

**Acoustic Analysis of English Pure Vowels in Clear  
and Conversational Speech : An Experimental  
Study at the University of Basra**

Asst. Prof. Hamid Majid Al-Hamadi (Ph.D.)      Mr. Firas F.Ali (M.A.)  
English Dept., College of Arts,                                  College of Education  
University of Basra

**Abstract**

The present study aims at examining the acoustic properties of English pure vowels produced by native and non-native speakers in clear and conversational speech (henceforth CLR and CNV speech respectively). This study concentrates on the most important aspects of acoustic phonetic research, acoustic analysis, vowel intelligibility, sex-related differences, as well as comparing clear to conversational speech. It is hypothesized that the acoustic properties of English pure vowels in conversational speech are different from those in clear speech. Moreover, the strategies employed by non-native speakers, as far as exhibiting the acoustic properties of the vowels involved in both styles, are different from those of native speakers. The researchers have followed the precise procedures presented by some investigators, particularly those procedures followed by Ferguson and Port (2002) as far as the work on speech styles and vowel intelligibility is concerned. Two recording sessions are conducted for both groups of talkers and all recordings were saved and analysed using Praat, which is a free scientific, authentic, and academic software for doing phonetics by computer. The results revealed that all talkers produced the vowels with longer durations, higher values of  $F_0$ , greater values of pitch range, and larger vowel space areas in clear than in conversational speech. Most female talkers produced higher values than males for the four acoustic measures. The findings of the experiment have shown that native and non-native speakers differ in the strategies they employ, and that clear speech is more intelligible than conversational speech.

---

<sup>o</sup>This paper has been extracted from an M.A.dissertation submitted by Mr.Firas F.Ali, under the supervision of Dr.Hamid Al-Hamadi to the Council of the College of Arts, University of Basra.

## تحليل فيزيائي لأصوات العلة الإنجليزية البسيطة في الكلام المتأني وكلام المحادثة: دراسة تجريبية في جامعة البصرة

### الملخص:

تهدف هذه الدراسة إلى تحليل الخصائص الفيزيائية لأصوات العلة الإنجليزية البسيطة كما ينطقها متكلمو اللغة الإنجليزية الأم ومتعلموها الأجانب في أسلوب الكلام المتأني وكلام المحادثة. كما تركز على أهم الجوانب في أبحاث علم الصوت الفيزيائي، وتحليل الاصوات اللغوية تحليلاً فيزيائياً ووضوح الكلام وأصوات العلة والإختلافات المتعلقة بالجنس، مع مقارنة الخصائص بين الكلام المتأني وكلام المحادثة، حيث يفترض الباحثان أن الخصائص الفيزيائية لأصوات العلة الإنجليزية البسيطة في كلام المحادثة تختلف عن تلك الخصائص في الكلام المتأني. فضلاً عن ذلك، فإن الإستراتيجيات التي يتبعها متكلمو اللغة الإنجليزية الأجانب في عملية نطق أصوات العلة في أسلوب الكلام، تختلف عن الإستراتيجيات التي يتبعها متكلمو اللغة الإنجليزية الأم.

لقد أثبتت النتائج بأن كل المتكلمين قد نطقوا أصوات العلة الطويلة والقصيرة بمدة زمنية أطول وتردد أساسي أعلى ونطاق ترددي أكبر ومساحة صوت علة أوسع في الكلام المتأني من كلام المحادثة. وقد حصلت الإناث على قيم أعلى من الذكور فيما يخص الخصائص الفيزيائية الأربع التي تم تحليلها. ومن الجدير بالذكر أيضاً، فإن النتائج أظهرت بأن معظم الخطوات التي إتبعها المتكلمون الأجانب تختلف عن تلك التي يتبعها متكلمو اللغة الإنجليزية الأم في أسلوب الكلام، ولأصوات العلة الطويلة والقصيرة.

وأثبتت نتائج التحليل والتجربة أن الكلام المتأني أوضح وهو مفهوم بدرجة أكبر بكثير من كلام المحادثة العادي، واختتمت الدراسة ببعض التوصيات والمقترحات للدراسات القادمة.

## **1. Introduction**

To the best of the researchers' knowledge, most of the acoustic work done by researchers at the Iraqi university levels, according to the available literature, lacks some instrumental studies, specifically in examining the acoustic properties of speech sounds. As a matter of fact, the acoustic studies, in Iraq, are very limited in number and very restricted in scope. Non-native speakers of English often come across some difficulties in producing speech sounds, particularly vowels. Accordingly, determining the acoustic characteristics of English pure vowels and comparing the results with the norms produced by English native speakers, will help a great deal in understanding vowel sounds, improving vowel intelligibility, specifying sex differences, and recognizing the similarities and differences between CLR and CNV speech.

The present study follows all of the precise procedures presented by Ferguson and Port (2002) in examining the acoustic properties of pure vowels. It also follows those procedures in determining the main differences between CLR and CNV speech as far as the production of these vowels is concerned.

The goals of the present study are:

1. performing an acoustic study of English pure vowels produced by selected native and non-native speakers, and introducing the main procedures to conduct an experimental acoustic phonetic study.
2. presenting a new tool and speech analysis software, new computer and digitized procedures, and a new recording technique.
3. examining the acoustic properties of English pure vowels in both CLR and CNV speech produced by both native and non-native speakers.

4. exploring and investigating sex-related differences in both groups.
5. highlighting the most important aspects of acoustic phonetic research, acoustic analysis, vowel production and intelligibility, as well as comparing CLR with CNV speech.

This study hypothesizes that the acoustic properties of English pure vowels in CNV speech are different from those in CLR speech, and the strategies employed by non-native speakers in applying the acoustic properties of the vowels involved in both speaking styles from those employed by native speakers.

The data of the study has been obtained from an instrumental analysis of English pure vowels which complement the acoustic descriptions of these vowels as produced by native and non-native speakers in CNV and CLR speech. This study helps build a more comprehensive knowledge of acoustic-phonetic analysis and recognize the acoustic differences to be encountered in the two speech styles by both sexes. The output of the present study is of a great importance to the teaching staff members and the EFL learners at the Iraqi departments of English to improve their intelligibility and to exhibit the appropriate characteristics of both speaking styles when communicating with native speakers. The value of the present study lies in the statistics and results of the conducted experiment. This study may also be regarded as a guidance or a tutorial that helps develop a good knowledge and acquaintance of the acoustic subfield of phonetics.

## 2. Theoretical Background

### 1.1 Acoustic Phonetics: Basic Remarks

The term 'acoustic', according to Fry, is normally used in relation to physics, but one general sense of the word itself is

related to the sense of hearing (Fry 3)<sup>•</sup>. Since our principal interest is with the physics of speech sounds as related to the language system, then the major purpose is to demonstrate the links between the physical input and its sensations to the ear which constitute the stimulus, and also how these sensations are further organized parallel with the language system.

A good number of linguists and phoneticians, such as O'Connor (1973), Ladefoged (2000), Roach (2000), Crystal (2003), Johnson (2003), Birjandi and Nodoushan (2005), Finch (2005), Stranzy (2005), Bussmann (2006), and Malmkjar (2006), have defined acoustic phonetics as that technical area of linguistics and branch of phonetics that studies the physical parameters of speech sounds as transmitted between mouth and ear, and as represented in variations of air pressure that spread in waves through the air and can be heard, recorded, visualized, and measured.

Sources of sound are those which cause variation in air pressure. The changes in the variations of air pressure come as a result of small frequent movements of the air particles. The source of sound which we are concerned with is the human voice. Here, variations in air pressure are caused mostly by the rapid opening and closing of the vocal folds. Consequently, speech consists of fluctuations in air pressure due to physical disturbances of air particles caused by the airflow out of the lungs. This airflow causes the air particles to accumulate together and move apart (oscillate), creating increases and decreases, respectively, in air pressure. The resulting sound wave transmits these transformations in pressure from the speaker to the hearer. As speech sound waves are difficult to describe, it will be helpful to

---

• The 2009 MLA documentation style has been followed in this paper.

represent them as visible shapes with the aid of microphones, computers, and computer softwares. The most fruitful and productive shape will be the one which shows how the air pressure at a given place changes over a particular period of time. Accordingly, sound waves can be described in physical terms such as cycle, period, frequency, and amplitude (Ladefoged, Elements 1-3).

It is very important to make a distinction between the sound itself, that is the stimulus, and the sensation. Frequency is a feature or a quality of the stimulus, while pitch is an attribute of the sensation. Pitch is directly affected by frequency; an increase in the frequency of vibration causes a similar increase in pitch (Fry 11). In an attempt to distinguish between pitch and frequency, Ladefoged (Elements 22) states the following:

We use the word **pitch** when we are referring to that aspect of a sound whereby we can, using only our ears, place it on a scale going from low to high. When we are discussing actual rates of vibration or rates of fluctuations in air pressure, we speak of the **frequency** of the sound.

This means that the frequency of a vibration is a purely physical phenomenon. Therefore, it is the pitch of a sound that we hear and not the number of vibrations per second.

The lowest frequency of the wave is the fundamental frequency ( $F_0$ ) which together with other frequencies constitute the harmonic components of the wave (Clark et al. 224). In other words, the tone which results from successive vibrations of the sound wave in its full length or size is the fundamental frequency. It is the fundamental frequency ( $F_0$ ) which determines the pitch of a sound, i.e., the pitch of a given sound is always the pitch of the fundamental. Other frequencies that result from vibrations above the fundamental are called the harmonics or overtones (formants).

Loudness, intensity, and amplitude are three significant features of sound waves where great fluctuations of air pressure exist.

These great fluctuations or variations of air pressure cause the eardrum to move and are interpreted by the listener as loud sounds. If these fluctuations are much smaller, they cause a softer sound (Ladefoged, Elements 16).

"The extent of the maximum variation in air pressure from normal during the production of any speech sound at any given time," is the amplitude of the sound (Awaness 18). The sound wave is accompanied by certain variations in air pressure. The highest distance these variations travel from zero to the farthest point is the amplitude of the sound wave. As the distance of the air molecules is increased, they strike the eardrum, and thus we perceive a louder sound. If the distance or the peaks of the variations of air pressure decrease, then the sound becomes less loud (Birjandi & Nodoushan 157-158).

Sawusch (2005) as quoted in Pisoni and Remez (7-9) who states that the acoustic speech outcome in humans results from a combination of a source of sound energy modified by a filter determined by the shape or geometry of the vocal tract. This model is referred to as the "source-filter" theory of speech production. When the larynx acts as a source of sound energy, voiced sounds may be produced by a rapid opening and closing of the vocal folds – voiceless sounds are produced with the air passing through without any vibration. The frequency of the pulses of the vocal folds determines the fundamental frequency of the laryngeal source which is the lowest of its correspondent harmonics. After passing through the larynx, the airstream enters the vocal tract. The vocal tract shapes that airstream and the sound that is produced represents a combination of the effect of the

larynx (source) and the effect of the vocal tract (filter). Finally, there is the radiation effect (baffle effect) that exerts the final shaping on the acoustic signal. This radiation takes place as an echo or a response at the walls of the vocal tract or at the lips.

The source-filter model of speech production, accordingly, can be put as follows: "The sounds of speech are the result of a source (voiced or voiceless) that is passed through a filter (the vocal tract)" (7-9).

Ladefoged (A Course 161) states that acoustic phonetics gives a better understanding of how computers analyse and manipulate speech and how speech recognition works, as the best way to analyse speech acoustically is to work from a recording . Doing an articulatory, acoustic, or auditory analysis of the recordings of speech falls under the domain of Digital Speech Processing (DSP). Rabiner & Schafer (1-2) point out that DSP gives a comprehensive overview of the different analytical techniques. It ranges from the basic nature of the speech signal which carries the message information (the acoustic waveform), through a variety of methods of representing speech in a digital form, to applications in voice communication, automatic synthesis and recognition of speech, and acoustic-phonetic research.

The sound waveform, as stated by Ladefoged (A Course 162-163), is a two-dimensional presentation with the horizontal axis representing the time domain and the vertical axis representing the frequency domain. The sound spectrogram is another graphical display for analysing speech signals acoustically. It is defined by Holmes and Holmes (286) as a graphical display of sound in which time is on the horizontal axis and frequency is on the vertical axis. Intensity is shown by a grey-scale representation (the darker the display, the higher the intensity) or by a colour display (the brighter the colour, the higher the intensity). It is by the use of



these graphical displays where the acoustic properties of speech sounds can easily be observed and described.

### **1.2 Clear and Conversational Speech**

This study explores the acoustic phonetic properties of English pure vowels in two speaking styles; namely, clear and conversational speech.

CLR speech is just one in a large class of speaking styles adopted by talkers as the situation demands. This large class of speaking styles may include: infant-directed speech, speech produced in noise, shouted speech, speech produced during simultaneous communication, speech produced while under stress or a cognitive workload, speech produced to a hearing-impaired person, talkers instructed to produce highly enunciated speech, and speech produced during human-computer error resolution (Uchanski, 1996 as cited in Pisoni & Remez (208)).

CNV speech, on the other hand, refers to that type of speech produced under casual or typical environments especially when no particular speaking effort or instruction is made. CNV speech is best described in relation to CLR speech. In other words, CNV speech is usually described *relative* to the same samples of CLR speech in order to distinguish the main differences and the most important characteristics of each speaking style (Pisoni and Remez 208).

Rather than training an individual talker to produce speech in a clear or a conversational style, the primary focus of the present study is the main acoustic and intelligibility parametric changes in the production of English pure vowels when a speaker shifts from a conversational to a clear speaking style, i.e., intra-talker differences. Inter-talker differences will also be examined in

relation to the description of some important acoustic and physical measures.

### 1.3 Speech Intelligibility

There is no plain and direct answer to the question 'What is speech intelligibility?' Ball et al. (569) state that speech intelligibility counts on speaker and listener characteristics, speaking and listening conditions, and a multitude of other factors. Goldberg and Riek (136) note that speech intelligibility refers to the listener's proper identification of the message content that is spoken. Others, as stated by Jenkins (70), relate intelligibility to 'comprehensibility', 'interpretability', 'recognition', or 'identification' and to 'understanding', i.e., absorbing the expressive content of the message, while the remainder associate 'intelligibility' with the effect of recognizing the linguistic forms. Ball et al. (269) define speech intelligibility as follows: "[A] relative measure of the degree to which a speaker's speech signal is understood, the relativity depending at a minimum on the identities of speaker and listener, what is spoken and where it is spoken."

It has been well documented by different studies such as Schum (1996), Ferguson (2004), Smiljanic and Bradlow (2005), Ferguson and Port (2007) and others, that the most important acoustic-phonetic factors which should be examined because they affect speech intelligibility in CLR and CNV speech are: vowel duration, fundamental frequency, pitch range, and vowel space area. All of these factors affect the intelligibility of speech and prove that CLR speech is more intelligible than CNV speech if they acquire higher values in the former style. Intelligibility is usually measured when the clarity of consonants and vowels is the only prominent and contributing variable. Clark et al. (305) state that "vowels, the most intelligible components of syllables, have

received more attention in later and more phonetically oriented studies." This has a close correspondence and association with the present study where the main objective is to build up a way of describing English pure vowels acoustically in two speaking styles.

### **1.4 Acoustic Properties of Vowel Sounds**

Clark et al. (289-191) state that the production of vowels is the result of an action composed of acoustic and articulatory mechanisms. The larynx functions as a source generator of the sound wave, and the vocal tract performs the vital role of an acoustic filter of the source sound. Vowel articulation is the setup for the articulatory organs to determine vocal tract shape for each vowel. The main articulatory factors that influence the production of vowels and eventually their acoustic properties are the opening and closing of the jaw, tongue body movement and tongue body shape, the action of the soft palate (lowering and rising), lip rounding which determines the shape of the vocal tract towards the open end, and lowering of the larynx as well as the vibration of the vocal folds. All of these factors have, in one way or another, their physical impacts on the vowel produced. They increase and decrease the harmonics or resonances of the fundamental frequency (formants) which differentiate one vowel sound from another.

According to Ladefoged (A Course 170-173) and Ladefoged (Vowels & Consonants 31-35), the most important properties to be recognized in describing the acoustic structure of vowel sounds are quality, frequency, intensity, formants, duration and plot. Quality is what distinguishes one vowel from another. It is based on the variations within each cycle of the sound wave. In other words, it depends on the overtone pitches 'formants' that give the vowel its identifiable quality. The formants of a vowel sound are estimated usually by measuring the first two or three characteristic

frequencies that we perceive as overtone pitches. The first formant is the lowest in frequency than the other two. These formants are formed because of varying the shape of the vocal tract by different articulatory processes. Apparently, the formant frequencies (regions) of the vowels as well as the intensity that we comprehend as loudness are shown in the spectrogram by darkness. Duration can also be measured for different vowels by counting the time consumed in seconds or milliseconds. By defining the first and second formant values and sometimes the third, one can draw the vowel plot area which represents the vocal tract and the occurrence of vowel sounds, and facilitates the description of vowels of different languages.

## 2 Experimentation

### 2.2 The Talkers

The experiment took place between April and June 2010, and lasted approximately for two months and a half. It consisted of two groups of talkers who served as participants in this experiment. The first group consisted of 20 Iraqi adult non-native speakers of English (10 females and 10 males). The second group consisted of only two adult native speakers of BBC English (1 female and 1 male) whose speech analysis results are considered a norm to compare to.

The non-native speakers were all teaching-staff members from the Departments of English, College of Education and College of Arts, Department of Translation, College of Arts, and two English language teachers from the College of Historical Studies, all at the University of Basra. The two native speakers were chosen from Basra International Airport by sending a request letter to the airport from the University of Basra/College of Arts. They were academically highly educated and were born and raised in London and speak the BBC accent which was checked by the researchers.

All talkers who participated in the experiment were free of any known voice and speech anomalies or any neurologic or muscular condition that would be expected to affect their ability to produce speech.

The non-native participants do not represent a random sampling of talkers. All individuals were selected based on their language experience and their pronunciation skills and clarity as instructors of English. Furthermore, all talkers were not aware of the details of the study and the purpose of the recordings when they were selected for the experiment. None of the talkers was remunerated, and all of them participated according to their willingness and readiness to help.

### **2.3 Speech Materials and Stimuli**

Two types of materials were used in this experiment: Material A and Material B.

Material A was prepared for the CNV speech recording session. It consisted of the twelve BBC English pure vowels carried by 12 target words embedded in a dialogue. The dialogue was presented in two different sensory modalities, audio and visual. To put it differently, Material A consisted of a recorded dialogue taken from O'Connor and Fletcher (54) to be played to the talkers in the experiment to listen to it. The dialogue was also written on paper so the talker could see it at the same time before starting the conversation.

Material B was prepared for the CLR speech recording session. It consisted of a printed list which comprised the twelve target words; the carriers of the twelve BBC English pure vowels. Some distracter vowels were also included to make the talkers feel more comfortable and unaware of the vowels to be analysed.

It is noteworthy to state that the researchers, by their choice of the dialogue with *easy* and *familiar* target words, have made sure that almost all talkers would produce them properly. This is also to have good recordings to examine and analyse.

### 2.3. Procedure

Two recording sessions were conducted on both groups of talkers, i.e., the native and non-native speakers. For all individual talkers, the CNV speech recording session was conducted prior to the CLR speech recording session.\*\*

In the first recording session, the researchers invited two native speakers of BBC English (1 female and 1 male) to come and talk with them individually for about 10 to 15 minutes each, as they are BSc. holders from University College, London. Time was distributed equally among all the participants of the experiment. Each one of the talkers first listened to an original recording of a dialogue spoken by native speakers, then they were requested to speak in a typical conversational manner, as if talking to their family members or friends. The instruction set given to the talkers in CNV speech was as follows: "**It is important that your speech be as much like your normal conversational style as possible.**" As mentioned before, the dialogue carried 12 target words that the researchers wanted the talkers to produce without their awareness. Therefore, the researchers prepared a list of

---

°°The researchers have made sure to conduct the experiment with its sessions by following the accurate and academic procedures being applied by well known scholars like Hillenbrand(1995), Ferguson and Port (2002), and Bradlow et al.(2005),

questions to use in the conversation to naturally evoke these target words.

Then, after listening to the recordings, the researchers asked the same talkers back for another recording session. Talkers were asked to read a list of target words selected from the conversation, and they were requested to speak clearly and carefully and to do whatever they felt necessary in order to be better understood. The instruction set given to the talkers in CLR speech was as follows: "**It is important that your speech be as clear and careful as possible.**" The prepared list consisted of 12 words, and the talkers were asked to produce each word three times to select the best recording, later on, for the analysis technique. All instructions in both sessions were given in English.

For the non-native speakers who were 20 in number, the same procedure was applied in both CNV and CLR speech. In CNV speech, talkers were instructed to speak in a normal conversational style as if talking to native English friends, classmates, etc. In CLR speech, talkers were instructed to do their best to be clear and better understood.

All speech was recorded using a high quality Sony MP3 IC Recorder which is normally used by journalists, and then saved on a computer (Section 3.5). For the native speakers, their CNV and CLR speech were recorded in a very quiet room in Basra International Airport. For the non-native speakers, all recordings were done in different places: the language lab of the Department of English of the College of Education, the language lab of the Department of Translation of the College of Arts, and in some quiet rooms where access to the labs was not available at that time. Each talker was seated approximately 10 inches from the recording device. The output

was saved in the recorder and then was fed to and saved in a laptop computer in an MP3 format.

After conducting all the required recording sessions, each target word in both speaking styles was saved manually in a single file to be acoustically analysed using Praat (Version 5.1.43). Four acoustic properties were selected to assess the primary characteristics of the vowels involved. The first property is vowel duration: measuring the duration of a vowel after identifying the onset and offset of that vowel. The second property is the fundamental frequency ( $F_0$ ): averaging the pitch values contained within the vowel.

The third property is the pitch range: determining the value of pitch change after finding the maximum and minimum pitch values. The fourth property is the vowel space area: determining  $F_1$  and  $F_2$  from LPC formant tracks at 50% of the vowel duration for each of the vowels produced by the talkers in each speaking style (following Ferguson and Port, 2002).

The entire procedure consisted of many recording sessions: one session per talker. It is also important to mention here that the experiment took approximately three weeks to complete as the non-native speakers were all teaching-staff members from different departments at different colleges. The time spent in doing all the acoustic analyses was nearly two months and a half.

## **2.4 Timing and Validity of the Experiment**

Time was distributed equally among all the talkers of the experiment. The time consumed in the CNV speech recording session with the native speakers ranges from 10-15 minutes for each talker, while the time consumed in producing the target



words in the CLR speech recording session was roughly one minute and a half. The researchers have made sure to conduct all the individual recordings within the same time limit.

The researchers have displayed all the research items, plans and procedures to a number of phoneticians such as Asst. Prof. Dr. Sarah H. Ferguson who is in charge of the Speech Acoustics and Perception Lab at the Department of Speech-Language-Hearing: Sciences and Disorders, University of Kansas. Another Phonetician was Prof. Dr. Keith Johnson who works in the UC Berkeley Phonology Laboratory at the Department of Linguistics, University of California – Berkeley. The third one was Prof. Dr. James M. Hillenbrand, Department of Speech Pathology and Audiology, Western Michigan University. They discussed many points concerning the methodology being followed, the design of the experiment, the selection of stimuli and talkers, the kinds of devices and equipments to be used, as well as discussing the subject matter, the language remarks and the specific terminology to be used. In fact, they acted as the native jury, whereas the non-native jury consisted of the following: Prof. Dr. Majeed Al-Mashta, Prof. Dr. Majeed H. Jasim, Prof. Dr. Balqis I.G. Rashid, Asst. Prof. Dr. Ala' Husein Oda, and Dr. Abdul-Kareem Talib.

Moreover, in order to ensure the validity of the experiment and the reliability of the measurements, the researchers have considered the methods followed and discussed by many phoneticians and researchers in the related studies. The studies include: Peterson and Barney (1952), Gagné et al. (1994), Hillenbrand et al. (1995), Schum (1996), Bradlow and Bent (2002), Ferguson and Port (2002), Bradlow et al. (2003), Ferguson (2004), Smiljanic and Bradlow (2005), Ferguson and Port (2007), among many others.

The value and significance of the study is also obtained from the statistics and results of the experiment. The researchers attempted to follow precise and scientific statistical techniques and analytical procedures.

#### **4. Results: Statistical Analyses and Discussion**

The four acoustic measures (vowel duration, fundamental frequency, pitch range, and the acoustic vowel space) are estimated and analysed for each sex for the two groups of talkers in both CLR and CNV speech. For vowel duration, the vowel onset and offset are marked, then duration is measured in (ms).  $F_0$  was measured by averaging the pitch contained within the vowel sound. In measuring pitch range, the minimum and maximum pitch values are determined and the change is displayed in numbers. As for the acoustic vowel space,  $F_1$  and  $F_2$  values are determined at 50% of the vowel duration and the result is displayed in a two-dimensional formant space.

Mean values are calculated for vowel duration, fundamental frequency ( $F_0$ ), and pitch range. For these three acoustic measures, the data were subjected to a factorial statistical analysis (ANOVA) with a Complete Randomized Design (CRD). The analysis involved three factors: Talker Group (TG), i.e., native or non-native, Talker Sex (TS), i.e., female or male talkers, and Speaking Style (SS), i.e., CLR or CNV speech. These three factors were considered the 'Between-Subjects Factors' of the experiment. For the two-dimensional acoustic vowel space, comparison was made to examine the vowel space expansion for all talkers in both speaking styles. Moreover, the ANOVA analysis was done at the (0.05) criterion of probability (P) to check the significance (Sig.) of the three acoustic measures mentioned earlier. In other words, for these three acoustic measures, values approach significance if (P . 0.05).

## 4.1 Vowel Duration

Mean vowel duration (ms) for long and short vowels produced clearly and conversationally by the female and the male talkers are shown in figures 1, 2, 3, and 4 respectively. Table (1) shows the ANOVA results of the long and short vowel durations.

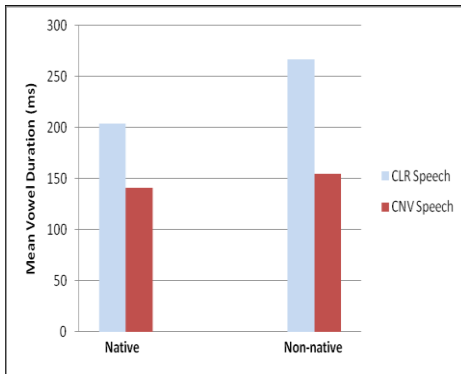


Figure (1): Mean vowel duration (ms) of long vowels produced clearly and conversationally by the female talkers

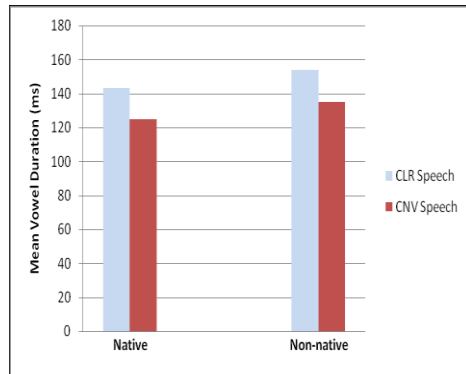


Figure (2): Mean vowel duration (ms) of short vowels produced clearly and conversationally by the female talkers

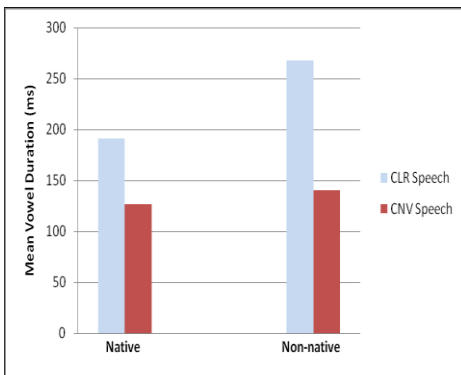


Figure (3): Mean vowel duration (ms) of long vowels produced clearly and conversationally by the male talkers

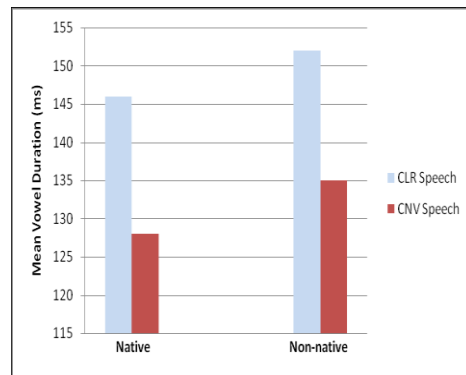


Figure (4): Mean vowel duration (ms) of short vowels produced clearly and conversationally by the male talkers

**Table (1):** ANOVA results of the duration of long and short vowels

Long Vowel Duration		Short Vowel Duration	
S.O.V	P(Sig.)	S.O.V	P(Sig.)
TG	0.009	TG	0.283
TS	0.894	TS	0.502
SS	0.000	SS	0.021

For both female and male talkers, the results have yielded significant SS effects where  $P(\text{Sig.}) = 0.000$  for long vowels and 0.021 for short vowels. This indicates that both long and short vowels are significantly longer in CLR speech than in CNV speech. This is due to the clear and carefully articulated manner of CLR speech.

However, within each sex, neither the effect of TS in native speakers, nor that of TS in non-native speakers for both long and short vowels approaches a remarkable significance because  $P(\text{Sig.}) > 0.05$ .

As for the TG effect, the results have shown that the non-native speakers produced longer vowels than the native speakers for both sexes and in both speaking styles,  $P(\text{Sig.}) = 0.009$ . While vowel durations of short vowels are not significantly different in both speaking styles,  $P(\text{Sig.}) = 0.283$ , and that both groups of talkers performed similarly in each style.

## 4.2 Fundamental Frequency

Figures 5, 6, 7, and 8 show the mean values of fundamental frequency ( $H_z$ ) in vowels produced clearly and conversationally

by the female and male talkers. Table (2) presents the ANOVA results for the fundamental frequency of the vowels involved.

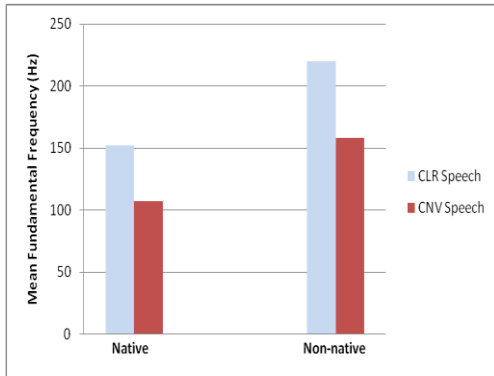


Figure (5): Mean fundamental frequency values ( $H_z$ ) of long vowels produced clearly and conversationally by the female talkers

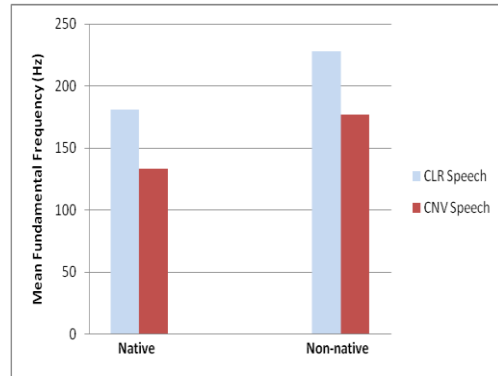


Figure (6): Mean fundamental frequency values ( $H_z$ ) of short vowels produced clearly and conversationally by the female talkers

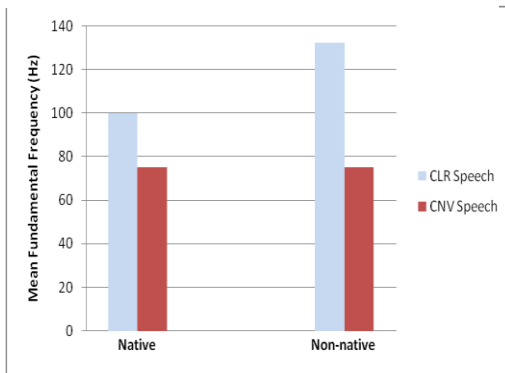


Figure (7): Mean fundamental frequency values ( $H_z$ ) of long vowels produced clearly and conversationally by the male talkers

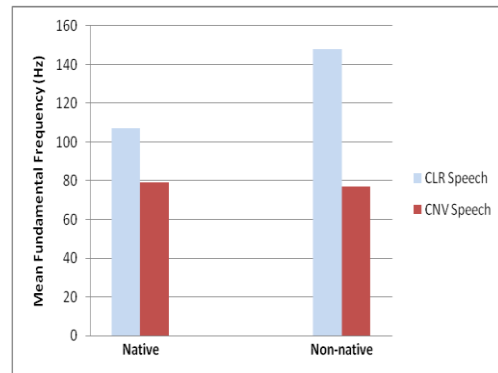


Figure (8): Mean fundamental frequency values ( $H_z$ ) of short vowels produced clearly and conversationally by the male talkers

**Table (2):** ANOVA results of fundamental frequency of long and short vowels

F <sub>0</sub> of long Vowels		F <sub>0</sub> of short Vowels	
S.O.V	P(Sig.)	S.O.V	P(Sig.)
TG	0.008	TG	0.007
TS	0.013	TS	0.011
SS	0.000	SS	0.029

For both male and female talkers, the results have shown a significant SS effect for both long and short vowels,  $P(\text{Sig.}) = 0.000$  and  $0.029$  respectively. The results have illustrated that all talkers have exhibited mean  $F_0$  values in the vowels that are higher in CLR speech than in CNV speech.

Moreover, within each sex, the effect of TS in native and non-native speakers is significant for both long and short vowels,  $P(\text{Sig.}) = 0.013$  and  $0.011$  respectively. For both native and non-native speakers, females tend to produce vowels with higher  $F_0$  values than males in both speaking styles.

As for the TG effect, the results have yielded significant effects where  $P(\text{Sig.}) = 0.008$  for long vowels and  $0.007$  for short vowels. Non-native speakers (both females and males) produced vowels with higher  $F_0$  values than native speakers in both styles. In other words, both groups of talkers didn't perform in the same way.

### 4.3 Pitch Range

Mean values of pitch range in vowels produced clearly and conversationally by the female and male talkers are shown in

figures 9, 10, 11, and 12 below. Table (3) displays the ANOVA results of pitch range values for both long and short vowels.

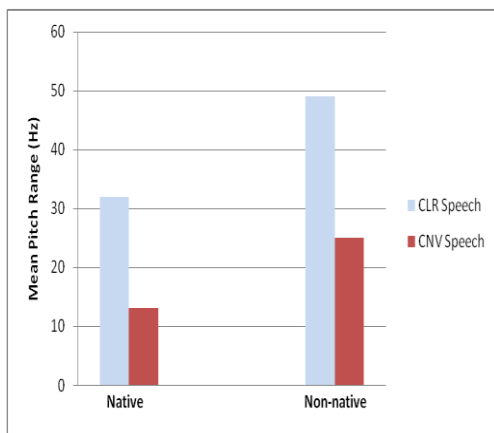


Figure (9): Mean values of pitch range in long vowels produced clearly and conversationally by the female talkers

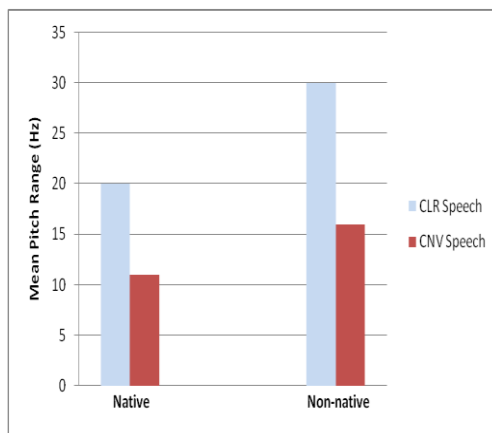


Figure (10): Mean values of pitch range in short vowels produced clearly and conversationally by the female talkers

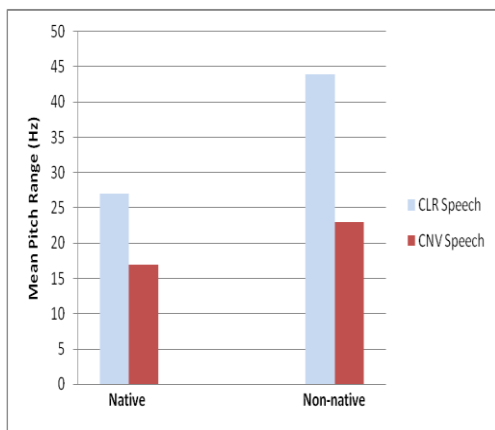


Figure (11): Mean values of pitch range in long vowels produced clearly and conversationally by the male talkers

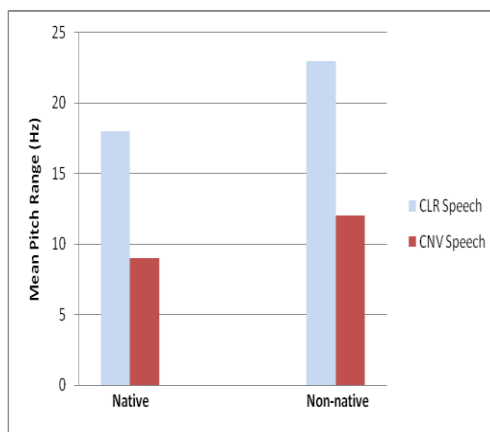


Figure (12): Mean values of pitch range in short vowels produced clearly and conversationally by the male talkers

**Table (3):** The ANOVA results of pitch range values for both long and short vowels

Long Vowels		Short Vowels	
S.O.V	P(Sig.)	S.O.V	P(Sig.)
TG	0.003	TG	0.479
TS	0.063	TS	0.840
SS	0.000	SS	0.000

As shown in the results of vowel duration and mean values of  $F_0$ , the main effects of SS have reached significance in both long and short vowels produced by both female and male talkers,  $P(\text{Sig.}) = 0.000$  for long vowels and  $0.003$  for short vowels. This indicates that all talkers expanded their  $F_0$  range in CLR speech to a greater extent than in CNV speech.

No significance was found within each sex in both groups of talkers because  $P(\text{Sig.}) = 0.063$  for long vowels and  $0.840$  for short vowels.

Moreover, a significant TG effect,  $P(\text{Sig.}) = 0.003$  is found between both groups of talkers in the production of long vowels. On the other hand, no significant differences in pitch range. ( $P(\text{Sig.}) = 0.479$ ), between both groups of talkers are found in the production of short vowels and that both groups performed similarly.

#### 4.4 Acoustic Vowel Space

Figures 13-20 below show the  $F_1$ - $F_2$  acoustic vowel space produced in CLR and CNV speech by the female and the male native and non-native speakers.



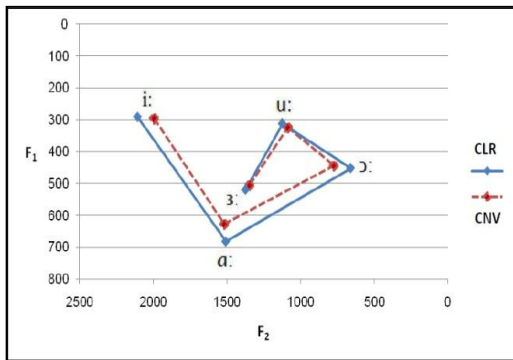


Figure (13): F<sub>1</sub>-F<sub>2</sub> acoustic vowel space of the female non-native speakers in CLR and CNV speech (long vowels)

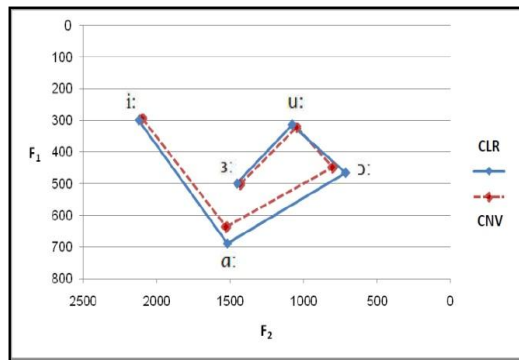


Figure (14): F<sub>1</sub>-F<sub>2</sub> acoustic vowel space of the female native speakers in CLR and CNV speech (long vowels)

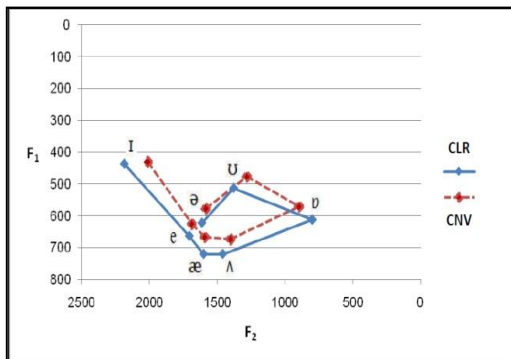


Figure (15): F<sub>1</sub>-F<sub>2</sub> acoustic vowel space of the female non-native speakers in CLR and CNV speech (short vowels)

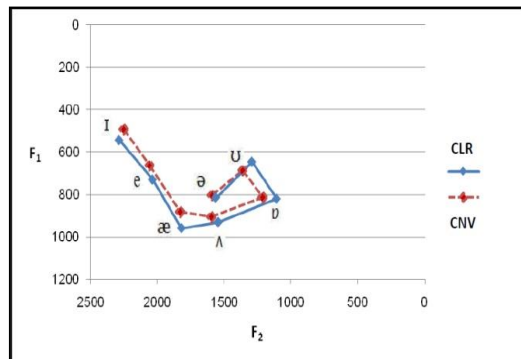


Figure (16): F<sub>1</sub>-F<sub>2</sub> acoustic vowel space of the female native speakers in CLR and CNV speech (short vowels)

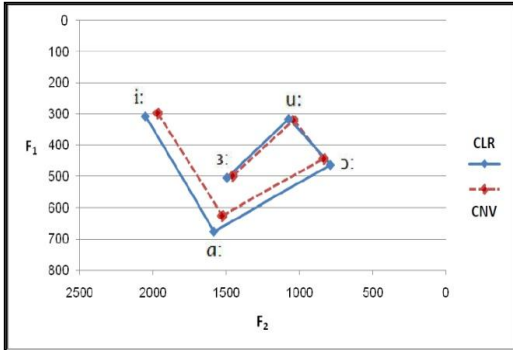


Figure (17):  $F_1$ - $F_2$  acoustic vowel space of the male non-native speakers in CLR and CNV speech (long vowels)

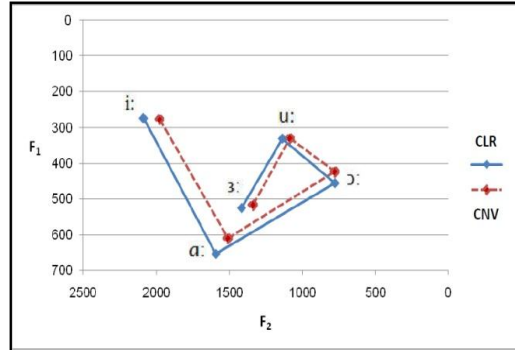


Figure (18):  $F_1$ - $F_2$  acoustic vowel space of the male native speakers in CLR and CNV speech (long vowels)

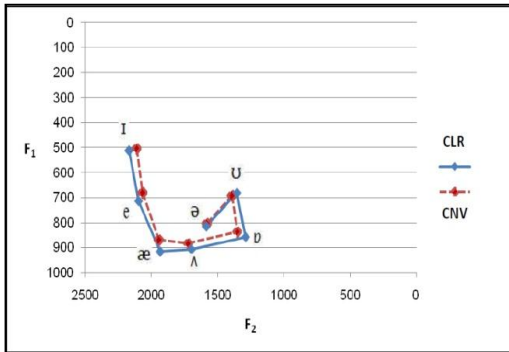


Figure (19):  $F_1$ - $F_2$  acoustic vowel space of the male non-native speakers in CLR and CNV speech (short vowels)

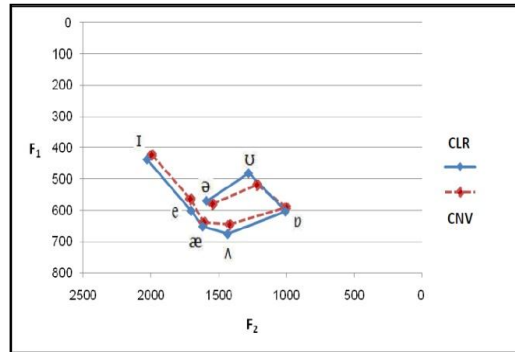


Figure (20):  $F_1$ - $F_2$  acoustic vowel space of the male native speakers in CLR and CNV speech (short vowels)

As shown in the figures above, the vowel space areas for both long and short vowels for all the talkers have been expanded in CLR speech. In other words, clear vowels occupy a larger area in the  $F_1 \times F_2$  vowel space than conversational vowels.

In general, and for both long and short vowels, the acoustic vowel space areas for the native and non-native female speakers have increased to a lesser extent than those of the male speakers. In other words, the female talkers had larger vowel space areas, larger  $F_1$  and  $F_2$  ranges, and greater distance among vowels than the male talkers.

Moreover, little differences have been found in vowel space expansion as far as the TG effect is concerned in both styles and for both female and male talkers.

### **4.5 Vowel Intelligibility**

It has been stated earlier in this study that speech is more intelligible if the sounds and words are produced clearly. This means that sounds, words, or even utterances are distinguishable and more understandable if they are produced clearly.

It has also been mentioned in 2.4 that there are various factors that affect speech intelligibility. The most important of these factors are the ones that were acoustically analysed and statistically displayed in the current study: vowel duration, fundamental frequency, pitch range, and the acoustic vowel space. These factors are all related to vowel sounds and they are tackled and studied for two main reasons. First, this study is designed to acoustically analyse English pure vowels in two speaking styles. Second and most important, vowel sounds are distinctive because they differentiate words from each other. The clearer the vowels are, the more intelligible speech will be.

The findings of this study have demonstrated that pure vowels have longer durations in CLR speech than in CNV speech. Moreover, all talkers of the experiment have produced the vowels with higher mean values of  $F_0$ , greater pitch range values, and larger vowel space areas in CLR speech compared to CNV speech. All these acoustic measures and enhancements are assumed to allow for more accurate identification of each individual sound.

This leads us to conclude that the vowels produced by the talkers are more intelligible in CLR speech than in CNV speech. Although this is not part of our investigation, CLR speech recordings appeared to be louder than CNV speech recordings, which also add to improved speech intelligibility.

In general, the instructional sets that were given to the talkers in each speaking style at the time of the recording sessions, are considered a viable strategy for improving speech intelligibility. These instructions made the talkers adopt each style perfectly and also made them decrease and increase their speech rates in CLR and CNV speech respectively.

#### 4.6 Discussion

Out of the statistically significant differences and as shown in the previous tables and figures, all talkers of the experiment exhibited the CLR speech properties. All talkers produced longer vowel durations, higher mean values of  $F_0$ , greater values of pitch range, and larger areas of vowel space in CLR speech.

ANOVA test using the SPSS statistical software is done for these acoustic measures, and the following table summarizes all these results:

**Table (4): The ANOVA results of the first three acoustic measures**

Vowel Duration				Fundamental Frequency				Pitch Range			
Long Vowels		Short Vowels		Long Vowels		Short Vowels		Long Vowels		Short Vowels	
S.O.V	P(Sig.)	S.O.V	P(Sig.)	S.O.V	P(Sig.)	S.O.V	P(Sig.)	S.O.V	P(Sig.)	S.O.V	P(Sig.)
TG	0.009	TG	0.283	TG	0.008	TG	0.007	TG	0.003	TG	0.479
TS	0.894	TS	0.502	TS	0.013	TS	0.011	TS	0.063	TS	0.840
SS	0.000	SS	0.021	SS	0.000	SS	0.029	SS	0.000	SS	0.003

For analysing the results of the fourth acoustic measure which is the acoustic vowel space, the researcher has depended on comparing the expansion in the vowel areas in both CLR and CNV speech.

The statistical procedures have revealed some significant differences in the strategies employed by both native and non-native speakers in the production of long and short vowels in both speaking modes. In other words, the results of the experiment demonstrated different values for vowel duration,  $F_0$ , pitch range, and the acoustic vowel space. Many differences are found between native and non-native speakers as far as these four acoustic measures are concerned.

Where there are significant SS and TG effects, there appears to be very slight TS effect between females and males in general. However, females often tend to produce longer vowel durations, higher  $F_0$  values, greater pitch range values and larger vowel space expansion than males do.

Moreover, because of the distinctive differences between both speaking styles and the characteristics they have exhibited, one can prove that CLR speech is clearly more intelligible than CNV speech.

## **5. Conclusions, Recommendations and Suggestions**

### **1.2 Conclusions**

It is mentioned earlier, and as hypothesized, that the acoustic properties of the target vowels which are produced in CLR speech are different from those in CNV speech. It has also been hypothesized that the strategies employed by both groups of talkers in exhibiting the acoustic properties of the vowels involved in both speaking styles are not the same. To that end, the researcher has conducted an experiment with a series of scientific investigations and analyses, and the results of that experiment have revealed the following:

1. Four main acoustic properties of vowels have been measured to determine the differences between CLR and CNV speech. These are vowel duration, fundamental frequency ( $F_0$ ), pitch range, and the acoustic vowel space. The results indicate that all talkers of the experiment, for both long and short vowels, produced longer vowel durations and higher  $F_0$  values in CLR speech. They also produced the vowels with greater values of pitch range, and larger vowel space expansion in CLR than in CNV speech. This means that there are large, diverse, and often competing acoustic differences between both speaking styles as hypothesized.

For vowel duration, the vowel onset and offset are identified by hand from high-resolution digital spectrograms using the software Praat. Duration is then measured in (ms). The duration of each vowel is not the same across the talkers of the experiment. Overall mean values of vowel duration have shown that long and short vowels are significantly longer in CLR than in CNV speech for all talkers and in both groups, i.e., native and non-native.

$F_0$  is measured by averaging the pitch over all the vowel region. As in vowel duration, talkers differ in their  $F_0$  values in both speaking styles. However, all of them produced higher  $F_0$  values in CLR than in CNV speech.

As for pitch range, the range is obtained by taking the difference between maximum and minimum  $F_0$  values over the entire vowel. As in vowel duration and  $F_0$ , there is a considerable across-talker variability in the amount of pitch range expansion. However,  $F_0$  range is increased in CLR speech than in CNV speech for all talkers. Mean values of pitch range have shown that this feature is increasing in

CLR speech, and it is one of the many consistent acoustic features of the CNV-to-CLR speech transformation across talkers in both groups.

For the acoustic vowel space,  $F_1$  and  $F_2$  frequencies are taken from the midpoint of each vowel and displayed in a two-dimensional figure for description. All formant measurements are made using an LPC formant tracking algorithm in Praat. The results have shown that the talkers differ in their acoustic vowel spaces in CLR and CNV speech. This is, of course, due to the vocal tract anatomical and physiological differences between the talkers in general and between the females and males in particular. Differences in vocal tract result in differences in formant frequencies, and hence different vowel positions in the acoustic vowel space. Mean values of the first two formants are taken and the results have shown that there is a significant vowel space expansion in CLR speech for all talkers and in both groups, i.e., native and non-native.

2. Mostly, and for all the acoustic measures, the female talkers produced longer vowel durations, higher  $F_0$  values, greater  $F_0$  ranges, and larger vowel space areas in both speaking styles than did the male talkers. This is of course because of the physiological differences between the vocal apparatus of females and that of males. This suggests that sex is an integral factor and plays a remarkable role in defining the acoustic properties of sounds.
3. As hypothesized, the findings of the current study have also demonstrated that the strategies employed by the non-native speakers are different from those of the native speakers, and that both groups did not perform similarly in roughly all of the acoustic properties.

The non-native speakers, both females and males, produced longer vowels in CLR and CNV speech than did

the native speakers. There are no significant differences in vowel duration between native and non-native speakers as far as the production of short vowels is concerned. The female and male non-native speakers produced vowels with higher  $F_0$  values than native speakers in CLR and CNV speech.

For pitch range, a significant TG effect is found between both groups of talkers in the production of long vowels. As in vowel duration, no significant differences in pitch range are found between both groups of talkers in the production of short vowels. As for the acoustic vowel space, slight differences are found in vowel space expansion between native and non-native speakers.

4. Most studies of CLR speech, in which talkers are instructed to speak clearly, have shown significant improvements in intelligibility. Instead of applying the diagnostic rhyme test or any statistical test to measure intelligibility, the researchers have performed a different strategy. The researchers have found a clear evidence that CLR speech is more intelligible than CNV speech. The present study has illustrated the multi-dimensional nature of speech and has displayed that a variety of acoustic factors can affect vowel intelligibility. The most important factors of these have been presented and measured here. Recall that vowels produced in CLR speech were longer, had higher  $F_0$  values, greater values of pitch range, and larger acoustic vowel areas than vowels produced in CNV speech. Hence, as these factors contribute to improved vowel intelligibility, one can easily infer that CLR speech is more intelligible than normal CNV speech.
5. The instructions that were given to the talkers in each speaking style played a markedly notable role in improving



6. speech intelligibility for they made the talkers adopt each style perfectly. With good and careful instructions, talkers can easily exhibit the characteristics of the speaking style as required.

## **5.2 Recommendations**

1. The output of the present study is very important to the teaching staff members and the EFL learners at the departments of English. It helps improving their speech intelligibility whenever needed as well as exhibiting the exact characteristics of CLR and CNV speech when communicating with native speakers and in the teaching classrooms as well. Therefore, it is highly recommended that speech styles and the speech intelligibility effect should be introduced in detail and be given more and more attention, especially in speaking skills courses.
2. As CLR speech is more intelligible or understandable than CNV speech, the researchers recommend using it especially in teaching pronunciation skills or in word-list reading.
3. A considerable interest has been given, these days, to acoustic phonetics as it is a promising field of research and because of its huge applications and great importance to speech research, technology and everyday life. Accordingly, the researcher suggests that this particular branch of phonetics should be taught in early years of university courses to get a good knowledge of it and its implication to speech, and correct pronunciation by means of introducing the most important acoustic procedures, properties,

4. and measures. No less important is to introduce this topic with the aid of computers and speech enhancements, and language laboratories.
5. The researchers also recommend using “Praat” and other speech analysis and synthesis softwares in doing acoustic analyses and in teaching language in general and pronunciation skills and properties in particular.

### **5.3 Suggestions for Further Research**

1. The present study is not intended to acoustically examine and analyse each vowel sound individually. Thus, it is suggested that future studies incorporate acoustic data related to all vowels and consonants in both CLR and CNV speech.
2. A detailed study of whether or not talkers differ in vowel intelligibility within a single speaking style is worthwhile investigating.
3. Conducting an acoustic analysis and incorporating articulatory properties of Arabic vowels is also worthwhile investigating.
4. This study focuses mainly on vowel production and its acoustic properties in CLR and CNV speech. A future study that integrates vowel production and perception, or that which include talkers and listeners, is recommended. Not only will this help talkers improve their speech, but also can help listeners in unfavourable listening situations.
5. Acoustic phonetics can also be used in clinical practice as in examining and analysing the acoustic properties of vowels and consonants produced by talkers with speaking

disorders. A similar work can be conducted on normal-hearing and hearing-impaired listeners.

6. It is important to note that the results of the present study are specific for the talkers of the current experiment. Other talkers may use different strategies for speaking clearly or conversationally, and the results may not be generalized to a larger population. However, the procedure and strategies employed by the researchers here can be followed and applied. The researchers suggest that examining a large number of talkers and comparing the results to the present study might be a useful idea.

### **Bibliography**

- Awaness, L. M. A. *Spectrographic Study of the Arabic Pharyngeal Consonants*. Unpublished M. A. dissertation, University of Baghdad, College of Arts, 1976.
- Ball, M. J., Perkins, M. R., Muller, N., and S. Howard (Eds.). *The Handbook of Clinical Linguistics*. Oxford :Blackwell Publishing, 2008.
- Birjandi, P. and M.A.S. Nodoushan. *An Introduction to Phonetics*. Tehran: ZatanKadeh Publications, 2005.
- Boersma, P. and D. Weenink. *Praat: Doing Phonetics by Computer* (Version 5.1.43) [Computer Program], 2010. Retrieved August 04, 2010, from <<http://www.praat.org/>>.
- Bradlow, A., Kraus, N., and E. Hayes. Speaking clearly for children with learning disabilities: Sentence perception in noise. *Journal of Speech, Language, and Hearing Research*, 2003:46, 80-97.
- \_\_\_\_\_, and T. Bent. The clear speech effect for non-native listeners. *Journal of the Acoustical Society of America*, 2002: 112 (1), 272-284.

- Bussmann, H. *Routledge. Dictionary of Language and Linguistics*. London: Routledge Group, Ltd., 2006.
- Clark, J., C. Yallop, and J. Fletcher *An Introduction to Phonetics and Phonology*. Oxford : Blackwell Publishing, 2007.
- Crystal, D. *A Dictionary of Linguistics and Phonetics*, 5<sup>th</sup> Edition. Oxford: Blackwell Publishing, 2003..
- Ferguson, S. H. Talker differences in clear and conversational speech :vowel intelligibility for normal-hearing listeners. *Journal of the Acoustical Society of America*, 2003: 116, 2365-2373.
- \_\_\_\_\_ and K. Port. Vowel intelligibility in clear and conversational speech for normal-hearing and hearing-impaired listeners. *Journal of the Acoustical Society of America* 2002:112(1), 259-271.
- \_\_\_\_\_ and K. Port. Talker differences in clear and conversational speech: Acoustic characteristics of vowels. *Journal of Speech, Language, and Hearing Research*, , 2007 50, 1241-1255.
- Finch, G. *Key Concepts in Language and Linguistics*, 2<sup>nd</sup> Edition. New York: Palgrave Macmillan, 2005.
- Fry, D. B. *The Physics of Speech*. Cambridge: Cambridge University Press, 1979.
- Gagné, J. P., Masterson, V. M., Munhall, K. G., Bilida, N., and Querengesser, C. Across talker variability in auditory, visual, and audiovisual speech intelligibility for conversational and clear speech. *Journal of the Academy of Rehabilitative Audiology*, 1994:27, 135-158.
- Goldberg, R. and L. Riek. *A Practical Handbook of Speech Coders*. Florida: CRC Press, LLC, 2000.
- Hillenbrand, J., Getty, L. A., Clark, M. J., and K. Wheeler. Acoustic characteristics of American English vowels. *The*

*Journal of the Acoustical Society of America*, 1995: 97, 3099-3111.

Holmes, J. and W. Holmes. *Speech Synthesis and Recognition*. London :Taylor & Francis, 2001.

Jenkins, J. *The Phonology of English as an International Language: New Models, New Norms, New Goals*. Oxford: Oxford University Press, 2000.

Johnson, K. *Acoustic and Auditory Phonetics*, 2<sup>nd</sup> Edition . Oxford: Blackwell Publisher Inc. 2003.

Ladefoged, P. *Elements of Acoustic Phonetics*, 2<sup>nd</sup> Edition. Chicago: The University of Chicago Press, 1996.

\_\_\_\_\_ *A Course in Phonetics*, 4<sup>th</sup> Edition . Boston: Thomson Wadsworth, 2000.

\_\_\_\_\_ *Vowels and Consonants: An Introduction to the Sounds of Languages*. Oxford: Blackwell Publishers, 2001.

O'Connor, J. D. *Phonetics*. Harmondsworth: Penguin Books Ltd., 1973.

\_\_\_\_\_ and C. Fletcher. *Sounds English: A Pronunciation Practice Book*. London: Longman Group Ltd., 1989.

Peterson, G . E., and H . L. Barney. Control methods used in a study of the vowels. *The Journal of the Acoustical Society of America*, 1952: 24, 175-184.

Rabiner, L. R., and R. W. Schafer. *Digital Processing of Speech Signals*. New Jersey: Prentice-Hall, Inc., 1978.

Roach, P. *A Little Encyclopedia of Phonetics*. London: University of Reading Press, 2000.

Sawusch, J. R. "Acoustic Analysis and Synthesis of Speech." In D. B. Pisoni and R. E. Remez (Eds.). *The Handbook of*

*Speech Perception*. Oxford: Blackwell Publishing, 2005: 7-27.

Schum, D. J. Intelligibility of clear and conversational speech of young and elderly talkers. *Journal of the American Academy of Audiology*, , 1996: 7, 212-218.

Smiljanic, R., and A.R. Bradlow. Production and perception of clear speech in Croatian and English. *The Journal of the Acoustical Society of America*, , 2005: 118(3), 1677-1688.

Stranzy, P. (Ed.). *Encyclopedia of Linguistics*. New York: Taylor & Francis Group, 2005.

Uchanski, R. M. "Clear Speech." In D. B. Pisoni and R. E. Remez (Eds.). *The Handbook of Speech Perception*. Oxford: Blackwell Publishing, 1996: 207-236.