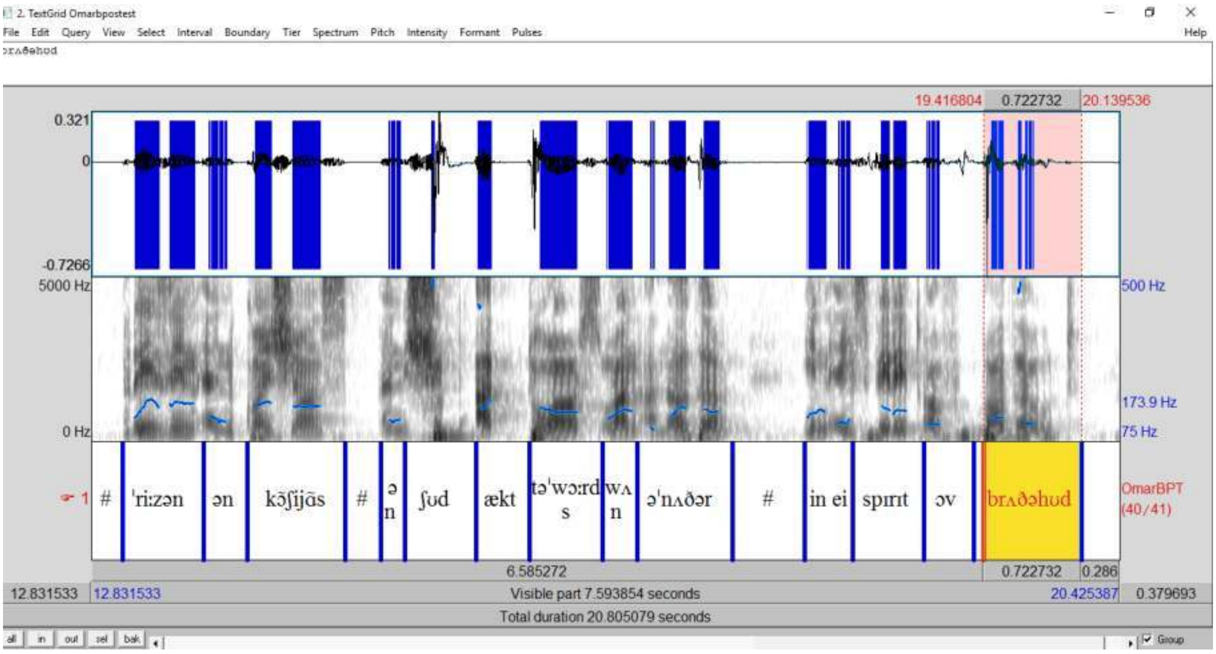




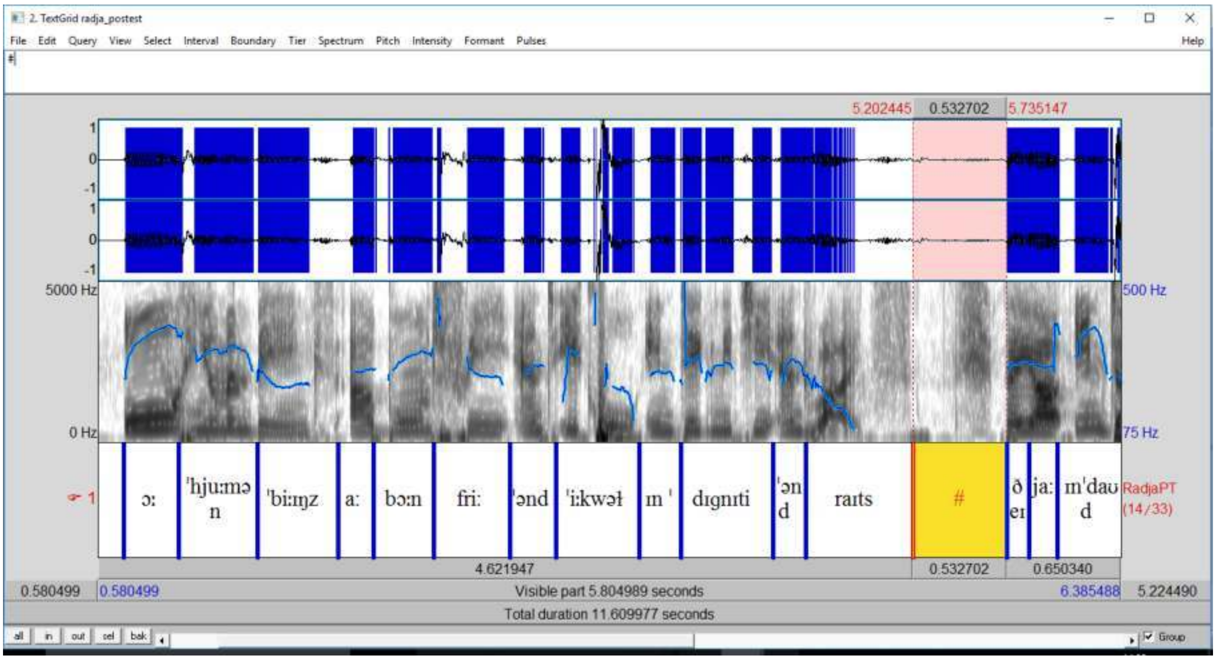
14. Omar Boubakar (Part1, 2&3)

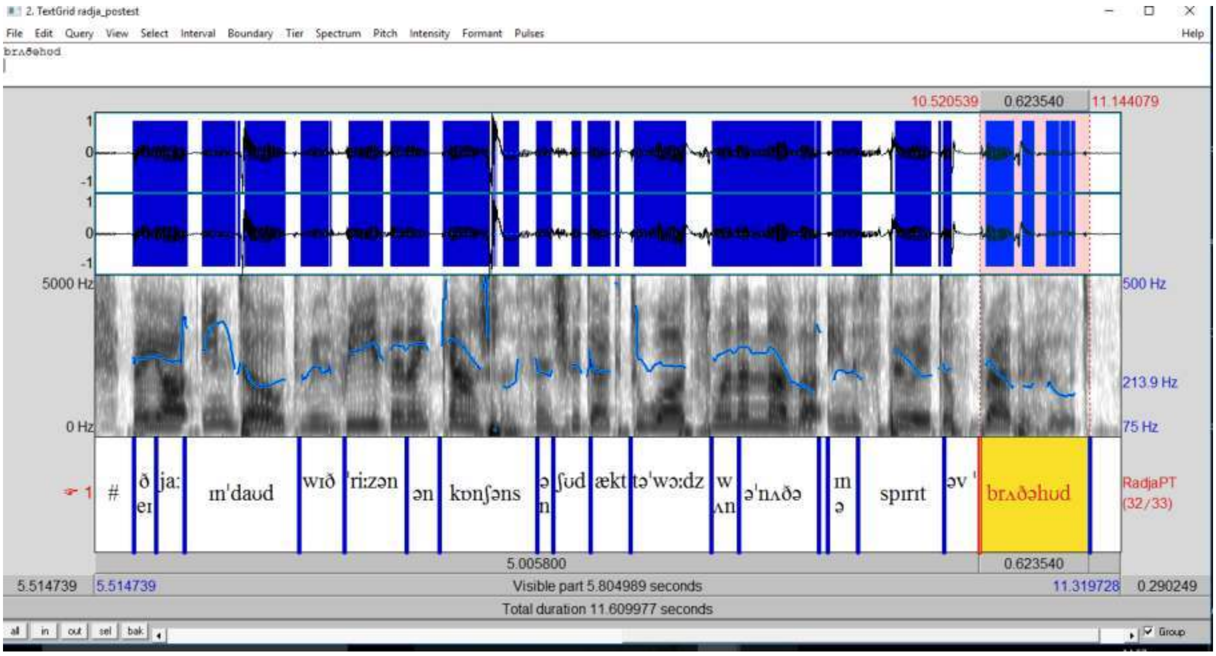




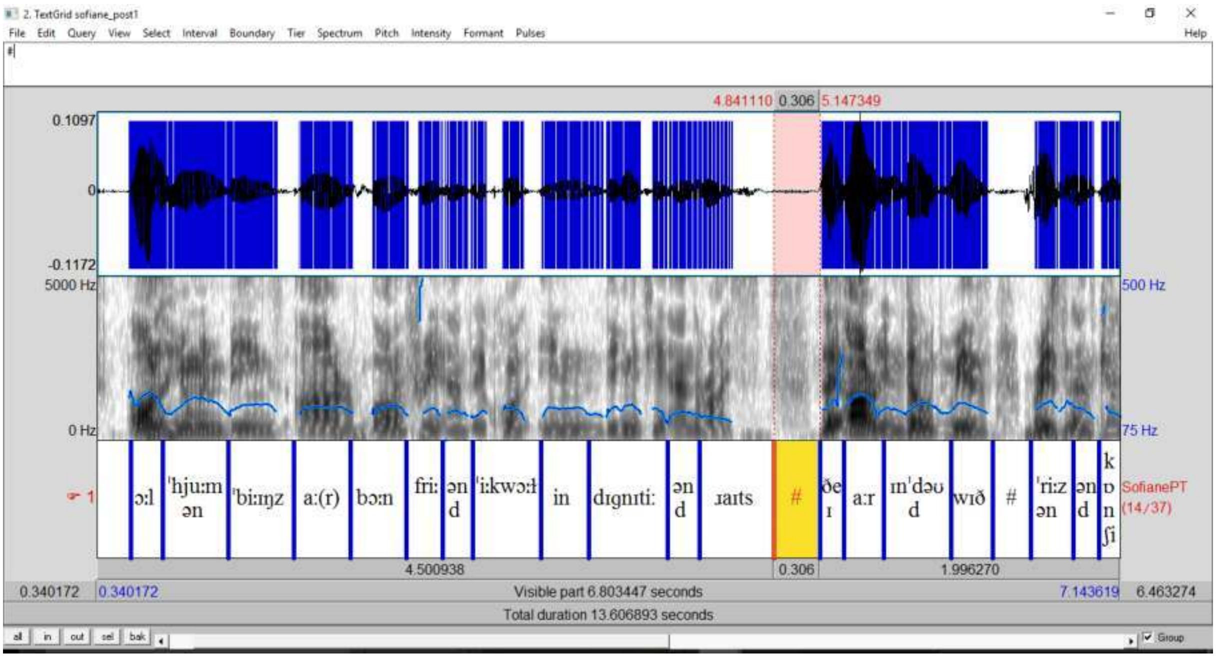


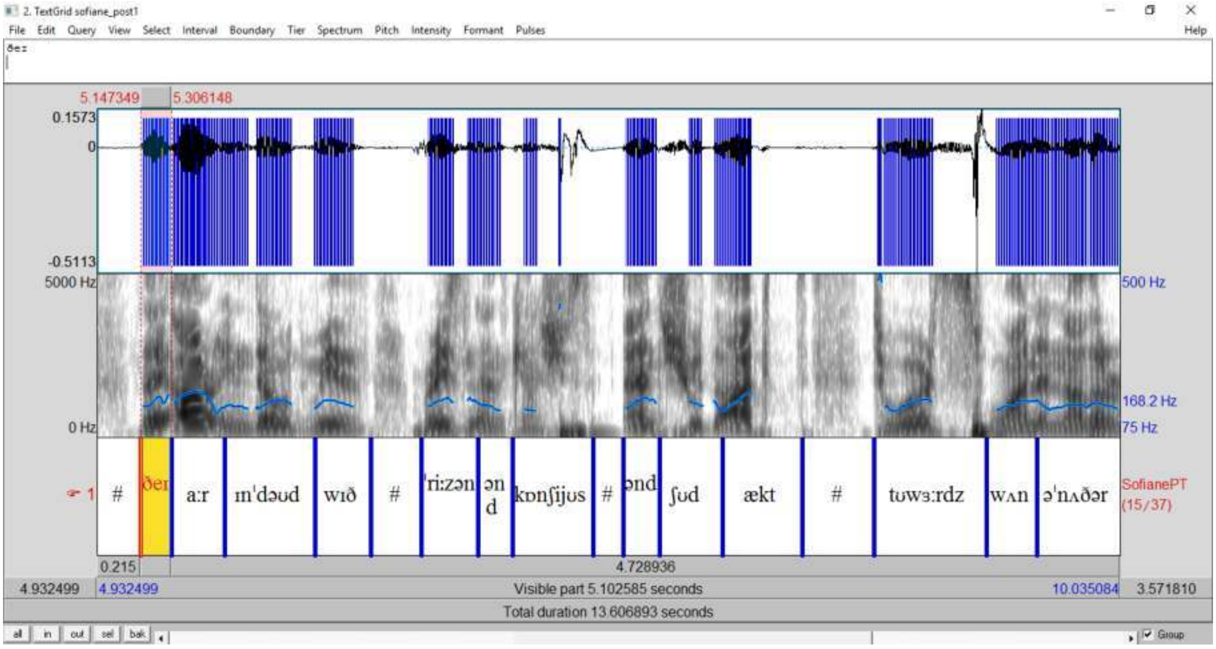
15. Radja (Part1&2)

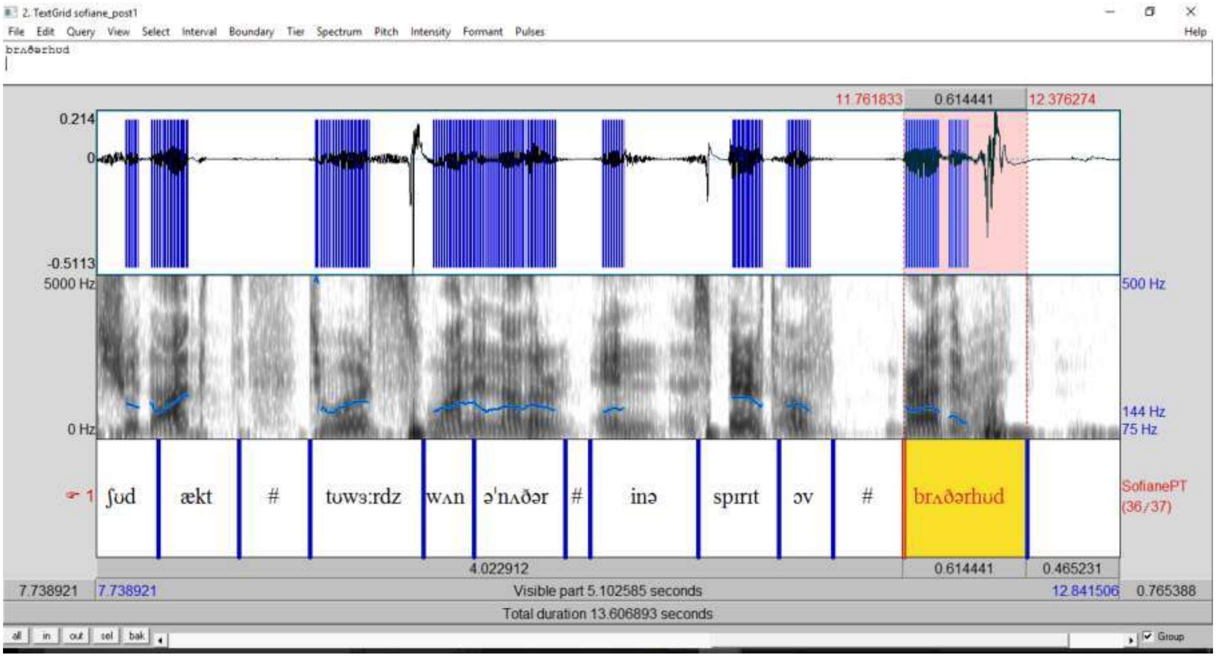




16. Sofiane (Part 1, 2 &3)




2. TextGrid sofiane_post1
File Edit Query View Select Interval Boundary Tier Spectrum Pitch Intensity Formant Pulses Help
ðeɪ
5.147349
5.306148
0.1573
0
-0.5113
5000 Hz
0 Hz
500 Hz
168.2 Hz
75 Hz
ðeɪ a:r ɪn'dəʊd wɪð # 'ri:zən ən d kɒnʃijʊs # ənd ʃʊd ækt # tʊwɜ:rdz wʌn ə'nʌðər
SofianePT (15/37)
0.215
4.728936
4.932499
4.932499
Visible part 5.102585 seconds
10.035084
3.571810
Total duration 13.606893 seconds


2. TextGrid sofiane_post1
File Edit Query View Select Interval Boundary Tier Spectrum Pitch Intensity Formant Pulses Help
brʌðərhʊd
11.761833
0.614441
12.376274
0.214
0
-0.5113
5000 Hz
0 Hz
500 Hz
144 Hz
75 Hz
ʃʊd ækt # tʊwɜ:rdz wʌn ə'nʌðər # inə spɪrɪt ɔv # brʌðərhʊd
SofianePT (36/37)
4.022912
0.614441
0.465231
7.738921
7.738921
Visible part 5.102585 seconds
12.841506
0.765388
Total duration 13.606893 seconds

6. **Students' and Native Speakers' Recordings: (CD)**

7. **Teachers' Interview (CD)**

8. **A History of Speech Science in People and Ideas**

Date	Person or Idea
19th cent.	Alexander Ellis (1814-90) developed the idea of phonetic transcription as a way to 'indicate the pronunciation of any language with great minuteness'. He also based hsi phonetic symbols on the roman alphabet, just the IPA does today. Henry Sweet (1845-1912) introduced the idea that the fidelity of transcription within a language could be based on lexical contrast, the basis of the later phonemic principle. He also transcribed the "educated London speech" of his time, giving us the first insights into Received Pronunciation. Hermann von Helmholtz (1821-94) was a physicist who applied ideas of acoustic resonance to the vocal tract and to the hearing mechanism. Thomas Edison (1847-1931) invented sound recording which allowed careful listening and analysis of speech for the first time. Paul Passy (1859-1940) was a founding member and the driving force behind the International Phonetic Association. He published the first IPA alphabet in 1888.
1900-1929	Daniel Jones (1881-1967) founded the Department of Phonetics at UCL and was its head from 1921-1947. He is famous for the first use of the term 'phoneme' in

	its current sense and for the invention of the cardinal vowel system for characterising vowel quality.
1930s	The First International Congress of Phonetic Sciences was held in 1932 in Amsterdam. Between 1928 and 1939, the *Prague School* gave birth to phonology as a separate field of study from phonetics. Key figures were Nikolai Trubetzkoy (1890-1938) and Roman Jakobson (1896-1982).
1940s	Invention of the sound spectrograph at Bell Laboratories, 1946. For the first time speech was made *visible*. 
1950s	Invention of the pattern playback (a kind of inverse spectrograph) at Haskins Laboratory, 1951. This was used for early experiments in speech perception. Development of the *Motor Theory of Speech Perception* by Liberman and Cooper at Haskins Laboratories, 1952. This marked the development of the idea that speech perception is "special" - it operates to deliver a discrete phonological level

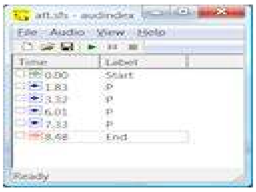
	explanation for the continuous signal. Publication of R. Jakobson, G. Fant and M. Halle's “Preliminaries to Speech Analysis, 1953. This presented the idea of distinctive features as a way to reconcile phonological analysis with acoustic-phonetic form. Denis Fry (1907-1983) and Peter Denes built one of the very first automatic speech recognition systems at UCL, 1958. John R. Firth (1890-1960), working at UCL and at SOAS, introduces prosodic phonology as an alternative to linear-segmental phonemic analysis.
1960s	Early x-ray movies of speech collected at the cineradiographicfacilty of the Wenner-Gren Research Laboratory at Nortull's Hospital, Stockholm, Sweden, 1962.  First British English speech synthesized by rule, John Holmes, Ignatius Mattingly and John Shearme, 1964:

1970s	Gunnar Fant (1919-2009) publishes *Acoustic Theory of Speech Production*, a ground-breaking work on the physics of speech sound production, 1970. Development of the electro-palatograph (EPG), an electronic means to measure degree of tongue-palate contact, 1972. Development of Magnetic Resonance Imaging (MRI), 1974. Adrian Fourcin, working at UCL, describes the Laryngograph, a non-invasive means of measuring vocal fold contact, 1977. Graeme Clark (1935-), working in Melbourne, Australia led a team that developed the first commercial cochlear implant (bionic ear), 1978.
1980s	Dennis Klatt (-1988) develops the KlattTalk system, which became the basis for the DecTalk product,1983: Famously, DECTalk has been used by Stephen Hawking for so long that he is now recognised by its synthetic voice. First automatic speech dictation system developed at IBM by Fred Jelinek, LalitBahl and Robert Mercer among others, 1985. Kai-fu Lee demonstrates Sphinx, the first large-vocabulary, speaker-independent, continuous speech recognition system, 1988.
1990s	Development of Electro-magnetic articulography (EMA), a means to track the motion in real time of small coils attached to the articulators. Publication of the TIMIT corpus of phonetically transcribed speech of 630 American talkers, 1993. This has been used as the basis for much research in

	acoustic-phetics. Jim and Janet Baker launch Dragon Systems *Naturally Speaking* continuous speech dictation system, 1997. Increasing use of functional MRI to study the neural basis for language production and perception, 1997. Ken Stevens (1924-2013) publishes *Acoustic Phonetics*, the most complete account yet of speech acoustics, 1999.
2000s	Discovery of FOXP2, the first gene shown to have language-specific actions, 2001. Microsoft includes free text-to-speech and speech recognition applications for the first time in Windows Vista, 2006.
2010s	Development of real-time MRI that can be used to image the articulators during speech, 2010.

9. Software in Speech, Hearing and Phonetic Sciences

Windows Software

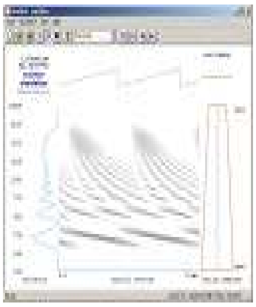


Audindex –Real –time audio indexing tool

AUDINDEX is a Windows PC program for quickly adding an index to an audio file. AudIndex will allow you to replay the audio while recording the times (relative to the start of the audio) when keys are pressed.

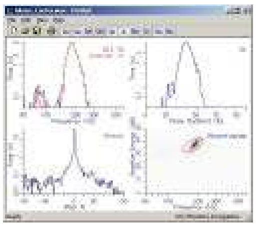
Browse-Simple tool for browsing audio files

BROWSE is a program for browsing audio recordings. With BROWSE you can drag and drop audio files in a range of formats onto the display to see and hear their contents. BROWSE allows you to zoom, scroll and save the audio files to other formats.



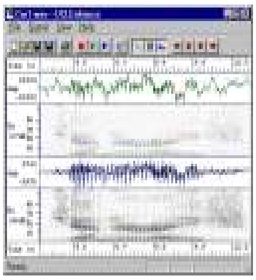
CochSim- Cochlear Simulation teaching tool

CochSim is a dynamic simulation of the time and frequency analysis performed by the ear. Sound signals such as sinewaves, pulse trains, saw tooth waves and vowels can be fed into an auditory filterbank and the output monitored in a moving animated display. The program shows the vibration of the oval window and the basilar membrane, the haircell activity against filter frequency and time, and an average excitation pattern across the cochlea.



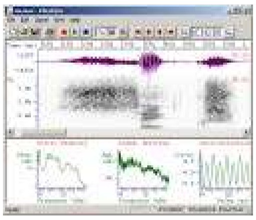
EFxHist-Fundamental Frequency Histograms

EFxHist is a program to analyse joint speech and laryngograph recordings. EFxHist locates each pitch period in the laryngograph waveform and produces statistical analyses including distributions of fundamental frequency, closed-quotient, jitter and shimmer.



Enhance-Speech Signal Enhancement

UCL Enhance is a program for the enhancement of the intelligibility of speech recordings. The program incorporates a number of standard techniques for manipulating the overall amplitude of the signal and for the removal of steady-state additive noise. However, the program is unusual in its ability to detect and enhance selective regions of the speech signal based on their phonetic properties.

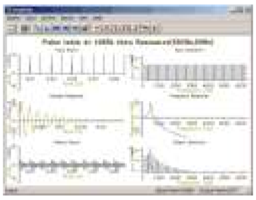


ESECTION-Speech Signal Cross-Sections

ESection is a free program for calculating and displaying spectral and other related analyses of sections of a speech signal. It can be used to demonstrate the different spectral properties of elements of speech. It can also calculate an LPC spectrum, autocorrelation and cepstrum analyses, and can display the signal as a waveform or as a spectrogram. It automatically finds formant and fundamental frequency values.

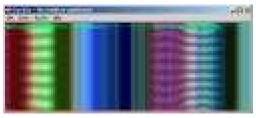
ESYNTH-Harmonics Analysis/Synthesis Teaching Tool

ESynth is a free program designed to explain the harmonic analysis and synthesis of signals. With ESynth you can create signals by adding together individual sinusoidal waveforms (sinewaves) and study the resulting waveform and spectrum. You can also perform an analysis of an input waveform, to see how a given sound can be represented in terms of a sum of sinewaves.



ESYSTEM-Signals and Systems Teaching Tool

ESYSTEM is a free program for experimenting with signals and systems. With ESYSTEM you can see the effect of simple systems on a range of simple signals. You can generate simple signals such as sinewaves, pulses, pulse trains, sawtooth and noise; you can pass them though systems such as an amplifier, a resonator, a low-pass, high-pass or band-pass filter, or a vocal tract model. You can observe the effect on the input and output waveforms and the input and output spectra. Now includes extensive tutorial!



FAROSON-The Auditory Loghthouse

FAROSON is a free program for displaying a real-time scrolling coloured pattern from speech sounds. The aim is to construct a pattern that reflects our subjective sensations of loudness, pitch and timbre. The program may be useful in teaching about the nature of sound sensation.

Hearloss –Hearing Impairment Demonstrator

HearLoss is a free interactive Windows PC program for demonstrating to normally hearing people the effects of hearing loss. With HearLoss you can replay speech, music and noise under a variety of loudness, filtering and masking conditions typical of hearing impairments. Best of all you can interactively change the settings and demonstrate their consequences.

PROREC-Speech Prompt & Record System

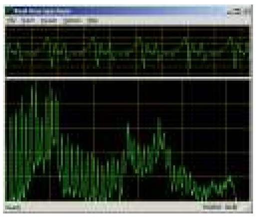
Prorec is a free Windows PC program for prompting speakers with lists of words or sentences and recording their speech back to disk. ProRec is suited to linguistics field workers wanting to use their laptop as an audio recording system.

RTGRAM-Real-time Spectrographic Display

RTGRAM is a free program for displaying a real-time scrolling speech spectrogram on Windows computers. RTGRAM is optimised for speech signals and has options for different sampling rates, analysis bandwidths, temporal resolution and colour maps.

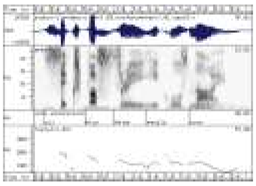
RTPITCH-Real-time Pitch Track Display

RTPITCH is a free program for displaying a real-time fundamental frequency (F0) track for a speech signal on Windows computers. With RTPITCH you can monitor the waveform and pitch of speech sounds being played into the computer's microphone or line input ports. RTPITCH also has an option to display two pitch tracks and switch between them for intonation practice.



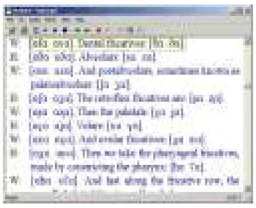
RTSPECT-Real-Time Waveform and Spectrum Display

RTSPECT is a free program for displaying a real-time waveform and spectrum display of an audio signal on Windows computers. With RTSPECT you can monitor the waveform and spectral shape of sounds being played into the computer's microphone or line input ports. RTSPECT can display one or two-channel audio signals.



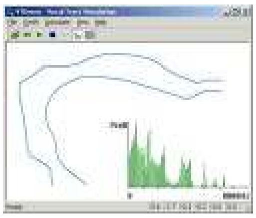
SFS-Speech Filing System Tools for Speech Research

The Speech Filing System (SFS) is a free software suite running on Windows and Unix systems for the storage and analysis of speech data. It provides a comprehensive environment for speech signal processing, synthesis and recognition.



VOISCRIPT-Scripted Audio Playback System

VoiScript is a program to synchronize display of a script with replay of an audio recording. VoiScript allows you to move around a scripted recording simply by clicking on the script. Since VoiScript also allows you to record your own voice, it is useful in many educational situations. VoiScript comes with instructions about how to author your own audio applications.



VTDEMO –Articulatory Synthesis Teaching Tool

VTDemo is a free program designed to demonstrate how the vocal tract generates different sounds. With VTDemo you can play larynx buzz through a vocal tract that can be manipulated in shape and hear the results.

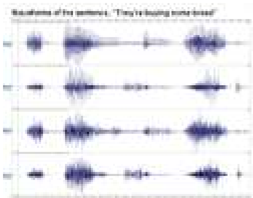
WASP- Waveforms Annotations Spectrograms and Pitch

WASP is a free program for the recording, display and analysis of speech on personal computers. With WASP you can record and replay speech signals, save them and reload them from disk, edit annotations, and display spectrograms and a fundamental frequency track.

WTUTOR – Web Tutorial Authoring System

WTUTOR is a system for the authoring and delivery of interactive teaching and learning used. WTutor allows you to create tutorials containing multimedia and quiz questions using standard word-processing software in combination with freely available tools. No additional software is required at the student end.

<u>Matlab Software</u>



Spectral Rotation

A suite of Matlab scripts that perform spectral rotation, of the kind described by Blesser B (1972) Speech perception under conditions of spectral transformation: I. Phonetic characteristics. *J Sp& Hear Res*, 15: 5-41. For. The software is available via anonymous FTP free of charge but with no guarantees and copyright retained by UCL.

Speech Analysis Tools

1. Anvil-Video annotation tool
2. Audacity-Spectrogram visualization and audio recording and editing
3. CeedVocal-Speech recognition for the IPhone
4. Colea- Speech analysis freeware
5. Computerized Speech Lab (CSL) –from Kay pentax
6. CSLU Toolkit- Mulkit- Multiple tools for speech analysis, recognition, generation , and display
7. Janus-Speech translation system
8. Julius- open-source large vocabulary continuous speech decoder
9. Let's Go-info on the mobile speech analysis project from CMU
10. Open Mind Speech-open source crowd sourcing project for speech analysis project.
11. Phonology Assistant-Tool using IPA (International Phonetic Alphabet) characters to index and display data.
12. the Praat- Open source speech analysis program
13. SFS/RTSPECT Version 2.4- Windows tool for real-time waveforms & spectra
14. Speech analyzer-From SIL International
15. Talk Bank CMU- Speech technology project
16. The CMU Sphinx Group-Open source speech recognition engines
17. Video Phonetics database and program –From KayPentax
18. Wavesurfer- Open source tool for sound visualization and manipulation
19. WinCECIL-Tool for viewing speech recording,automatic pitch contours ,and spectrograms
20. WinPitch LTL-Voice processing software

21. Winsnoori

Language Learning Software and Projects

1. Better accent Tutor-Speech analysis
2. CAMMIA (A Conversational Agent for Multilingual Mobile Information Access) –from Language Technologies Institute ,CMU
3. CandleTalk-Conversation tool for English
4. Fluency – foreign language pronunciation training CMU
5. Intelligent dialog overcomes speech technology limitations: THE SENECa example
6. MyET-MyCT3- Uses speech analysis software for tutoring English or Chinese
7. Review of Tell Me More Chinese- From Calico Review
8. Sakhr Software-Mobile voice applications
9. Speak GoodEnglish- English pronunciation program
10. SPICE (Speech Processing Interactive Creation and Evaluation)-Speech processing models
11. Tactical Language and Culture-US Army language software
12. TransTac-Spoken language communication and translation system for tactical use

10.Glossary

Accent	the unique speech patterns of a person or group
Affricate	a speech sound (consonant) that contains a stop followed by an immediate fricative, as in the ch /ʧ/ in "chair"
air flow/airstream	the flow or passage of air out of the mouth

Alveolar	sound formed by touching the tip of the tongue to the upper alveolar ridge, as in /t/ or /d/
alveolar ridge	the bony region at the roof and bottom of the mouth behind the front teeth; contains the tooth sockets
Approximants	consonants with a partial obstruction of airflow, as in /w/ and /r/
Articulation	the act of making speech sounds
Aspiration	a small "explosion" of air when you make a sound
Auditory	hearing (not seeing)
Bilabial	consonant sounds formed using both lips, as in /p/ or /b/
close vowel (sometimes called "high" vowel)	a vowel sound that is pronounced with the tongue close to the roof of the mouth (but not close enough to constrict the air and make a consonant), as in /i:/ in the word "free"
Consonant	a speech sound made when there is complete or partial obstruction of air in the mouth, as in /v/, /h/, /d/ (compare vowel)
Clusters	blended sounds put together to make a single sound
Curl	a position of the tongue where the tongue is shaped in a curve, not flat
Dental	a consonant sound made when the tongue touches the upper teeth, as in /t/ and /n/
Dialect	unique vocabulary, pronunciation and usage that is typical of a certain group of people
Diphthong	a sound made by the combination of two vowel sounds in a single syllable, as in "boy", "loud" or "wide", where the sound starts as

	one vowel and moves towards another vowel
Flatten	a positioning of the tongue where the tongue is flat not round
Fricative	a speech sound (consonant) in which air is forced to pass through a small opening and creates friction, as in /f/ and /v/
glide/slide	moving the tongue while saying a word
glottal stop	the sound that is made when the vocal folds are closed very briefly; as in the middle of the word "uh-oh" (common in American English)
Gum	the tissue around the base of the teeth
hard palate	hard part of the roof of the mouth
Intonation	change in pitch of a sentence, up and down; the music or rhythm of speech
Labiodentals	sounds that are made with the lower lip and upper teeth, as in /f/ and /v/
Larynx	the hollow, muscular organ in the throat that holds the vocal chords; the voice box
Lateral	a speech sound that is made by touching the tongue to the middle of the alveolar ridge, allowing air to pass on both sides
lengthen sound	make the duration of the sound longer
Linking	the joining of words when speaking, as in "Ca-nI-ha-va-bi-to-fegg?" (Can I have a bit of egg?)
lips spread	lips are open slightly and pulled back

Lower	bottom of mouth
minimal pairs	two words that differ only in terms of one sound, as in "cat and bat" OR "fine and vine"
Monophthong	a single vowel sound that does not change in auditory quality; also called a "pure vowel"
nasal consonants	consonant sounds made by pushing air through the nose, as in /m/, /n/ and /ŋ/
non-pulmonic	when the air comes from a source other than the lungs
Obstruction	a blockage of air flow
open vowel (also called "low" vowel)	a vowel that is produced with the tongue far down from the roof of the mouth, as in the /a:/ sound in "far"
Palatal	a sound that is made when the tongue is near or touching the roof of the mouth
Palate	the roof of the mouth
Phoneme	an individual speech sound
phonetic alphabet	an alphabet that represents the sounds of speech
phonetic transcription	a form of notation that uses symbols to identify the individual sounds (phonemes) in a word
Plosive	a consonant sound produced when there is a complete obstruction of air followed by its sudden release, as in the /p/ of "pot"
Pitch	amount of highness or lowness of a sound or speech
Postalveolar	a consonant sound made with the tip of the tongue slightly back

	from the alveolar ridge, as in /ʃ/ in "shut"
pressed lips	top and bottom lips touching
protruded lips	rounded lips, pushed out
Pulmonic	a sound that is made using the airstream directly from the lungs
Raised	higher than the neutral position
Reduction	the natural shortening of sounds when speaking (e.g. "going to" reduced to "gonna")
Rhotic	a variety or dialect of English in which "r" is pronounced before a consonant (as in "hard") and at the end of words (as in "car"); Midwestern American English, for example, is "rhotic"
Roof	the inside top part of the mouth
rounded lips	lips formed into the shape of a circle
rounded vowel	a vowel made with rounded lips
Sentence stress	the placement of emphasis on specific words within a sentence or phrase
shorten sound	make the duration of a sound shorter
soft palate	soft part of the roof of the mouth
Sonorant	sounds that are made when air is impeded only slightly, as in /m/, /n/
stop (stop consonant)	a consonant sound that is produced when the airflow is (temporarily) stopped entirely by the lips or tongue, as in /p/
Syllable	a single unit of sound that creates one beat in a word; the word

	"coffee" has two syllables (cof-fee)
syllable nucleus	the central part of a syllable, usually a vowel
Tap	touch quickly
tone	the emotion that is conveyed through the sound of speech (e.g. anger or sadness)
tongue	muscular tissue in the mouth used for tasting and articulating
tooth ridge	the hard area directly behind the top front teeth
Trill	a vibrating sound made with a flapping tongue, as in the rolled "r" sound made when people roll their r's
upper	top of mouth
velar	of a sound that is made with the back of the tongue near the soft palate, as in the the /ŋ/ in "sing"
velum	a soft membrane on the roof of the mouth (also called "soft palate")
vocal chords (AmE cords)	two muscles inside the larynx that vibrate and create the voice
vocal tract	the entire apparatus that produces voice, starting in the lungs and ending at the lips and nostrils (openings of the mouth and nose)
voiced	of a sound made with the vocal chords (voice box) vibrating
voiceless/unvoiced	of a sound made without the vocal chords (voice box) vibrating
vowel	a speech sound made when air is free to pass through the mouth with little or no obstruction, as in sounds made with the letters a, e, i, o, u, and sometimes y as compared to consonants.

vowel backness	position of the tongue in relation to the back of the mouth when making a vowel sound (positions include front, near-front, center, near-back, back)
vowel height	distance between the tongue and the roof of the mouth when pronouncing a vowel sound (IPA has 7 heights: close (highest), near-close, mid-close, mid, open-mid, near-open, open (lowest)
Word stress	the placement of emphasis within a word that has more than one syllable

https://www.englishclub.com/pronunciation/terms.htm#vowel

11. What Materials Should a Language Laboratory Contain?

Introduction

Laboratoire de Phonétique et de Phonologie (LPP)[i] is specialized in education and research in experimental phonetics and in phonology. LPP inherits these missions from the first phonetics laboratory created in Paris for Pierre-Jean Rousselot (1846-1924), who is often considered as the founder of experimental phonetics. Currently, LPP is a multi-disciplinary research team, which includes medical doctors and speech therapists, along with phoneticians and phonologists.

The long-term goal of LPP is to achieve an integrated model of phonetics and phonology. It also works on applications to clinical phonetics and language learning. LPP has recently assembled a research platform for investigating the behaviour of each speech organ involved in speech production. The platform is elaborated by a joint effort between engineers, phoneticians and clinicians. It provides tools to investigate various phenomena of coordination and compensation across speech organs that are observed in the production of speech by normal or pathological speakers, foreign language learners, and singers. Non-invasive instrumentation techniques are located at LPP to be open to a broader public of researchers. Mildly invasive techniques are applied at the Georges Pompidou European Hospital (otolaryngology department)for diagnostic purposes and for research by the members of LPP.

Table 1 lists the instrumentation techniques that are being used at LPP to examine the speech organs: larynx, tongue, velum, lips, jaw, and face. Electroglottography (EGG) and photoelectro-glottography (PGG) are used to evaluate the degree of contact between the two vocal folds, and the degree of opening of the glottis, respectively. Endoscopy is used for direct observation of the laryngeal components. Ultrasonography (USG) allows us to visualize tongue shapes, and static or dynamic palatography methods to acquire data on tongue-palate contact. Nasal acoustic and vibratory signals are recorded using a nasal microphone and a contact transducer. Nasal and oral airflow rates are measured using an aerodynamic sensing system (EVA2^{TM}).Visual estimation of velar height is conducted by endoscopy. Photonasography (PNG) evaluates the opening of the velopharyngeal port. Video and motion capture systems are used to study movements of the lips, jaw and face, in particular to investigate visual cues to labial articulation. Several instruments are used simultaneously whenever possible, depending on the instruments.

Organs	Instrumentation Techniques	Parameters to beMeasured
Larynx	Electroglottography (EGG)	Contact between the two vocal folds
	Photoelectric glottography (PGG)	Degree of opening of the glottis
	Rigid and flexible endoscope	Visualization of laryngeal components
Tongue	Ultrasonography (USG)	Visualization of contour of the tongue
	Static and dynamic palatography	Degree of linguo-palatal contact
Velum	Nasal acoustic microphone, and Piezoelectric transducer	Nasal acoustic signal, and intensity of skin surface vibrations
	Aerodynamic sensors, EVA2™ [1]	Nasal and oral airflow, and air pressure
	Videofiberscopy	Height of the nasal surface of the velum
	Photonasography (PNG)	Opening of the velopharyngeal port
Lips and face	Video	Visualization
	Motion capture system	Marker tracking in 3D dimensions

Table 1: Target Organs, Instrumentation Techniques, and Parameters to be Measured

[1] http://www.sqlab.fr/evaSensFR.htm

Angélique Amelot. Institute of Linguistics, Phonetics and Phonology. University Sorbonne Nouvelle Paris 3

[i] Phonetics and Phonology Laboratory in Paris (http://lpp.univ-paris3.fr/).LPP is associated with University Sorbonne Nouvelle (University of Paris III) and Centre National de la Recherche Scientifique (National Center for Scientific Research), France.

Summary in French and Arabic

Résumé

Ce travail vise à introduire une technologie récente et à explorer les moyens par lesquels elle peut être utilisée afind'améliorer l'enseignement et l'apprentissage de la prononciation. Cette technologie est représentée sous forme de logiciel d'analyse et de traitement de la parole qui fournit aux étudiants, en temps réel, des données visuelles et auditives, et leur procure une évaluation immédiate et authentique deleur production en utilisant le logiciel Praat d'analyse phonétique. L'hypothèse émise estque l'introduction de cette technologieet son utilisation dans la formation des étudiants,permettrait une amélioration considérable de leur production orale. La recherche est basée sur une étude pilote menée à l'Université de Sorbonne. Elle consistait à déterminer expérimentalement les gains en performance des étudiants ayant utilisé cette technologie. Pour tester cette hypothèse, nous avons opté pour l'utilisation de différents outils de collecte de données, à savoir le questionnaire, l'enregistrement de la production des étudiants et l'entrevue des enseignants. Les premiers résultats ont montré que l'approche a permis d'améliorer et d'affiner la prononciation des étudiants ainsi que leur aptitude à détecter et corriger leurs erreurs de manière autonome. En ayant la disponibilité instantanée, visuelle et automatique affichée dans des spectrogrammes, les étudiants ont été en mesure effectivement de détecter aisément les erreurs qu'ils auraient pune pas remarquer par l'écoute seule.Le travail démontre également que les aptitudes acquises grâce à l'utilisation de Praat se concrétisent par une amélioration concrète de la production orale, puisque les étudiants ayant bénéficié de cette approcheont étécapables de produire les énoncés plus clairement et plus naturellement dans des exercices subséquents de lecture. Les résultats finaux de cette étude ont montré comment les enseignants et les étudiants peuvent bénéficier grandement de ce logiciel pédagogique gratuit et facilement accessible. Il permet notamment aux étudiants de progresser à leur rythme tout en stimulant leur confiance en soi lorsqu'ils s'expriment oralement.

ملخص

يهدف هذا العمل إلى إدخالتكنولوجيةحديثة واستكشاف الطرق التي يمكن استخدامها لتعزيز تعليم النطق و تحسينه.وتتكون هذه التكنولوجيا من برنامج تحليل الكلام الذي يوفر للطلاب كل من البيانات البصرية والصوتية ويعطيهم تقييماآنياعلى ذلك باستعمال البرنامج الحاسوبي برات للتحليل الصوتي. ويفترض في هذا البحث أنه إذا تم إدماج الطلاب وتدريبهم باستخدام هذه التكنولوجيا، فإن التدريب الذي توفره هذه الطريقة من شأنه أن يحسن إلى حد كبير تحصيلهم الصوتي. ويعتمد البحث علىدراسة تجريبية أجريت في جامعة السوربون والتي تتمثل في تدريس النطق عن طريق اختبار فعالية استخدام هذه النظم المرئية القائمة على الحاسوب لمساعدة الطلاب على تحسين النطق. لاختبار هذه الفرضية، اخترنا استخدام أدوات مختلفة لجمع البيانات وهي الاستبيان، وتسجيل الطلاب و مقابلة مع بعض الزملاء الأساتذة.وتظهر النتائج الأولية أن البرنامج برات بوسعهأن يمكن الطلاب من صقل نطقهموعلاج أخطاءهم. فمن خلال وجود تقييمات فورية بصرية وتلقائيةمعروضة على الشاشة، استطاعالطلاب رصد أخطاء النطق التي لم يكن بوسعهم رصدها لو اكتفوا فقط بالاستماع فقط. ويبينالعمل أيضا أن التدريب بواسطة برات يمكن الطلاب منالتحدثبشكل أكثر وضوحا وبتلقائية خاصة. وتظهر النتائج النهائية لهذه الدراسة كيف تساعدهذه البرامج التعليمية المجانية والمتاحة بسهولة للجميعالأساتذة والطلاب على إدراكأخطاءهم و القدرة على تصحيحها بصفة عفويةوعلى اكتساب الثقة أثناء التعبير الشفهي.