CHAPTER ONE
INTRODUCTION

1.1 Background of the Study

Speech results largely from articulating strings of phones. These phones could be abstracted as individual concrete entities that can be perceived and given visual representation through phonetic symbols. Besides, there are some other aspects of speech which cannot easily be divided into segments. Rather, they are more or less perceived as the speaker performs his speech act. These other features are non-segmental which manifest in different variations as a speaker produces the sound segments in a continuous speech flow. Furthermore, these non-segmental features are important in, among other things, determining the level of competence of the speaker in a given language. They also help a language to give clear and precise information.

The non-segmental features identified which are relevant to this study are pitch, stress, tone and intonation. Each of these features, otherwise known as suprasegment, plays significant roles in speech and can bring about change in the meaning of utterances. They are therefore distinctive and contribute to the meaning of an utterance. Consequently, speech in terms of the sound component can be analysed from two linguistic points of view – segmental and suprasegmental.

When the native speaker of a language speaks, certain variations occur. These variations cannot be overlooked as mere verbal dexterity; they are also the result of the pitch pattern. Furthermore, they make the information more elucidating and also add to the aesthetics of discourse.

In a normal speech, for instance, some syllables are pronounced with greater intensity than others. It is for this reason that Roach (2000) says that in the production of stress more muscular energy is exerted than in an unstressed syllable. Besides, there are different degrees of stress such as the primary, secondary, tertiary. The pattern of pitch variation in any language helps to describe the melody and meanings in the language. For instance, in the English language, pitch occurs in a regular pattern over long domains. In such a situation, pitch is said to be performing intonational

function. Hence, the English language is regarded as an intonation language, and is also classified as a stress-timed language. That is, the intervals between stressed syllables in speech are either equal or more nearly so than the intervals between the nucleus of each syllable and the next.

In some other cases, as in most African languages, pitch variation occurs in a different pattern. Here, each syllable is assigned one form of pitch variation or the other in such a manner that such languages are described as syllable-timed languages. Syllable-timed languages, by the way, display a particular type of rhythm in their pronunciation. Such timing of syllables tends to be equal. The syllable-timed languages as Igbo and Yoruba are also known as tone languages.

Apart from the above fundamental principles of pitch variation in speech, the style of pitch variation a speaker applies at times may be as a result of the speaker's linguistic background and competence. This is so because one's competence in speaking a language depends largely on one's disposition and even exposure. For instance, one's native language can interfere in one's speech.

From the foregoing, it can be observed that pitch variation is used in different languages and in different patterns for different reasons of conveying meaning or achieving required linguistic function. Also, the application of the non-segmental features depends largely on the type of language. Nonetheless, different languages sometimes share the same non-segmental features, although the pattern may differ. Consequently, the level of competence of a tone language speaker in an intonation language may be a factor of his linguistic background.

Not many works have treated the interface of tone and intonation, especially from speakers of non-stress languages, hence the study of Igbo speakers of English to determine the points of relatedness of the suprasegments in these languages through acoustic study. The acoustic study will present accurate and authentic feature of the pitch patterns in the speech.

As a result of the above background, it is therefore necessary to investigate the interface of tone and intonation using the speech of English speakers of English and Igbo speakers of English.

1.2 Statement of the Problem

As there are languages so are there patterns of speech of these languages. This is so because each language of the world in some respect is spoken in a manner peculiar to it. Therefore, in speech, the native speaker of a language is expected to exhibit all the idiosyncratic qualities of the language. These linguistic peculiarities typify the language and often times mark it out from other languages. Good performance in speech and effective communication in any language, therefore, require knowledge of the suprasegmental features which play major roles in the language and their functions.

Although languages differ in one way or another, researches have also continued to report that certain linguistic features are shared by them. For instance, English and Igbo languages belong to different language types – intonation and tone; yet, evidences abound that they share some non-segmental features, although their pattern may differ in these languages. This study is set to identify some of these evidences and present acoustic analysis of them.

Some arguments hold that stress and tone are synonymous. But, typologically, stress is easily understood to feature in English language as tone is taken to be an important suprasegmental feature in Igbo language. Again, the question of whether there is tone in intonation languages, and intonation in tone languages has continued to excite arguments, many of which have left many scholars confused. Furthermore, an acoustic distinction of pitch patterns in the speech of various languages does not appear to have been clearly conducted. This has rather left a foggy notion about the tone and intonation features. Furthermore, only very few acoustic studies on the determination of interface of tone and intonation have been carried out, while a good number of attempts were carried out through ear perception. Similarly, acoustic study of speech has remained more of a proposal due to its high technical demands and few available scholars.

Therefore, ability to determine the features that correlate in two or more languages will greatly enhance good speech and effective communication. And to perform well in speech, the non-native

speaker of any language has to be conscious of the linguistic features of the language and their pattern of occurrence.

Generally, when the features and their pattern of occurrence are the same with those of one's native language, one may perform well. If, however, they are largely at variance, the speaker may be faced with great problem. This research, therefore, aspires to determine those areas of correlation of the suprasegments of pitch, stress, tone and intonation in the Igbo and English languages. This will be investigated acoustically using elaborate data collected from the speech in English of Igbo and English native speakers. Furthermore, it will investigate if the native language of a tone language speaker influences his speaking of an intonation language.

1.3 **Purpose of the Study**

Intonational characteristic in language is a universal feature employed by every language to distinguish meaning. But the patterns differ in languages and may even differ between varieties of the same language (Richards and Schimdt, 2002). When people speak, there are modulations of the voice which give rise to different shades of meaning.

However, the presence of intonation features in Igbo (tone language), or tone in English (intonation language) is not actually the contention of this study. Rather, the concern of this study is to determine the acoustic nature and points of connection of these linguistic (phonetic) features.

Notable works like Ikekeonwu (1993), Uguru (2000), Uzochukwu (2001) and Uguru (2006) and a few other works have made giant efforts at investigating the existence of intonation in tone languages and vice versa. This acoustic study is to investigate in detail the relatedness of tone and intonation examining the English speech (utterances) of Igbo native speakers. Thus, with the measurement of wavelengths plus the analysis of different acoustic properties, notably frequency and duration of the English speakers and the Igbo speakers of English, the necessary information about

the interface of tone and intonation will be authenticated. Then, the speculation surrounding their relatedness will be confirmed.

1.4 Significance of the Study

Effective communication in any language requires, among other things, knowledge and competence in the use of the suprasegmental features of the language. Our findings in this study of tone intonation interface have shown that Igbo (tone language) and English (intonation language) share such suprasegmental feature as tone. It also reveals that although stress is an independent linguistic feature, it can exist in a tone language as different forms of tone, and in intonation language as either word or sentence stress.

The study has highlighted the role of suprasegmental features of pitch, stress, tone and intonation in communication. Findings made in the research can therefore help comparative linguists in making categorical statements on interface of tone and intonation, particularly in Igbo and English languages. It will also identify the performance of a tone language (Igbo) speaker in English (an intonation) language.

The study will go a long way in providing literature for the study of linguistics, as well as serve as a reference material to phonologists, researchers and students that have interest in acoustic phonetics. It is also significant as it will provide avenue for further studies in instrumental analysis of speech. It will be a great asset to the applied linguists and second language teachers, particularly in the teaching and learning of tonal and intonational features of a second language, and improve teaching and learning of the intonation of English (Nigeria's official language) and Igbo (a major indigenous language of Nigeria).

The study will equally inspire researchers to engage in similar investigative study in these languages, and in other languages, as well as try other approaches. The study will benefit newscasters and actors/actresses a lot and encourage Nigerian English speakers that they can use the

intonation of English language and other languages correctly. The study then, will provide support to the assertion that tone languages have intonation just as intonation languages equally have tone (Clark, Yallop, Fletcher, 2007).

1.5 Scope and Limitation of the Study

The meaning that a language structure gives is not derived only from the inherent meaning of the individual words used in a speech. Rather, it also derives from the non-segmental or prosodic features that act upon the segments, syllables, words, or even a stretch of utterance. Hence, Ashby and Maidment (2007) say, "… all languages as far as is known use pitch variation to communicate". In other words, pitch (a non-segmental or prosodic feature) contributes to the meaning derivable from any speech.

In this study, therefore, the suprasegmental aspect of speech is investigated and the primary interest is the interface of tone and intonation in Igbo and English languages. Similarly, pitch variation in Igbo and English is our focus in order to use it to analyze the performance of tone language speakers of English. However, for convenience and effective study, the work is centred on the speakers of Igbo which is one of the major tone languages in Nigeria. Igbo is the major language spoken in the south eastern part of Nigeria.

English speakers' pronunciation comprising single words, compounds and sentences drawn from Peter Roach's and J. C. Wells' pronunciation tapes were used as control group. The utterances of speakers within the contexts are investigated instrumentally to determine where tone and intonation correlate. The study further investigates the performance of these respondents (native standard Igbo speakers) in speaking English, an intonation language and today Nigeria's second/official language and also spoken widely in the world.

Specifically, the acoustic properties, particularly duration and fundamental frequency of stress patterns of selected two-syllable, three-syllable and compound words, and their correlating tone

patterns were investigated. Also investigated are the default nuclear tones of selected sentences such as statements with definitive fall, tentative statements and yes-no questions and their correlating tones in Igbo (tone) language. In doing this, the study reserves detailed discussion on the linguistic feature, stress as it is not a major concern of this work.

Some challenges were, however, experienced in the course of this study. Notable among them were the dearth of resources, both human and material in the area of acoustic phonetics. as the area is yet the least patronised in phonetic studies. Other challenges include the high technical manipulations involved and the senstive nature of gadgets used in collecting data and for the analysis of data. Further to the challengess were the difficulty in sourcing native English speakers from our immediate environment, which led to resorting to tapes/CDs of recorded utterances of the English. The recordings of the utterances were not without their own challenges, as it was very difficult to get a totally noise-free environment for our recordings.

1.6 Research Questions

The following constitute the research questions for this study:

a. What are the different modes of realisation of intonation?

b. What are the regular suprasegmental features of tone?

c. Which tones and stress correlate in tone and intonation languages?

d. In what areas can we expect correlation between tone and intonation, particularly in the speech of a non-native speaker of English (an intonation language)?

e. Do tone language speakers encounter any problems in the articulation of the non-segmental features of English?

CHAPTER TWO

REVIEW OF RELATED LITERATURE

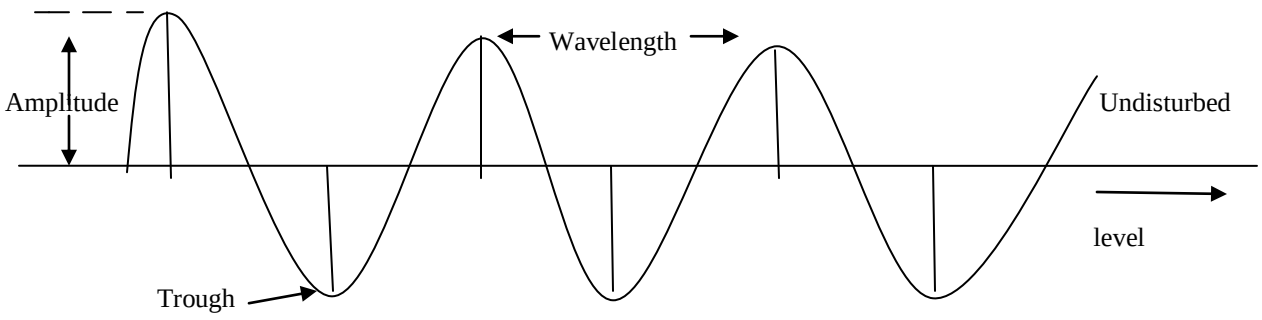
This chapter presents the theoretical framework that is adopted in this work. Secondly, some important works which discuss issues about pitch, stress, tone and intonation were reviewed. Particular attention was paid to such works that carry out analysis of stress reflection in speech. Also reviewed are works that discuss tone and intonation as suprasegments and their correlation.

Furthermore, works on general phonetics were reviewed to provide required background information. Review of some empirical studies on related studies already conducted on the topic was also done. All these shall provide aid for easier appreciation of the phenomena to be encountered in this work.

2.1 **Theoretical Framework**

The sound wave theory which Partee, Meulen and Wall (1990) recommended for acoustic analysis is applied in this study. Anyakoha (2010:248) defines wave as "a disturbance which travels through a medium transferring energy from one point to another without causing any permanent displacement of the medium". The source of every wave is vibration. Waves such as sound are transferred through a process known as wave motion. Sound wave is an example of mechanical wave; and a mechanical wave is an example of waves that require material mediums for their propagation. Below is an example of a wave.

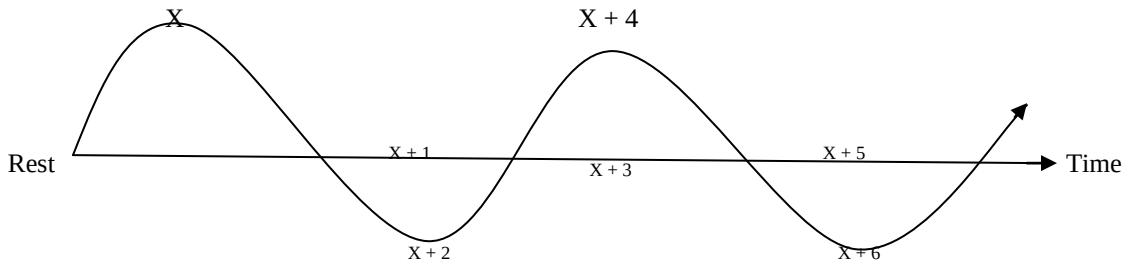
1.



Wave representation (Adapted from Anyakoha, 2010)

Anyakoha (2010) further explains that with every wave, the vibrations of the material medium which produce the wave, and the wave itself that carries the energy from the source of disturbance must be distinguished. In the case of sound, air is the medium or material in which the energy travels. Therefore, as the sound travels, the air buzzes about a mean position, thus passing the vibrations from one particle of the medium to another. Therefore, Anyakoha asserts: "The direction in which the vibration takes place is significant in classifying the type of wave" (p. 252). This is illustrated with the diagram below:

2.



A waveform indicating the pressure wave built up when particles move. The horizontal line represents the passage of time. (Adapted from Crystal, 1997).

In the acoustic analysis of speech sounds, this approach can be used to account for waveforms and wavelengths. It can also be used to determine voicing of segments and ultimately account for syllable weight. Ashby and Maidment (2005) note that the waveform of voiced and voiceless sounds are different and further points out that in periodic wave, the pattern of a wave repeats regularly. They observe that a period of simple sine wave runs from one identifiable point to another in an upwards-and-over excursion, and one downwards-and-up again excursion, returning to the zero line. Such complete movement is called a cycle. The length of one period is therefore the periodic time, *T*.

Sound wave theory is used to measure sound pressure and goes on to describe sound wave, how it is generated and the factors that affect it. It also measures the characteristics of sound waves such as sound velocity, frequency and wavelength. Hence, Anyakoha (2010) observes that there is a

very important relation between frequency (*f*), wavelength (ג) and velocity (*v*) of all waves, and gives the quotient of the relationship as followed:

v = fג; that is Velocity = frequency × Wavelength

During speech production, vocal cords come together and pull apart thus, producing some air pressure in the sound wave which is reflected in the frequency of the sound produced. The wave motion travels at a velocity depending on the frequency of the sound. When the pressure is high the frequency also becomes high while the wavelength becomes short. But if the pressure is lower, the frequency also becomes lower and the wavelength longer. The pulmonic airstream produces this pressure and while the vocal cords are together and pull apart, they interfere with the airstream, thus producing short pulses of air by chopping. The rate at which the chopping of the air occurs determines the pitch of the voice (fundamental frequency, Fo). Thus, the measurement of the *Fo* and its harmonics, according to Crystal (2003) is important in the analysis of intonation patterns, hence, its adoption. In this study, the wavelengths of the Control and Consultants were measured to enable us make authentic comparison between the utterances (languages). The major acoustic parameters measured are duration and frequency.

Harmonics, which is an acoustic property of sound, has been described by Crystal (2003) as a regular (periodic) waveform accompanying a fundamental frequency which helps to identify a complex tone. Therefore, the combination of a fundamental frequency and the amplitude of its various harmonics combine to give a sound its characteristic tone and quality. Hence, Crystal (1997:133) earlier notes: “This kind of framework is especially useful in analysing vowels, certain consonants, and intonation patterns”.

2.2 Theoretical Studies

This theoretical review is arranged in the following order - the prosodic features of speech (here, we shall review matters concerning pitch, stress, syllables, tone, intonation, relationship

between intonation and tone, the prosodic features in language: transfer and interference); the acoustics of speech production, and the acoustic analysis of speech.

2.2.1 The prosodic features of speech

Speech is not just a string of speech sounds, rather, it contains other features that are not easily observable but are more or less abstract elements. These non-segmental features, according to Ashby and Maidment (2007), spread over a number of speech sounds or a number of syllables or even whole utterances. These features referred to as prosodies or suprasegments would include pitch, stress, tone and intonation. They are as important in speech as the segments and are distinctive in speech. Lyons (1995), in his own account, adds that most linguists apply the term suprasegmental to refer to such things as stress, tone and length. But he contrasts this opinion with the assumption of American phonemics whose basic assumption was that the structure of words and sentences could be fully accounted for in terms of sequentially ordered phonological elements. He also says that tone, like stress, serves to distinguish forms in much the same way in tone languages that stress does in English, though not very extensively. Furthermore, Clark, *et al* (2007:327) describe prosody as "a set of higher level organisational structures that account for variations in pitch, loudness, duration spectral tilt, segment reduction and other associated articulatory parameters".

Catford (1994) refers to what he calls phonatory prosodies as those prosodic features that relate to phonation which include features of voice quality that in some languages affect stretches of speech longer than the segment, and features of pitch. He explains that phonation occurs when vocal folds vibrate at different frequencies; hence, voice is produced at different pitches. He further explains that change in pitch can take place in the following two ways – by stretching and tensing vocal folds in such a way that the tenser the folds, the higher the pitch; and by changing the pressure below the vocal folds, that is, the sub glottal pressure. Also, the higher the sub glottal pressure, the higher the pitch.

The above conditions support the fact that strongly stressed syllables have a tendency to be pronounced on a higher pitch than the unstressed ones. This tendency may, however, be language bound, hence Catford (1994:183) says, “there are languages, including dialects of English, in which most stressed syllables are actually pronounced on a lower pitch than the neighbouring unstressed syllables”. Such languages are the English of Aberdeenshire in Scotland and some varieties of Welsh English, Jamaican English and Ludian English.

In conclusion, therefore, the prosodic features contain both the resultant speech perceived consequent upon phonation and, at times, the meaning and or function of the word or utterance. Also, these suprasegmental or prosodic features do not only affect speech and the meaning but also they influence speech act and speech melody.

(i) Pitch and pitch patterns in speech

Richards and Schmidt (2002:402) define pitch as “the relative height of speech sounds as perceived by a listener”. Thus, when people speak, we hear sounds or groups of sounds in their speech to be relatively higher or lower than others. What we can hear as pitch, they go on, “is produced by the vocal cords vibrating (and) the faster the vocal cords vibrate, the higher the pitch”. Ladefoged (2003:75) sees pitch as an auditory process, and defines it as “the rate at which vocal fold pulses recur, and thus the fundamental frequency of sound wave”.

In speech, people use different melodic patterns. These patterns are distinctive as they are peculiar to languages. The melodic pattern of speech results from fluctuation in the pitch of the voice which Abercrombie (1967) refers to as “voice gesture”. He also observes that voice pitch fluctuation carries different implications which could be indexical or linguistic. Among the linguistic functions of pitch variation is what Abercrombie (1967) and later Collins and Mees (2006) refer to as “speech melody” which function is part of the spoken form of language. Therefore, pitch fluctuation (speech melody) can be found in all languages and it helps to classify languages. Hence, when pitch differences distinguish the dictionary meaning of words, the function of pitch is tone and languages

so affected are referred to as tone languages (Collins and Mees, 2006). Ashby and Maidment (2007) also affirm that pitch variation is used by all languages to communicate. But, the way pitch is used varies from one language to another. According to them, when pitch variation does not affect the meaning of individual words, but rather the way utterances are interpreted, the function is intonational. Languages that operate in this way are therefore classified as intonation languages. Catford (1994) affirms that pitch and pitch changes are utilised in languages in two distinct ways. According to him, "… variation of pitch may be related to relatively long stretches of speech, which may be syllables in length, and which correspond to relatively large grammatical units such as the sentence. Pitch variation used in this way is called intonation". He illustrates use of pitch in intonation with the following sentences in English:

3a. ˈJane was here ˈyesterday. [‾ · · ‿ . .]

b ˈJane was here ˈyesterday? [‿ · · — · ·]

c Jane was here yesterday? [\ · · — · ·]

In the above examples, the pitch patterns or intonation contour does not change the meaning of the sentence instead, it changes the function of the sentence as a whole from a statement (3a) to a question (b), and to a question with incredulous emphasis on 'yesterday' (c). Illustrating the above demonstration, Catford (1994) therefore says: "intonation as here, has a 'pragmatic rather than semantic' function".

Laver (1994) observes that pitch can also be word-based or syllable-based, and affirms that most African languages show syllable-based variation. However, he says that both word-based and syllable-based pitch variation fulfil tonal function. In other words, languages which manifest pitch this way can be said to be tone languages. Hence, Pike (1948:3) says of pitch reflection on tone languages: "A tone language may be defined as a language having lexically significant, contrastive, but relative pitch on each syllable". Continuing, Pike (1948) observes that for words to be spoken at all, they must be spoken on some pitch, whether in tone or in intonation language. This statement is

important and noteworthy for every word spoken is done on one type of pitch or another to accomplish the intended meaning.

Catford (1994) also observes that the pitch variations of a language may be related to short sketches of speech, typically of syllable length, and to grammatical units as words and morphemes. He refers to pitch variation used in this way as tone; and illustrates this phenomenon with this piece from Mandarin Chinese:

4a. ba [ˊ] 'eight' (b) ba [ˊ] 'to import'

(c) ba [ˇ] 'to hold' (d) ba [ˋ] 'a harrow'

The examples above show that the pitch differences occur on a single syllable. Secondly, they completely change the meaning of the word. He then concludes, "in other words, 'tone', as in these examples, has a 'semantic' rather than a pragmatic function". Therefore, these facts confirm that both intonation and tone involve pitch. What distinguishes one from the other therefore is the nature of pitch variation applied. Hence, in English, pitch helps to spell out the attitude or mood of the speaker. That is why Clark *et al* (2007) say that pitch is functional in English.

Ladefoged (2006:247-248) says that all languages use pitch to mark the boundaries of syntactic units. As he puts it, "in nearly all languages, the completion of a grammatical unit such as a normal sentence is signalled by a falling pitch". By this, he shows the universal aspect to ways in which languages use pitch differences to convey linguistic information.

In tone languages such as Igbo and Yoruba, pitch has a definite mode of occurrence on syllables and it influences the lexical relevance of the word. Hence, Clark *et al* (2007:342), citing Cruttenden (1977), state, "…in a tone language, tone is a feature of the lexicon, being described in terms of prescribed pitches for syllables or sequences of pitches for morphemes or words". That is, pitch is used to distinguish the meaning of words. Therefore, tone is regarded widely as a significant part of a syllable (or morpheme or word). Laver (1994), however, asserts that all languages whether they display tonal function or not, show intonational use of pitch. This observation has also been

made earlier by Schatcher (1965), Hyman (1973) and Abramson (1975) all of who claim that there is certain amount of movement displayed by the so-called level tones. The movement is at times in the form of rises or falls. Similarly, Hornby (1975: iii) observes that, "stress and pitch are closely associated". He states that "pitch is the relative height or depth of the level of the voice … (which) may be sustained (at a high or low level) or it may rise or fall". He further says that on any syllable with principal stress, there is typically a change in pitch which is usually either a rise or a fall.

Different languages therefore, manipulate pitch in different ways, usually for emphasis or for distinction between syllables (Mbah and Mbah, 2010). However, there is usually a particular realization of pitch for some syllables of an utterance thus giving rise to pitch prominence. This heightening or lowering of the pitch to a particular aspect of utterance is meant to draw attention. Therefore, any syllable that enjoys more prominence in a multisyllable word is referred to as the nucleus.

In speech, an individual's performance can be as a result of his pitch range. This pitch range gives rise to the perceived melody of a speaker's voice on any given occasion. Laver (1994) identifies three types of pitch range: the organic range, the speaker's paralinguistic range and the linguistic range. The organic range is the maximum range which the speaker's voice is physically capable of, given his or her laryngeal anatomy and physiology. This category borders on speech and gender. Ashby and Maidment (2005) and Clark *et al* (2007) identify the following Fo ranges for English speakers:

5.

		Ashby & Maidment (2005)	Clark, Yallop & Fletcher (2007)
Adult males	-	80 – 210 Hz	80 – 200 Hz
Adult females	-	150 - 320 Hz	150 – 300 Hz
Children	-		200 – 500 Hz

Laver (1994) describes the speaker's current paralinguistic range as the adjustment that occurs within the organic range which gives rise to a particular attitudinal information. Such situation, which is

also affirmed by Mbah and Mbah (2010) and Anagbogu, Mbah and Eme (2010), gives rise to the exhibition of surprise, sorrow, impatience, concern, etc in any given utterance. Chang (1972), Laver (1994), Onwudiwe (2014) contend that cultures vary in the way people utilize pitch-settings as part of paralinguistic communication. For the linguistic range of pitch, Laver (1994: 457) says, it is the range "within which the phonologically relevant pitch of the speaker's voice habitually varies in paralinguistically unmarked attitudinally neutral conversation". He also observes that different languages place this linguistic range at different degrees. Hence, it could be lower in some languages such as Spanish, but may be higher in French. He, however, says that Mandarin is spoken at a greater degree of prominence than French.

Celik (2001:9) refers to the linguistic range as 'key' and says, "The term 'key' can be described as utterance pitch; specific and/or meaningful sequences of pitches in an intonation unit. Keys that are linguistically meaningful and significant are worth being included in a syllabus". He then lists qualities of a significant key to include the following: (1) it should be under speaker's control, (2) it should be perceptible to ordinary speakers, and (3) it should represent a contrast. He also observes that three keys are usually identified. These are high, mid and low from which a speaker must choose from for conversation. He contends that most speech for a speaker takes place at the mid (unmarked) key. This key is employed in normal and unemotional speech. But, the high and low keys are marked. For the high pitch, "speakers generally exploit high pitch when they exclaim". Exclamation and indeed high pitch are used to express strong feelings like excitement, surprise, anger, agitation, cheer, merriment, fun, etc. And he continues, "high key is used for emotionally charged intonation units while use of low key indicates an existence of equivalence (as in appositive expressions), and relatively less significant contribution to speech". He illustrates the use of the various keys with the following examples:

6. Use of high key to express strong feelings –

high She: oh <u>GOD</u>
mid

low He: / have you GUESSED? /

7. Use of high pitch to indicate contrastivity –

high BOGnor /
mid / we're going to MARgate this year / not
low

8. Use of low key for co-reference, appositives -

high
mid / I TOLD you already /
low DUMmy /

In (8), the word *dummy* in low key is co-referential with *you* in mid key.

9. Use of low key to indicate non-defining relative clauses, parenthetical statements, expression of disagreement, reduced clauses, etc.-

high
mid / my DOCtor / / is very WELL-known /
low who's neuROlogist

10. Use of low key to indicate statement of opinions –

high
mid / the GOvernment / / will agree with our deMANDS
low I THINK

According to Celik, there are times when short statements of opinion that involve clarification, certainty/uncertainty, are attached to propositional statements as illustrated in (10) above.

Celik, however, makes this distinction between pitch and key: "the relationship between pitch and key is a comparative one in that syllabic pitch is always higher than the utterance pitch; in some sense, syllabic pitch is one step ahead of the utterance pitch". From the above discussion, it would be deduced that pitch contributes in different ways to language and speech act. On one hand, pitch sequences in language contribute to the melody of the language. This, of course, depends on the nature of the pitch sequences. On the other hand, pitch plays vital linguistic function in languages by contributing greatly to the meaning of utterances. Furthermore, pitch indicates the frequency of the vibration of the vocal cords, as well as help in determining the degree of prominence in the course of production of different syllables in a speech.

(ii) Stress and stress patterns in speech

Stress and pitch have been observed to have natural physical connection (See Pike (1948) and Laver (1994)). Consequently, stressed syllables are often pronounced on a different pitch from neighbouring unstressed syllables. Richards and Schimdt (2002) say that stress is the pronunciation of a syllable or word with more respiratory energy or muscular force than other words or syllables in the same utterance. They continue, "A listener often hears a stressed word or syllable as being louder, higher in pitch, and longer than the surrounding words or syllables". Hence, Blake (2008) describes stress as a phenomenon of loudness (pitch) and duration. The same view about stress is expressed by Ashby and Maidment (2005) and Wells (2006). Roach (2000) adds that stress is applied to, or is a property of syllables and asserts that stress is part of the suprasegmental phonology of English.

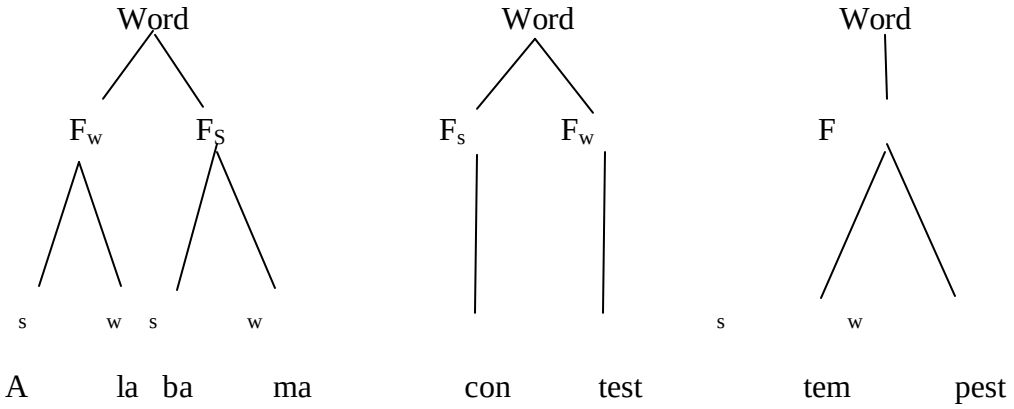
Wells (2006) further states that some languages as Greek use stress lexically (to distinguish between different words in dictionary). But, some other languages like French do not use stress lexically; in French, there are no pairs of words of different meaning distinguished by stress placement. He says, "English, like Greek is a stress language as stress is an important part of the spoken identity of an English word" (Wells, 2006:3). Some examples to illustrate this claim will include:

11. (a)	`abstract (noun)	(b)	ab`stract (verb)
	`pervert (noun)		per`vert (verb)
	`compound (noun)		com`pound (verb)

In (11a) above, more force is exerted in the first syllable of the words (nouns), while in (11b), extra force is on the second syllables. Thus, in line with Ladefoged (2003), "in every case the stressed word has a higher pitch and a greater length". To further illustrate stress assignment, http://www.uni-potsdam.de/u/anglistik/carroll/exte/SS/SS_lectures_2-5.htm cites Halle & Vergnaud's (1978) and Selkirk's (1980) introduction of the unit *foot* as the domain of stress assignment. According to them,

"each foot has a head (its strong syllable) and may also include optional weak syllables". He illustrates stress assignment to syllables with the metrical grid below using the words *Alabama, contest* and *tempest*:

12 (a) (b) (c)



He describes the foot as a prosodic category which groups syllables together. Foot, therefore helps us describe stress and rhythm.

To say that a syllable is stressed or prominent, Roach (2000) says that it needs to have the following attributes: loudness, length, pitch and quality. For Collins and Mees (2003), indicators of stress are intensity, pitch variation, vowel quality and vowel duration.

According to Roach (2000), a syllable is considered prominen t or stressed if it is made louder than the others. This is supported *in toto* by Collins and Mees (2003) where they say: "Intensity in physiological terms is the greater breath effort and muscular energy associated with stressed syllables. It's closely related to what is perceived by the listener as loudness." He also says that a stressed syllable is expected to be made longer than others during production. This aligns with Collins' and Mees' view of duration of vowel as one of the indicators of stress in English.

For vowel quality, Roach (2000) and Collins and Mees (2003) believe that when a syllable contains a vowel that is different in quality from the neighbouring vowels, the syllable will tend to be prominent. Therefore, they take stressed syllables to be those occurring against a background of the weak syllables, so that their prominence is increased by contrast with these background qualities.

The weak syllables, by the way are such contains short vowels as /i, u, ɪ, ə/ in English. Hence, two major types of stress exist – strong (primary) and weak stress. Therefore, words which have more than one syllable will have stress contrast (stressed and unstressed situations).

Gleason (1979:41), however, observes that monosyllabic words attract heavy stress like the stressed syllables of multisyllabic words. Therefore, he says: "We accordingly conclude that all monosyllables, when said in isolation, have primary stress". But, according to him, when words with more than two syllables are considered, it will be discovered that yet a third category of stress exists. He illustrates these three types of stress with the word 'dictionary'. Here, the first and second syllables receive the primary and weak stresses, while the third syllable receives the 'tertiary stress' represented thus / ˋ /. The primary and weak stresses, he concludes, are on their own represented as / ˈ / and / ˇ/ respectively.

Trask (1996) observes stress as a certain degree of prominence which is present in some languages upon certain syllables. He says that native speakers and phoneticians find it easy to determine the syllables that bear stress and even to distinguish varying degrees of stress but, however, says that the phonetic characterization of stress remains a very difficult task. While considering the ability of native speakers and phoneticians in determining stressed syllables, it may appear quite logical to add that non-native speakers will therefore find it difficult to determine which syllables bear stress as well as the varying degrees of stress. Consequently, realisation of stress varies from individual to individual. Drawing from his (Trask's) view, it can be observed that stress realisation varies from one individual to another. During speech, we do not have systematic variations that correspond to "syllable pulses", instead it is influenced by speaker's respiratory control. Consequently, Lieberman and Blumstein (1988) observe that when a speaker stresses or emphasizes a syllable or word, local "prominences" do occur. This, they say, is because "different speakers… regulate their pulmonary pressure in different manners". But, they say that the pulmonary

air pressure must fall at the end of the expiration of the pressure in all cases. In other words, this is uniform for all speakers.

As already established, stress assignment is on the syllable. In affirmation of this and contributing to the discussion, Laver (1994) adds that in some languages "stress falls on a fixed syllable in the structure of the word", while in others it is variable. He therefore says of stress assignment: "Lexical word stress, as part of the make-up of individual words, should be distinguished from emphatic stress, which is a property of utterances, with optional location on the syllables of the utterance". This is a confirmation of the belief that stressing can be lexically based or reflected on a whole string of words. Thus, Inkelas (1999) observes that languages assign stress in special ways. In his study of Turkish stress, he observes the exceptional pattern of Turkish stress, and argues: "The specific problem to be addressed … is how to handle exceptional patterns of Turkish stress: in the lexicon, with underlying templatic metrical structure, or in the grammar through morpheme-specific constraints". In his findings, he lists the Turkish stress pattern to include the following: "Final stress", "Sezer stress", "Prestressing suffixes", "stressed suffixes" and "stressed roots". He asserts that Turkish exhibits two distinct and productive patterns of stress assignment. Each of those patterns, he says, has its own morphological domain, and that each is assigned by a different sub-grammar, or "cophonology" of Turkish.

Continues Inkelas in his report on Turkish stress pattern, one pattern which is imposed by the word-level cophonology assigns final stress, while the other which is imposed by a stem-level cophonology assigns what he refers to as a non-final stress which is termed the "Sezer" pattern after its founder, Eugin Sezer. He then gives the specific function of each of these stress patterns thus: the Sezer stress pattern "is imposed on place names and some borrowings, (and that) it is unfailingly non-final". The Sezer stress pattern rule adapted from Sezer (1981) in Inkelas (1999:139) says: "If the antepenultimate syllable is heavy (H) and the penultimate syllable is light (L), stress the antepenultimate syllable; otherwise, stress the penultimate syllable".

Inkelas (1999) also discovers that "the Sezer pattern is imposed on names from other languages even when the stress in the source language is on a different syllable". He illustrates the above point with the following names as contrasted below:

13. **Original Name** **Sezer stress pattern**

Árkansas Ar.kán.sas

Santa Mónica Santamoníka

He presents another pattern of Turkish stress: the final stress, which he says is found in mono-morphemic and suffixed words. According to him, "Regardless of the number of suffixes, stress is unfailingly final". Agbedo (2000) also accounts for the stress pattern/assignment in such languages as French, Hungarian, Polish, Swahili and English with each displaying its own pattern. According to him, in French, final prepausal segments are somewhat stressed. In Hungarian, the place of stress is determined by word boundaries, while in Polish and Swahili, stress placement occurs one place before the end of a word. In English, he goes on; stress placement varies from word to word – some initial, some medial and some final. This is what Ashby and Maidment (2005) call variable (free) stress language.

Oyebade (2006:6) observes that in English, stress assignment is said to be "quantity sensitive". This assertion means that stress is assigned normally to a heavy syllable, that is, a strong syllable (Roach, 2000), or syllable with long vowel or diphthong (Ashby & Maidment, 2005). He further reports that in the absence of a heavy syllable, stress is assigned to the (ante) penultimate syllable. Again, he reports that since heavy or stressed syllables invariably bear a high or low pitch, it then follows that a low pitch is normally banned on a heavy syllable. This report is an account of stress assignment in a stress or intonation language (English), and this account appears synonymous with what obtains in such languages as Igbo and Yoruba where greater prominence goes to syllable with high tone.

Uguru (2006) claims that stress features in tone languages spoken in Western Delta region in Nigeria. But, she observes that stress reflection in stress languages is different from that in tone languages. She reports that in stress languages, words have particular stress order which she refers to as lexical stress, but in tone languages, what obtains is contrastive stress. Hence, she says, "… stress in Ika involves the use of stress for distinguishing degrees of emphasis or contrasts in sentences". By this, she asserts that stress features in tone and continues, "…stress in tone language is mainly for emphasis". This assertion agrees with Ikekeonwu (1993), and also supports the stress pattern in vowel lengthening in Igbo. In this same study, Uguru identifies other tone languages which use stress, like Duixi which she observes makes extensive use of stress. Thus, she says, "Just like stress is used for grammatical distinction in English, a stress language, so also is it used for lexical distinction in Duixi Mixtec. Kratochvil (1968) shows that Mandarin (Chinese language) manifests stress".

On the other hand Celik (2001) while recognising the notion of stress in words goes ahead to identify and discuss four types of stress. They are, the tonic stress, emphatic stress, contrastive stress and new information stress. He continues, "words that carry higher information content in the utterance are given higher stress than those carrying lower input (information) and those that are predictable in the context" (p. 3). The words that are generally believed to carry higher information are the content words as against the function words. Secondly, he notes that stress applies to syllables and asserts that it is content words, that is, nouns, verbs, adjectives and adverbs that are polysyllabic and not function words (articles, prepositions, conjunctions and modal auxiliaries). According to him, an intonation unit almost always has one peak of stress called 'tonic stress' or 'nucleus'. This type of stress, he contends, is preferred to the kind of stress used in "referring, proclaiming, and reporting utterances (and) is almost always found in a content word in utterance final position" (Celik, 2001:3). For example:

14. I'm <u>go</u>ing
    I'm going to <u>Lon</u>don

I'm going to London for a holiday

On some occasions, tonic stress can be moved from its final position in an utterance. This happens when the speaker wants to assign an emphasis to a content word, usually a modal auxiliary, an intensifier, an adverb, etc. like *very, indeed, great, surely, especially, alone, mustn't,* and so on. Such type of stress is called the emphatic stress. Celik (2001) illustrates emphatic stress with the following examples (adapted from Roach (1983:144):

15. i. It was very BOring. (unmarked)
ii. It was Very boring (emphatic)

i. You mustn't talk so LOUDly. (unmarked)
ii. You MUSTN'T talk so loudly. (emphatic)

In the case of contrastive stress, Celik (2001:4) has this to say:

> In contrastive contexts, the stress pattern is quite different from the emphatic and non-emphatic stresses in that any lexical item in an utterance can receive the tonic stress provided that the contrastively stressed item can be contrastible in that universe of speech.

Here, he points out, there is no distinction between content and function words. See his illustrations of contrastive stress below:

16 (a) Do you like this one or THAT one?
(b) I like THIS one.

(c) She played the piano yesterday. (It was her who ...)
(d) She played the piano yesterday. (It was yesterday…)

The new information stress projects the information supplied, or according to Bolinger (1968) (cited in Celik, 2001), highlights the most important and informative idea in the sentence. Thus, Celik (2001:5) concludes:

> So the stressed lexical item is that which carries the information enveloping communicative intent and purpose. The information in the stressed item is the core of the message within the utterance. Therefore, it is the most important

element in the utterance.

He illustrates the new information stress with both responses from wh-questions and declarative statements below:

17 (a) What's your NAME
(b) My name's GEORGE.

(a) What do you DO
(b) I'm a STUdent.

In the above examples, the new information is supplied to the question with a question word. Then, for the new information declarative statements, he illustrates it with these examples:

18 (a) It sounds like there was an excitement yesterday
(b) Didn't you hear? There was a torNAdo in the area.

Therefore, stress of whatever type or form serves to distinguish one syllable or word from another.

(iii) Syllable and syllable weight in speech

During speech production, the pulmonic air is released in parcels. These parcels are 'breath breaks' take place at the end of syllables. In his account of the phonetic incidences that occur during speech, Catford (1994:180) says: "the activity of pulmonic initiator is parcelled out, as it were, into short chunks, or small peaks of initiator power separated from each other by slight retardations of the initiator. These chunks are syllables…" So, speeches are made in parcels and each parcel is separated from the other in a string of speech.

A syllable may be produced with greater or lesser initiator power (stress) than the neighbours. Hence, Ashby and Maidment (2005) say, "… usually a stressed syllable is longer and louder than its neighbours …" Catford (1994) demonstrates this phenomenon with the word *phonetics* which is rendered in three chunks of initiatory activity, each bounded by consonants. He claims that the second syllable receives the primary stress. In other words, the second syllable is stressed.

Abercrombie (1967) had earlier said that in analysis of speech, one way to describe and compare the pronunciation of different languages is to analyse the syllable structure of the different

languages. Abercrombie (1967:34) describes speaking as “modified breathing”. This process shows that as the pulmonic air-stream mechanism is in action, the air is not expelled from the lungs by a constant, regular muscular pressure. Rather, “the respiratory muscles alternately contract and relax at a rate of roughly five times per second … in a succession of small puffs. Each contraction, together with the resulting puff of air constitutes the basis of a syllable”.

This view by Abercrombie (1967) holds that syllable is a ‘movement of the speech organs, and not a characteristic of the sound of speech. The ‘movement’, however, produces different chunks of speech each of which receives one type of pitch variation (tone) as in Igbo language. This syllable-producing movement of the respiratory muscle he refers as ‘chest pulse’, or ‘breath-pulse’, or ‘syllable-pulse’ while Catford (1994) refers to such mechanisms as ‘peak of initiator power’. But a ‘reinforced chest pulse’, according to Abercrombie (1967) gives rise to a stressed syllable, while the extra muscular movement which results to this is called ‘a stress-pulse’.

Although the syllable is an audible movement, Abercrombie (1967) observes that it is possible for a chest- pulse to take place silently, thereby producing an inaudible syllable. Such a chest-pulse occurs in the speech of many English speakers. Silent stresses, according to him play an important part in the structure of the English verse. Such stresses are often revealed by ‘synkinetic nod’ or other gesture preceding the audible syllable. When someone speaks we perceive a string of utterances with virtually equal number of stressed syllables. In such situations, there is a tendency for stressed syllables to occur at roughly equal interval in time. This has been observed by Catford (1994), Collins and Mees (2003) and Ashby and Maidment (2007) as a feature of English. This situation occurs in sentences, and according to Ashby and Maidment (2005:161) “… some of the stressed syllables remain stressed and others may lose their stress”. They further say, “stress in many languages is what defines the rhythm of speech”; and therefore define rhythm as “the pattern of occurrence in time of relatively ‘strong’ and relatively ’weak’ events”.

Catford (1994) contributing says that in such a string of utterances, the peak in each set indicates stressing and this is a "stress group" or "foot". He says that English speech tends to be delivered in a series of feet, each of which contains from one to several syllables and power curve. Within short stretches of speech, feet in English tend to be "isochronous" (that is of equal duration) and "isodynamic" (that is each involving about the same output of initiator power). On the rhythmic organisation of English, Laver (1994:532) states that this is characterised by "foot-level shortening". He describes this phenomenon as the situation where adding an unstressed syllable to a foot shortens the duration of the stressed syllable. He ends his contribution with Dauer's (1983) conclusion of her study of speech rhythm using English and Spanish thus:

> … 'the rhythmic differences we feel to exist between languages such as English and Spanish are more a result of phonological phonetic, lexical and syntactic facts about the language than any attempt on the part of the speaker to equalize interstress or intersyllable intervals'…

Laver's report on Dauer's (1983) study then stresses that the differences between the terms 'stress-timed' and 'syllable-timed' refer to what goes on within rhythmic groups, the characteristics of successive syllables and their interrelationships which are ultimately a product of the entire linguistic system.

Catford (1994), Collins and Mees (2003) and Ashby and Maidment (2005) agree that not all languages or varieties of English share this characteristic. Therefore, Collins and Mees (2003:117) have this to say:

> Stress-timing appears to operate for all types of English spoken by native speakers, possibly with the exception of those strongly influenced by Creoles … Some types of English employed as a second language (e.g. The English used by many Indians and Africans) absorb the syllable-timing of the mother tongue of the speakers ….

Collins and Mees (2003) here contend that whereas English speech observes equal stress timing, other languages, including most African languages rather observe equal duration of syllables in their speech. Languages like this are therefore referred to as syllable-timed languages. Hence, Uzochukwu (2001:75) contends: "An Igbo utterance can be split into units or syllables, each bearing a tone and each being of equal duration in utterance". Continuing, and citing Leech (1969) he says, "In this way, it is unlike English whose unit or foot "is usually larger than the syllable and contains one stressed syllable, marking the recurrent beat, and optionally, a number of unstressed syllables"". Uzochukwu (2001) therefore asserts that it is not possible to establish the foot in Igbo and as such the rhythm of Igbo poetry is non-metrical and free. He however says that the reason that people think that Igbo rhythm is metrical is "their tendency to equate high tone with stressed syllable and low tone with unstressed syllable". He affirms that one of the three factors that give rise to the free rhythm in Igbo poetry is "the regular recurrence of equal time duration in consecutive utterance". On this phenomenon, he says,

> One of the most common features of Igbo poetic rhythm is the regular recurrence of equal time duration in consecutive utterances that have equal syllable and, therefore, equal beats. In other words, the equality of the number of syllables reflects the equality of beats and is therefore an indication of equal time duration. This is because, Igbo being a syllable-timed language… it is expected that the utterances with the same number of syllables should have the same number of beats, …, and should therefore exhibit a recurrence of equal time duration (Uzochukwu, 2001:82).

Uzochukwu (2001) illustrates this "equal-time-ness" with the following song excerpt:

19. S.10		No. of Syllables	No. of Beats
Qgxuri	:	O sie ikē any[ab[a	
10		10	

```
                Nd[  Okweuri:                O  sie  ikē  any[  ab[a
10                   10
                Og              :               Q jxq oyī any[ alaa
10                   10
                Okw             :               Q jxq oyī any[ alaa
10                   10
```

(Soloist : When it is tough we come
Chorus : When it is tough we come
S : When it lacks lustre we go
C : When it lacks lustre we go)

In the excerpt above, there is a correspondence between the number of syllables and the number of beats. This is a clear indication of equal time duration in the consecutive utterances, thus justifying the claim that Igbo is a syllable-timed language. But Ashby and Maidment (2005) say, "both types can be found in most languages", just as "it is also true that different events of the same language may have different characteristics". This means that the stress/syllable-timing may vary from one speech to another depending on the meaning meant to be conveyed by the speaker.

Speech is largely an activity of the pulmonic airstream mechanism; and this mechanism itself is a product of two processes – the periodic syllable producing movement (syllable stress) and a reinforcement of these movements at intervals (stress process). These two processes are, however, co-ordinated in different ways by different languages. The manner in which stressed and unstressed syllables succeed each other is, according to Abercrombie (1967:36) what produces the rhythm of the language. Therefore, he asserts:

> It is probable that the rhythm of a language is one of the most fundamental things about it, in the sense that it is among the earliest things learnt by the infant, and perhaps the most difficult thing for the adult speaker to modify, when he wants to learn to pronounce a foreign language.

Similarly, tempo may be distinctive in languages. For instance, in English, according to Laver (1994:158) :"a medium articulation rate is about 5.3 syllables per second, and a medium speaking rate a little over 200 words per minute". He also distinguishes three tempos of speech viz. slow,

medium and fast. But, anyone learning a foreign language first has the impression that its native speakers use an exceptionally rapid tempo. Thus impression may be anchored on the assumption that a native speaker of any language articulates an utterance more easily and so more rapidly than a non-native speaker. Hence, Laver (1994:539) describes articulation rate as "the tempo of performance of all audible, 'vocalised' speech within an individual utterance, whether that speech consists of the manifestation of linguistic units of paralinguistic signals of hesitation". It follows therefore that a non-native speaker of a language, while battling with what Laver (1994) describes as "paraphonological syllables of hesitant filled pauses", his articulation rate will be slowed down. This claim, however, is not to obliterate the fact that some people employ more variation in tempo than others.

Consequent upon the views described above, Abercrombie (1967) observes that the prevalence of the dual conditions has made it necessary for people to learn to listen differently in order to be able to analyse speech rhythm, whether of one's mother tongue or another language, and to describe it in general phonetic terms. If this point is neglected, he notes, the problem of empathy in appreciating another language may remain.

(iv) Tone and tones in languages

Ladefoged (1971) describes tone as variations of pitch to convey lexical information. It is the "height of pitch and change of pitch which is associated with the pronunciation of syllables or words and which affects the meaning of the word" (Richards and Schimdt, 2002:555). Abercrombie (1967:105) says, "Tone is speech melody when it is a property of the word …", and the type of variation of pitch on a language classifies the language. Significant pitch distinguishes the meanings of utterances. Blake (2008) observes that the pitch varies over the course of an utterance in all languages. But, in some languages like Chinese, Vietnamese, Thai and in numerous African languages, he says, "a particular vowel can be pronounced with a higher-than-normal pitch or a lower-than-normal pitch to yield a word with a different meaning. There is also the possibility of

rising or falling pitch" (Blake, 2008:139). Wells (2006) also notes that tone (differences in the pitch of the voice) can be high level, mid level, low level, rising or falling. Hence, Pike (1948:3) earlier said, "When pitch is lexical, it distinguishes the meanings of words". It follows therefore that pitch changes on words influence their meanings. Wells (2006:3-4) illustrates this with the following example from Thai:

20.

[kha:]	with tone 1 (a mid level tone)	→	'to be struck'
	with tone 2 (low level)	→	galingale (name of a plant)
	with tone 3 (falling)	→	'value'
	with tone 4 (high level)	→	'to trade'
	with tone 5 (rising)	→	'leg'

Pike (1948) classifies a language which has lexically significant, contrastive but relative pitch on each syllable as a tone language. Hence, Richards and Schimidt (2002:555) define a tone language as "a language in which the meaning of a word depends on the tone used when pronouncing it"; and to be lexically significant, according to him, implies that pitch distinguishes the meaning of words. For instance, in Igbo, a notable tone language in Africa, the following illustrations show the significance of tone which therefore distinguishes the meanings of the words:

21 (a) akụ → palm kernel

àkụ̀ → wealth

(b) ǹdụdọ (dialect) → epilepsy

ǹdụ̀dọ̀ (dialect) → headiness

In 21(a) and 21(b), the change of pitch level (tone) from either high or low brings about the difference in meanings. It can further be observed that the different pitch levels fall on the different syllables of the words. Furthermore, Abercrombie (1967) affirmed later by Udoh (2003) says that in a tone language, a change in speech melody is very likely to change the meaning of individual words.

Lieberman and Blumstein (1990:198) in their own contribution observe that tone features play important linguistic role in many languages. Therefore, "the study of these 'tone' languages reveals that speakers execute deliberate laryngeal manoeuvres to produce consistent *fo* contours". Roach (2000) says that a syllable can be said with either a level tone or a moving tone. He, however, observes that the English language seldom uses level tone; rather, "moving tones are more common". He attributes this to the fact that, "English,…, is not a tone language, and the function of tone is much more difficult to define than in a tone language". This, however, does not preclude the presence of tone in an intonation language, instead it "makes use of tone intonationally, but not lexically" (Wells, 2006:5). Ashby and Maidment (2005:169) refer to this type of tone as "intonational tone" and describe it as a group of pitch patterns (nuclear tones) "which begin at the nucleus and continue to the end of the IP". This implies that tone in English does not alter the lexical identity of the word; rather, variation of tones presents "a range of intonational meanings". Celik (retrieved 2005:6) earlier described tone as certain pattern of voice movement which is not an arbitrary one as it is meaningful in discourse. He lists the range of meanings generated by tones thus: "By means of tones, speakers signal whether to refer, proclaim, agree, disagree, question or hesitate, or indicate completion and continuation of turn-taking in speech". Each of these meanings is initiated by either of the tones.

In African languages, for instance, the different levels of tone – like the high, mid, low combine to give contours like rising and falling. From the foregoing, it can be observed that some problems may arise when words from other cultures are introduced into a tone language community. Illustrating this claim with singing, Abercrombie (1967:106) says: "When songs are imported from other cultures into a tone language community – translations of hymns introduced by Christian missionaries, for example – the tunes may lead the singers to express sentiments different from those intended". This also holds true in the event of pronouncing words of an intonation language by a tone language speaker.

The above discussions hold true only when we discuss and analyse the nature and characteristics of the pitch patterns of tone and intonation languages. But, in actual speech mechanism, the nature of speech melody may no longer be as distinct as they are said to be in these two types of languages. In line with this assertion, Abercrombie (1967:109) earlier noted:

> …we sometimes find, in what is predominantly a tone language, a component of speech melody that has to be analysed as intonation, i.e. as part of the structure of sentences. Similarly, we may find in what is predominantly an intonation language a component of speech melody that has to be analysed as tone, i.e. as part of the structure of words.

This statement presupposes that there is very close correlation between tone and intonation languages which is the main concern of this study. This claim about the correlation between tone and intonation enjoys the strong support of people like Laver (1994) where he says that both tone and intonation languages exhibit intonational use of pitch.

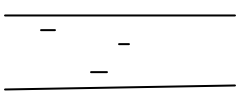
In his analysis of melody of language, using Yoruba tone systems, Laver (1994) citing Bamgbose (1969) says that Yoruba uses register tone system which has three level tones – high (′), mid (unmarked) and low (`). Laver (1994) also adds that there is a contextually conditioned rise to high. Abramson (1975:469) reveals that "… analysis of the fundamental frequency actually used confirms once again that the so-called level tone displays certain amount of movement through substantially less than the so-called rises and falls". Referring to Gandour's (1978) and Hombert's (1976) analyses of Yoruba tone system, Laver (1994:467) also says: "… analysis reveals, not surprisingly, that in real speech the pitches of the so-called 'level tones' are seldom strictly level".
It has to be noted, therefore, that rises and falls are characteristics of intonation.

Hyman (1975) describes downdrift as the phenomenon that results in high tones after low tones being phonetically less high than any preceding high tone in the utterance which can be

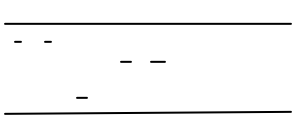
regarded as a kind of tonal assimilation with low tone influencing the height of the succeeding high tone. Then adds Laver (1994: 472), "To the extent that downdrift is a characteristic of potentially a whole utterance, it can be regarded as an intonational use of pitch". He, however, says that downdrift can result in several types of intonational patterns, and he demonstrates this claim with Urhobo as recorded by Kelly's (1969) study. According to the report, successive high tones in Urhobo remain at the same pitch height and successive low tones maintain a common level. Continued Laver (1994:473), while "high tones interspersed with low tones drift downwards … low tones interspersed with the high tones are represented by Kelly as maintaining their standard baseline".

This is illustrated as follows:

22. Tonal downdrift in Urhobo

(a) 

[ɔdìbɔ] 'banana'

(b) 

[uwekùgbeɹɔ] 'a nose and an ear'

Emenanjo (1978) and Mbah and Mbah (2010) state that Igbo belongs to the terraced level tone languages. But, Emenanjo (1978) contends that Igbo has two distinct tonemes – high and low. However, he observes that there is a tonal phenomenon known as downstep. While presenting the features of tone in speech, Emenanjo (1978:17) observes that Igbo has features that are related to what obtains in intonation. Therefore, he states:

> Outside the features of downstep and downdrift, there are a number of non-significant tonal features. One of these is the high falling-to-low-tone. This high falling-to-low-tone is a grammatically motivated tone which always occurs on the last syllable preceding the verb in the indicative verb form.

He illustrates the 'high falling-to-low-tone' with the following sentence:

23. ‘Ụlû gàwàrà ahịa → Ụlụ went to the market

Emenanjo, however, points out that in the Igbo system of writing, an extra (low tone) vowel of some quality as that preceding the verb would be preferred. The writing convention, according to him, would therefore give the above sentence as:

24. ‘Ụlụụ gàwàrà ahịa’ → Ụlụ went to the market.

It can, therefore, be deduced from (23) and (24) above that what happens here is vowel lengthening which gives rise to high falling contour tone which is a characteristic of intonation language. Similarly, Oluikpe (1979) observes that the Ngwa dialect (of Igbo) belongs to the register tone type of language. Besides this he, (Oluikpe 1979:203) states: “Tone glides are dominant feature of Ngwa tonemics. These occur primarily whenever there is a lengthened vowel in word-final position… There are two dominant tone glides – low-high / ˇ / and high-low / ˆ /”.

He subscribes to the theory that Igbo has two tones – high and low. He acknowledges that there is a feature known as downstep tone but argues: “there is no evidence that a mid tone begins a lexical tone; but a high and a low tone do begin a lexical tone in the language” (Oluikpe, 1979:204). In what appears an opposition to the views of Oluikpe (1979) and Emenanjo (1978), Mbah and Mbah (2010) say: “Igbo has three tones: the high, the low and the down-step. These terms describe their levels as regards their height or pitch level”. Nonetheless, we subscribe to the views of Oluikpe (1979) and Emenanjo (1978) that Igbo has two tones – high and low on the ground that the downstep tone is not an independent tone *per se*, but rather, a variant of the high tone.

Looking at the argument of Oluikpe (1979) and Emenanjo (1978) on the presence of intonation features in Igbo, and referring to earlier (and later to be followed) assertions, we corroborate that there are features of intonation in tone. It also nullifies the notion that tone languages are not rhythmic because of their popular pitch pattern of syllables. Hence, Ladefoged (1971) contends that the distinction between tone and intonation languages is not precise. Most tone languages have a

mixture of methods of using pitch. He then affirms: “There is probably no language, however tonal which does not have some features of intonation corresponding to a grammatical unit such as clause or sentence” (Ladefoged 1971:81); and further presents different types of tone languages with features of intonation thus: “(1) tone languages like Yoruba which have some lowering of the final tone; and (2) other tone languages like Hausa which have a falling or a rising intonation pattern over the whole sentence”.

He describes the above circumstance as “the superimposition of intonation on tone languages”.

Tone is undoubtedly of paramount importance in Igbo, as Igbo language itself is a tone language. Hence, tone performs important functions in the language. According to Mbah (2005), the function of tone has been agreed largely to be lexical and grammatical. He then says: “tone can and does perform only grammatical function, the exponent of which may be lexical, phonological, morphological, syntactic or semantic”. The importance of tone in Igbo had earlier been stressed by Mutaka and Tamanji (2000). Similarly, Okorji (2002) also supports the claim that tone is a very important suprasegment in Igbo. According to her, tones perform both lexical and grammatical functions. She identifies two phonemic tones in Igbo, viz: high tone (ˊ) and low tone (ˋ) and recognises the downstep (↓′) which she describes as a variant of high since it is a step lower than an ordinary high.

Furthermore, studies conducted on some dialects of Igbo have identified only two phonemic variants of tone. Hence, in her report, Okorji (1998:1) says, “Basically, Umuchu dialect has two phonemic tones: the high (ˊ) and the low (ˋ) tones. Every other tone that is applicable to the dialect is an allophone”. As in the standard Igbo, Okorji (1998) also observes that in Umuchu dialect, “tone performs lexical functions … and is used to distinguish the meaning of words which are morphologically similar”. These she illustrates with the following examples:

25. òny**ù**nyo ‘peeping’

òny**ù**nyò ‘shadow’

ǹtʰu ‘nail’

ǹtʰụ̀ ‘stinging powder’ (as in wasp or bee)

ushi̇̀ ‘odour/smell’

u!shi̇̀ ‘to warn sternly’

Again, in his study of Ngwa dialect of Igbo, Oluikpe (1979) equally asserts that tone is an important suprasegment that performs notable function in Igbo. According to him, tone glides occur in Igbo and they are a characteristic feature in some dialects of Igbo, such as Ngwa. Hence, he regards tone glides as “dominant feature of Ngwa tonemes”. Such glides, he continues, occur whenever there is a lengthened vowel in word final position. The dominant tone glides observed by Oluikpe (1979) are low-high (ˇ) and high-low (ˆ).

Mbah and Mbah (2010) advance some functions of tone in Igbo to include distinguishing between interrogative and declarative statements. For example, contrasting the H and L tones on the subject pronoun in the sentences below presents one of the phonemic qualities of tone.

26. O nyere ya ego ‘He/She gave him/her (some) money’

Ò nyere ya ego ‘He/She gave him/her money?’

Also, the placing of low tone on the resumptive pronoun equally turns the statement to a question. For example:

27. ‘Nchedo pụrụ apụ’ → Nchedo went out

‘Nchedo ọ̀ pụrụ apụ’ → Nchedo she went out?

Mbah and Mbah (2010) also say that tones are used to distinguish the cardinal from ordinal numbers as illustrated below:

28. ‘mmadụ̀ ìsiì’ (HHL LHL) → six persons

‘mmadū̀ īsiì’ (HH↓H ↓HHL) → sixth person

‘oche anọ’ (HH HH) → four chairs

‘oche ānọ’ (HH ↓HH) → fourth chair

Continuing, they say that tonal change in a construction that contains a relative (subordinate) clause helps to distinguish a sentence and a noun phrase containing a relative clause. Thus, in the following construction:

29. Ewu anyị bū rū ibù efùola ‘Our goat which is fat has got lost’.

In the above construct, the relative clause is introduced by the step tone on the verb *buru* which marks the relativised clause. Therefore, the ↓H tone marks the dependent clause, ‘buru ibu’.

Another remarkable function of tone in Igbo as identified in Mbah and Mbah (2010) is that it distinguishes nouns in associative constructions from possessive constructions. The following illustrations are used to present this function:

30 a(i) ìtè igwè (LL HL) ‘pot iron’ - ‘iron pot’

(ii) ìte Īgwē (LH ↓H↓H) ‘Igwe’s pot’

b(i) ùdù ntị̀ (LL HL) ‘drum ear’ – ‘ear drum’

(ii) ùdu Ǹtị̀ (LH LL) ‘Nti’s drum’

30a and 30b above illustrate the effect of tone on associative constructions in Igbo. Whereas the first set 30a(i) and 30b(i) which indicate more noun complements in the associative constructions retain their inherent tones, 30a(ii) and 30b(ii) which are possessive constructs are affected by tonal change that immediately changes their meaning and grammatical function. It is also worthy of note that the tones on the second syllables of the initial nouns of the second set of examples both changed from their inherent tone L to H as possessive constructions. Thus, H tone is again remarkable. (See illustrations about ordinal numbers and relative clause above). In these associative constructions 30a(i) and 30b(i); and 30a(ii) and 30b(ii) tone performs two functions – lexical and grammatical.

Ikekeonwu, Ezikeojiaku, Ubani and Ugoji (1999) observe that most often downdrifting occur when a high tone (H) precedes a low tone (L) in a chain. In such situation, the subsequent high tones will no longer be as high as the preceding ones in that order. This phenomenon according to them gives tone languages like Igbo and Yoruba the name terraced languages. See the example below:

31.

Ahà ya bụ Nnà

Name his is Nna

(His name is Nna)

In this illustration with the tone pattern HL H L HL, the initial H tone in ‘Aha’ is at the highest pitch level. But the subsequent ones in ‘ya’ and ‘Nna’ are no longer as high as the preceding one.

Another feature of tone which has been observed in Igbo is downstepping. This is the phenomenon in which when two high tones H occur in a word in succession, the second high tone drops slightly, but not to the level of a low. For example:

32. ‘egō’ → money

‘ezē’ → teeth

‘ajā’ → sand

In these examples, each of the two-syllable words has two high tones. But the second of each set drops slightly. Consequently, the pitch levels of the two tones also change slightly. However, Ikekeonwu, *et al (*1999) observe that in some dialects of Igbo, like Ọhụhụ-Ụmụahịa and Mbaise, the above tones would be realised simply as HH. (Also, refer to Oluikpe (1979) above). Nonetheless, Ikekeonwu, *et al.* (1999) warn that it is not each time the high tones occur in succession that the penultimate high tone steps. They support this argument with such examples as follows:

33. ‘isi’ → head

‘anya’ → eye

‘ewu’ → goat

‘egbe’ → kite

‘ekwe’ → wooden gong

The above discussions, therefore, show the prominent roles played by tone in Igbo as well as in some other tone languages. It also indicates that tone appears very significant in Igbo. Hence, in her contrastive study on Igbo and Hausa, Anidobe (2007) observes that tone is more complex in Igbo than in Hausa. She supports her claim by saying "… while a change of tone can give two interpretations to one Hausa word, a change of tone can give up to four interpretations to one Igbo word".

(v) Intonation and intonation patterns

Intonation is the change in the pitch of the voice as we speak. As an individual speaks, the pitch of his voice may be high or low, or as explained by Roach (2000:150), the pitch of the voice may be "ranging from "light" to "heavy" … or from "left" to "right"". Consequently, no two speakers speak on the same pitch level, as one may speak with lower pitch than the other. But, it is important to note that every individual has control over his or her pitch. Hence, he or she may choose to speak with a higher than normal pitch. Thus, speakers give some syllables in utterances a greater degree of loudness and change their speech rhythm giving rise to different intonation patterns. Each intonation pattern signifies a particular form of sentence such as, declarative, question, affirmative, annoyance, etc. or attitude. Each form of sentence is a consequence of either of these forms of intonation pattern enumerated by Roach (2000): level (moving) tone, falling tone, rising tone, fall-rise and rise-fall.

O'connor and Arnold (1973) highlight three basic characteristics of intonation which state: (1) Intonation is significant – in this case, intonation can bring about differences in meaning of an utterance. O'connor and Arnold (1973) here share the view of Abercrombie (1967) who has earlier said that in an intonation language a change of speech melody will likely change the meaning of the sentence as a whole. (2) Intonation is systematic – words, sounds and tunes used in speech are not invented arbitrarily, rather, they are learned and used to specification. According to them, there is a limited number of pitch patterns in any one language which are used to produce definite meaning

effect. Therefore, frequently recurring patterns of pitch can be described to give rules for their use. (3) Intonation is characteristic. This, they explain, means that the pitch patterns or tunes of an intonation language are not necessarily the same in form as those of other languages, and they do not produce the same effect as they would in other languages. These characteristics of intonation distinguish such languages as English which is a notable intonation language from other types of languages such as tone languages.

Ladefoged (1982) describes intonation as the 'up and down' movement/application of pitch in conversational utterance. Therefore, he states, "the intonation of a sentence is the pitch pattern of changes that occur; (and) the part of a sentence over which a particular pattern extends is called a tone group". He explains that a short sentence often forms a single tone group; longer sentences are made up of two or more tone groups. He also states that although the actual pitch changes that occur in speech are complicated, and often involving small pitch increases on each stressed syllable, the minor pitch changes are often neglected. The reason for the neglect is simply because intonation is considered to be only those changes that characterize a tone group as a whole.

Reiterating the point earlier made by O'connor and Arnold (1973), Ladefoged (1982) affirms that there is no systematic unit that exclusively corresponds to a tone group. Rather, when speaking slowly, say in a formal style, a speaker may choose to break a sentence up into a larger number of tone groups or 'intonation units'. This again shows that any word in any group may be important depending on the context and the speaker. Also, Roach (2000) explains that intonation units are very necessary in removing ambiguity in sentences. Illustrating this claim, Ladefoged (1982) uses the following sentence:

34 (a) "The 'girl / gave the 'money to her 'father"

This sentence, according to him, can normally be broken into two tone groups or according to Ashby and Maidment (2005) intonation phrases (IPs) as shown above. He however says that one can also break the sentence into three groups as follows:

34 (b)

The ˈgirl // gave the ˈmoney // to her ˈfather

Concluding his contribution, therefore, Ladefoged (1982) says that the way a speaker breaks up a sentence is purely idiosyncratic because it is what the speaker considers to be the important information points in the sentence that determines where he breaks up the sentence. He also says that a tone group is a unit of information, and not a syntactically defined unit. He then asserts that it is only in rapid conversational style that there is likely to be one tone group per sentence. The point made here by Ladefoged (1982) is a salient one for it affects the pattern of pitch and the flow of speech in general. It holds therefore that an individual can apply a pitch pattern that ordinarily may appear different from some other person's and in that wise present a different intonation.

Ashby and Maidment (2005) describe intonational phrase as different 'chunks' of an utterance which the speaker is at liberty to determine how many to use and where they should start and end. Within each IP, are a group of pitch patterns, otherwise called 'nuclear tone' "which begin at the nucleus and continue to the end of the IP" (Ashby and Maidment 2005:169). This group of pitch patterns they call "intonational tone" which according to them can be '(high) falling (˪)' or 'falling-rising' (˩˧).

Celik (2001:2) says of an intonation unit:

> An 'intonation unit' is a piece of utterance, a continuous stream of sounds, bounded by a fairly perceptible pause. Pausing in some sense is a way of packaging the information such that the lexical items put together in an intonation unit form certain psychological and lexic-grammatical realities.

In this presentation, Celik uses pausing to explain the breaks or units of utterances such as subordinate clauses and prepositional phrases all of which introduce variations in tone and meaning. Hence, Celik concludes, "Closely related to the notion of pausing is that a change of meaning may be brought about; certain pauses in a stream of speech can have significant meaning variations in the

message to be conveyed". To him, right pausing, that is right breaking of utterances into units may become a necessity to understand and to be understood well. He illustrates this view with the following examples which he adapted from Roach (1983):

35 (a) Those who sold quickly / made a profit
(A profit is made by those who sold quickly.)

(b) Those who sold / quickly made a profit
(A profit was quickly made by those who sold.)

We can see the change in meaning of the pair of the utterance occasioned by the pauses indicated with the slashes.

Tench (1996) in his own part, observes that intonation performs six distinguishable functions in English. He categorised these functions into two namely, melodic and communicative functions. He lists the melodic functions as:

(i) Marking of information status, e.g. the use of fall-rise tune pattern for non-finality;

(ii) Showing speaker's communicative intention. This is demonstrated by the type of utterance; it could be interrogative, declarative, exclamatory, etc.

(iii) Marking speaker's attitude. This can be explicated with tone in any one of these ways: anger, joy, friendliness, etc.

The communicative functions of intonation as highlighted by Tench are:

(i) Functions relating to information, reality and belief. This has to do with the speaker's certainty of utterance.

(ii) Suasion – that is, the effect of speaker's utterance on listener's behaviour.

(iii) Social interaction. This has to do with using tone to express thanks and farewell.

Similarly, Wells (2006) recognises six functions of English intonation as follows:

(i) The attitudinal function: This is the use of tone to express shock or surprise, pleasure or anger, interest or boredom, etc. He considers the most obvious role of intonation as the expression of our attitudes and emotions.

(ii) The grammatical function: He says that intonation plays the type of role played by punctuation in writing by identifying grammatical structures in speech. Here, intonation is used to mark the beginning and end of grammatical unit such as clauses and sentence. This specific function, he calls demarcative function using tonality. Also, intonation can be used to distinguish clause types such as question versus statement and disambiguates grammatically ambiguous structures. This aspect of grammatical function, he calls syntactic function, and he says that it is done mainly by tone.

(iii) The focusing/accentual/informational function. Here, intonation helps to determine what information in an utterance that is new and the one that is already known. It is used to bring some part of the message into focus and leave others; to emphasize or highlight some parts and not others. This, he says, is done by tonicity and the placement of other accents. He considers this the most important function of English intonation. He also adds that accentuation and the choice of tone are combined to present some longer stretches of the message which constitute the foreground of the picture we paint, leaving others. He refers to these as the pragmatic functions of intonation.

(iv) The discourse/cohesive function. This is a situation where intonation signals how sequences of clauses and sentences go together in spoken discourse, to contrast or to cohere. This function enables us to signal whether or not we have come to the end of the point we are making; whether we want to keep talking or are ready to give another speaker a turn.

(v) The psychological function of intonation, according to Wells, helps us to organise our speech into units that are easy to perceive, memorize and perform with the aid of tonality.

(vi) The indexical function: Here, intonation acts as a marker of personal or social identity. As he explains, “what makes mothers sound like mothers, lovers sound like lovers, lawyers sound like lawyers…officials sound like officials, partly, their characteristic intonation”.

Tench’s (1996) and Wells’ (2006) expositions on the functions of intonation are vital to our understanding of intonation as a suprasegmental feature. A careful study of these functions will show that four of Wells’ propositions will neatly knit into Tench’s versions. For instance, the melodic

function that marks information status and accentual function aspect of focusing can be merged. The function showing speaker's communicative intention and syntactic function aspect of grammatical function can go together, while the function marking speaker's attitude blends with Wells' attitudinal function. The function relating to information and reality and belief can go with the indexical function just as that on suasion is the same with the pragmatic aspect of the focusing function. Tench's social interaction of intonation and Wells' attitudinal function can be likened. Therefore, it should be noted that Wells' discourse and psychological functions of intonation are exceptions from those of Tench and they are as important as the rest of the others.

(vi) Interface of tone and intonation

As it has been strongly argued that stress features in tone languages, so also is it a strong claim that tone and intonation share so many features in common. For instance, Clark, *et al.* (2007:339) aptly state "… in fact tone languages have intonation too, as do all languages". Wells (2006) also accepts that all languages share this suprasegmental feature intonation but in different ways. As he puts it, "Like other prosodic characteristics, intonation is partly universal (= the same in all languages), but also partly specific (= differing from one language to another)"; and therefore concludes, "Languages differ in the intonation patterns they use, and in the extent to which they rely on intonation to convey aspects of meaning" (Wells 2006:5). It is in line with this that Clark, *et al.* (2007) again say that intonation is something superimposed upon the intrinsic meaning of words themselves, conveying the speaker's attitude and not any fundamental meaning. He observes that the same physical pattern of rises and falls may as well have different meanings or different pragmatic implication in different languages.

Uguru (2006) asserts that tone exists in intonation with the argument of Maidment (1990) who asserts that the different intonation patterns are made up of two tones – high and low – which are also existent in a tone language. However, she says that it is the direction of movement of these tones in speech that brings about intonation. The tones could be increased, reduced or left as they are. The

high tone, she says, could be increased to get a high rise, or reduced to get a high fall. She further observes that Goldsmith (1974) has been able to analyze English as a tone language allocating the tonic stress to the high tone. However, Uguru (2006) observes that it is not all the tone languages that have stress. She cites the acoustic and perceptual experiments conducted by Du (1988) which, according to her shows that Taiwanese (a tone language) does not have stress.

Citing Ikekeonwu (1993), Uguru (2006) argues that high tone is used specifically to reflect focus, and that the high pitch is borne by only one syllable in a tonal intonation group (TIG). Furthermore, the syllable with the focal high pitch (FHP) is usually higher than the preceding high tones, while the high tones of succeeding words are either downstepped high or low tones.

Accounting for stress in Ika, Uguru (2006) reports: "The variables of weak vowels and prominence may be used to account for stress in Ika". She recognises the fact that most syllables in Ika are produced with some amount of prominence because of the existence of lexical tone. But she adduces two reasons which can be used to prove that stress features in Ika. The reasons are: (1) that in most Ika utterances, all the syllables are not pronounced with equal amount of prominence. Rather, most of the time, the last (or at few other times the penultimate) syllable in an utterance bears a greater prominence than the other syllables. The greater prominence comes because it is the last syllable that bears the intonation pattern. She illustrates the above claim with the following example:

36. Ishì mānyā ǹ → ishi ma^nya n! – The stench of this drink

In (36), the prominence on 'manya' draws the rise fall intonation pattern on the emphasized syllable – 'nya'. However, it is observed in (36) above that low tone precedes a step tone which is not the convention in the tonal structure of the standard Igbo.

Another factor she gives which accounts for stress in Ika is (2) "the feature of compression effect in the alphabet'. According to her, this reduces the number of syllables in an utterance for rhythmic effect and therefore results in vowel reduction and production of such weak vowels as the schwa / ə / and the light vowel / ɪ / as can be found in English. For example:

37. Be wē jē → Bə we je - They will go.

Mbah (1996) observes that two main types of tone languages exist. These are contour and register tone languages. The contour tone language (CTL) includes languages where tone pitches glide from one pitch to another with their corresponding semantic change. Here, tones are not discrete and the semantic change generated from the pitch change is determined by the direction of the movement of tone glides. He gives examples of contour tone languages to include Mandarin Chinese and Thai.

On the downdrift tone languages DTLs, Mbah (1996) contends that this type of register tone language has varying tones which "criss-cross acoustically", the criss-crossing being attributed to fossilization. He quotes Williamson (1993) and Dimmendaal (1975) as positing that tonal fossilization is a by-product of augment shedding occasioned by elision at the segmental tier in some nominals whose protoforms can be reconstructed using historical evidence. Augment shedding influences tones to either reduce or increase in height. He then gives Yoruba and Kana as examples of languages with downdrift tone patterns.

Mbah asserts that terraced-level tones are consequences of downdrift. This is because tones influence one another across phonological units. Citing Clements and Goldsmith (1980), Mbah (1996) posits that "TLLs are characterised in terms of down-step of its converse "↑", the upset". He illustrates the above contention with the associative construction in Igbo in which if the initial word ends in any tone other than a high tone, its inherent tone must change but it never changes as a second word in an associative construction. For example:

38 (a) àlà Ìgbò → àla Ìgbò

land Igbo land of Igbo

(b) egō m → ego m

Money I money of me (my money)

Contributing to the discussion on downdrift and downstep, Ikekeonwu (1993:96) in her experimental study on intonation and focus opines that downstep is "an intonation feature arising as a result of the need for focus or emphasis on particular syllables within what we describe here as a Tonal intonation Group (TIG)". This claim agrees with that of Maidment (1990) which states that stress in English (intonation language) also serves in some cases as a prosodic feature on syllables which is introduced for emphasis or focus. Ikekeonwu (1993:102) continues:

> The basic element for the reflection of emphasis or focus in Igbo is the high pitch. This is reflected on one of the syllables of the TIG. It is generally on one of the syllables of the lexical item being being focused on, or on the syllable immediately to the right if this syllable has a high tone.

From the above, it can be observed that tone and intonation languages share this feature of stressing for focus or emphasis.

In her study on tone in intonation using Esan language, Ejele (2003:435) asserts that intonation and tone occur in languages. First, she says: "Pitch used distinctively in words is called tone". She, however, warns, "This does not in any way suggest that tone and intonation are mutually exclusive". She tries to argue the different circumstances that ascribe each prosodic occurrence of pitch its name. Citing Hyman's (1975:227-228) comparison of intonation and tone, Ejele (2003) says that (Hyman) used the two related notions of downdrift and downstep to explain the issue of relationship between intonation and tone. According to her, Hyman describes downdrift as "an automatic lowering process which in tone languages is common among those with two contrastive tones, high and low". Also citing Crystal (1997) Ejele notes that downdrift is a property of intonation, involving a sequential process whereby high tones after low tones become progressively less high throughout an intonation unit. She states that the downstep is a non-automatic lowering process that applies to the second of two high tone syllables in such a way that a downstepped high tone would be slightly lower than the preceding high tones but not so low as to be equivalent to a low tone.

It is in line with the above assertion that Ejele (2003:435) affirms: “It is widely but incorrectly believed that intonation is not normally associated with tone languages”. She makes this assertion following her study of Ekpoma dialect of the Esan language. To her, intonation is “a prosodic/suprasegmental feature which can be super-imposed on tones in a word, phrase, clause or sentence in tone languages, just as it can be super-imposed on stressed/unstressed syllables in stress-timed languages …” In other words, intonation can apply to tone languages to achieve various purposes. Such purposes are enumerated by Ejele (2003:435) where she says: “Intonation serves various purposes, including marking grammatical boundaries such as phrases and clauses, signalling sentence types such as statements and questions, and conveying all sorts of speaker’s attitudes such as the expression of joy, surprise, exasperation, irony, etc.”. She posits that it has two contrastive tones – high and low, both of which form rising and falling tones.

According to her, lowering high tones or vice versa is a process commonly associated with intonation. But she asserts that the process is also common among tone languages that have two contrastive tones, high and low. She then illustrates some of the intonation patterns found in Esan (a tone language) with the following examples:

(a) Intonation in statements: Ejele observes that Esan can be said to have neutral intonation in statement because, according to her “since there is no consistent pattern associated with them (Ejele 2003:436). Therefore, statements in Esan “can end on a low, high, falling or rising pitch”. Here, Ejele gives three varieties thus –

(i) Statements with final words ending on a low tone. For example:

39 (1) Intonational pattern

ọnù òmò mè bhúnù
the soap good Loc. mouth

‘The soup tastes delicious’ Ọ no mo mẹ bhu nu

(2) Intonational pattern

Ẹ̀fẹ̀ lamè ọ̀ghèdè
L L̇

'Efe fried the plantain' Ẹ fe la mọ ghẹ dẹ

In the above examples, she argues that it is the tone on the last syllable that determines the pitch on which the sentence ends. Therefore, all the sentences end on a low pitch, and as such observe downdrift where a following high tone is lower than a preceding high.

(ii) Statements with final words ending on a high tone. Here, she gives the following examples:

40 (1) Intonational pattern

ọni okhuò dẹ emì èso
the woman bought things for

'The woman bought some things' Ọ no khu o de mi e so

(2) Intonation pattern

Ẹ̀fẽ̀ kpe ukì ọ̀gọ̀
wash top bottle

'Efe washed the top of the bottle' Ẹ fẽ kpu kọ gọ

In the above examples, Ejele observes that the high tone in each case is realised as a lowered high. This development, according to her, is associated with intonation. Automatic lowering, otherwise called downdrift, again she affirms, is a feature of intonation which as obtained in the examples above involves a sequential process whereby high tones after low tones become progressively less high throughout an intonation unit.

(iii) Statements with final words ending on a falling or rising tone

Ejele describes these types of statements as where "the sentences end on HL and LH pitch respectively as dictated by the sentence-final words which end with HL and LH tones" (Ejele, 2003:437). She, however, observes that words ending in a falling or rising pitch are not common in the Esan language; hence, instances are very few. Examples:

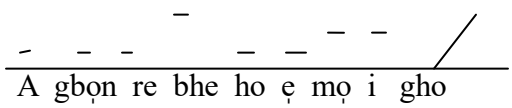
41 (1) Intonational patterns

ọnì ọ̀dùẹ̀kì sàa bọ̀ ùwâ
the trader atlast build house

'The trader is at last building a house' ọ no dwẹ ki sa a bu wa

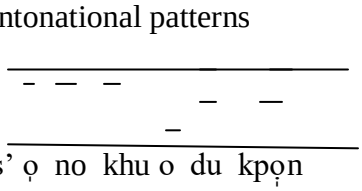
(2) Àgbọ̀n rèbhe hò ẹ̀mọ̀ ighŏ

people all love/like matter money
'Everybody loves money'

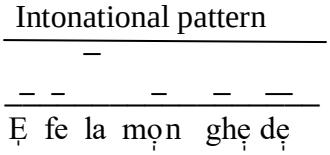


(b) Intonation in emphatic statements: In her discussion of this phenomenon in Esan language, Ejele (2003) goes on to say that "When statements are used with emphasis, differences show up, depending on the sentence – final tone". Thus, when statements with final low tone are used for emphasis, the low tone is raised, making it sound auditorily like a phonetic mid tone. For example:

42 (1) Intonational patterns
ọni okhuò dẹ u↑kpọ̀n
the woman bought cloth
'The woman did buy clothes' ọ no khu o du kpọn



(2) Intonational pattern
Ẹ̀fè la↑mẹ̀ ọ̀ghẹ̀dẹ̀
fried plantain
'Efe did fry plantain' Ẹ fe la mọn ghẹ dẹ

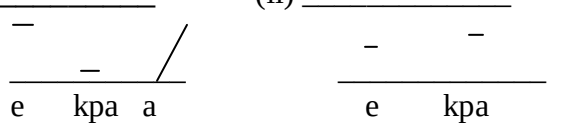


(c) Intonation in questions: Ejele (2003) explains that in Esan, words, phrases and sentences can be questioned by introducing the LH tones as Q-marker which combine with the last tone of the string to form contour tones, thereby necessitating vowel lengthening. She asserts, "Thus sentences can be questioned simply by raising the sentence-final tone through the introduction of the LH Q-marker"(Ejele, 2003:440). She illustrates this with such examples as these ones below:

Words Ouestions Intonational patterns

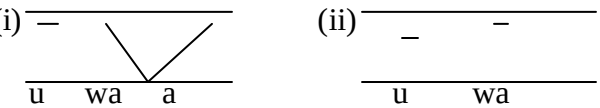
43 (i) (ii)

(1) ekpà (i) ekpàă
'vomit' (ii) ekpấ
'vomit?'



(i) (ii)

(2) uwâ (i) uwâă
'house' (ii) uwấ
'house?'



The above examples, according to her show that two forms of tonal marking of questions result when the LH tones interact with the last tone of the word as expressed in (43)(1) (i) and (ii). She explains that examples (43)(2) (i) involve contours, while (ii) involve the extra tone which is represented orthographically by a double acute on the particular vowel. However, it would be observed from the examples that whether the word ends on a low, high, falling or rising tone, the question ends on a high or extra high pitch. Secondly, it would be noticed that there are instances of elision in all the examples. Although the author does not explain why, it can be conjectured that it is a ploy for rhythm which also is a feature of intonation.

Chang (1972) conducted an investigative study of the relationship between intonation and tone using Chengtu, a dialect of Chinese. Chang describes intonation as "the fluctuation of the voice pitch as applied to the whole sentence. It is the sentence melody and is superimposed on the sentence as a whole" (Chang, 1972:401). Continuing, he says "… tones apply to individual syllables whereas intonation covers the whole sentence". Furthermore, Chang asserts: "a change of intonation does not affect the lexical values of words" rather "it only adds to shades of meaning to the sentence spoken and brings out the attitude of the speaker and the emotional state he is in". He goes further to observe that every community has its own intonation, that is, its own rules of changing the voice pitch when uttering the sentence; and the fluctuation of the voice pitch of the individual follows consciously as well as unconsciously, the intonation pattern. It then follows from Chang's (1972) observation that any individual whose intonation does not coincide with these patterns is considered a foreign speaker. In the same vein, any individual who is not familiar with these patterns expectedly misses the 'overtones' of the sentence spoken.

Speaking on the relationship of intonation to tones, Chang opines that ordinarily, one would imagine the pitch of each syllable in a tonal language to be fixed before hand, and therefore that it would be difficult for a tonal language to have intonation. But he contends that a closer examination reveals pitch phenomena that can only be regarded as intonation superimposed upon the tonal

system. He then points out that apart from pitch change due to the tonal environment, there still remain characteristics and modulations of the voice that bring out different shades of meaning. Therefore, he says: “The fact is that the sentence may be spoken in different ‘keys’ when representing different attitudes, and that the syllables go through perturbation …, thus giving the whole sentence a rising and falling tune” (Chang, 1972:401). By perturbation or tone sandhi Chang, means the replacement of one tone by another, that is, the interchange of allotones.

What Chang imputes is that though tones appear fixed in a tone language, it is not always so as some kind of modulation of the voice affects the tonal pattern in speech. These modulations are not only stylistic, but also grammatical as they suggest different meanings intended by the speaker. Hence, in his study, Chang advances series of illustrations using different types of statements like ordinary statements, ordinary questions, emphatic sentences expressing awe, sentences expressing contempt, etc. and then concludes that “intonation is indicated by the perturbation of one syllable which in the case of Chengtu dialect is the final syllable” (Chang, 1972:413).

Finally, in the findings of his study, Chang says that tones pronounced in isolation behave differently from the ones pronounced in connected speech. In connected speech, tones go through perturbation, and this is usually governed by the position they occupy in the phrase, or by the tonal environment. Chang’s (1972) study establishes the point that tone and intonation correlate, although Chengtu dialect is used in the study. All the same, Chengtu is a tone language like Igbo. Therefore, the findings of the study greatly support the assumptions of this research.

Oyebade (2006) accounts for the relationship between intonation and Yoruba, a tone language in his analysis of the speech act of a Yoruba speaker of English loanwords in Yoruba. In his discussion of the paper on preservation of loanwords from English, he says that in Yoruba language, “epenthesis is introduced to break up consonant clusters or to open closed syllables”. He cites examples from vowel elongation of words loaned from English into Yoruba, such as

44. [kilaàsì] → ‘class’

[dʒoònù] → ‘John’

[aloòmù] → ‘alum’

In an explanation for this style, Oyebade says that Bamgbose argued that lengthening was an attempt by Yoruba language to stay faithful to the falling intonation pattern of English stressed syllables. He continues: “since Yoruba disallows glide tone on short vowels, it adopts the strategy of lengthening to accommodate the gliding pitch of English stress” (Oyebade, 2006:2). This, in effect is to say that there is intonation pattern in Yoruba, a tone language.

Oyebade supports the above phenomenon with the contribution of Kenstowicz (2004) which he summarises thus: (1) stressed syllables of English source have their pitch adapted with Yoruba high tone, and (2) final syllable of English source is adapted with Yoruba low tone. He contends that this is so because as every word in English invariably has one (main) stress, a monosyllabic word is adapted with Yoruba high pitch. Again, he says that since the monosyllabic word is just one tone, Yoruba assumes that it is simultaneously an initial and final syllable, hence the assignment of low pitch.

Oyebade (2006:3-5) illustrates the different variations obtainable in words loaned from English into Yoruba with the following examples:

(a) Monosyllabic words

45 (i) With simple onset

English		Yoruba	
[paund]	→	[kpɔũ]	‘pound’
[fəuld]	→	[foòlù]	‘fold’

(ii) With complex onset

[brʌʃ]	→	[burɔɔʃì]	‘brush’
[drɔ]	→	[dùrɔɔ]	‘draw’

(i) Disyllabic Loanwords

46 (i) With final stress

English		Yoruba	
[ri'si:t]	→	[rìsiìtì]	'receipt'
[di'lei]	→	[dìleè]	'delay'

47. With initial word stress and open final syllable

English		Yoruba	
['titʃə]	→	[tiʃà]	'teacher'
['brɔðə]	→	[burɔdà / bùrɔda]	'brother'

48. With initial word stress and closed final syllable

English		Yoruba	
['pailət]	→	[kpailɔɔtù]	'pilot'
['tæblət]	→	[tabulɛɛtì]	'tablet'

(c) Polysyllabic Loanwords (with (ante) penultimate stress)

49 (i) Trisyllabic

English		Yoruba	
['eirəplein]	→	[ɛropuleèn]	'aeroplane'
[mə'kænik]	→	[mɔkaniìk]	'mechanic'

The above illustrations display the systematic variations noticeable in the source (English) and target (Yoruba) languages as identified by Oyebade (2006). The notable variations presented by Oyebode here include that "monosyllabic source lexical items entering into the target language with a falling pitch translate this pitch into a high HL tonal sequence in Yoruba" (Oyebade 2006:5).

Furthermore, “Epenthetic vowels bear a largely predictable tone, L when inserted after the ‘fall’ part of a gliding tone and H before the ‘rise’ part of the gliding tone”.

Another interesting exposition of Oyebade’s paper is the correlation of the English stressed syllables and the Yoruba H (high) tones. This phenomenon provides answers to the main question of this research. The presence of fall-rise, and rise-fall in Yoruba as revealed in example (40-45) support the claim earlier made by Chang (1972), Okorji (2002), Ejele (2003), Clark, *et al.* (2007), and others that there is intonation in tone languages. The account of Oyebode’s (2006) study also explains the peculiar Yoruba accent while speaking English, an intonation language.

All the arguments and discussions adduced above have gone a long way in throwing light on the state-of-the-art with regard to the nature of prosodies, on the one hand, and on the effect of prosodies on speech, on the other hand. Similarly, the submissions of notable scholars have helped to provide the much needed explanations to phonetic-phonemic and suprasegmental analysis and interpretation of speech. Furthermore, tone and intonation languages, though seen and presented as different types of languages, have been observed to share some common prosodic features. Therefore, although the feature stress before now appeared to be an exclusive attribute of intonation languages, it has been revealed through these studies that stress reflections also occur in tone languages. This assertion is fully substantiated by the cited authorities and this
has helped in charting a new approach to the analysis of suprasegmentals.

(vii) The prosodic features in languages: transfer and interference

Westerman and Ward (1990:133) observe that problems exist for speakers of a foreign language. Hence, they say: “Students with few exceptions do not pick a correct intonation naturally; their normal tendency is to use the intonation of the mother tongue in learning a new language”. They go on to say that the students are apt to transfer into the new languages the habits of their own

and they will continue to do so unless they are trained in the new intonation habits. In line with this thought, Wells (2006:12) postulates: "As in other areas of foreign or second language learning, learners of English will tend to start by assuming that English is like their own first language. They will **transfer** the intonation habits of the L1 to the L2".

He observes that transfer can be positive or negative. The positive transfer, according to him, occurs when the learner's own language (L1) and the target language (L2) share both universal and non-universal elements. A case similar to that is German and Dutch which he says have the same tonicity with English. He illustrates this with the example below:

50. *English:* I've 'lost my 'bag.
German: Ich habe meine 'Tasche verloren.
I have my bag lost.

According to Wells, the ordinary intonation pattern for the sentence involves a nucleus on *bag* which is both the last word and more importantly the last lexical noun phrase as *Tasche* (bag). He notes, however, that the different word order in the two languages is the reason for the lexical noun phrase not occupying the same (last) position. On the other hand, Wells cites French as an example of a language which does not use tonicity in the same way as English; and therefore concludes: "French learners have difficulties with English tonicity because of their **negative transfer** of the French system to English". He further observes that speakers of languages that make little or no use of tonicity are likely to make many inappropriate tonicity choices in English. Any such wrong accentuation often sounds absurd or bizarre to the listener as it sends out wrong signals about the end of the new information in the message, or about what the focus is or is not. His illustration of negative transfer is as follows:

51 (a) *English:* In 'most cases | 'standards have greatly im'proved.
French: Dans la plupart des 'cas | …
French English: In most 'cases | …

(b) *English:* I can 'see someone.
French: Je vois quel'qu'un.
French English: I can see some'one.

As Collins and Mees (2003:187) earlier observed, “Errors made by language learners frequently reflect the sound systems of their L1”. They opine that if the sound system of one’s L1 is compared with that of the L2, it can easily be conjectured the nature of errors that one will make. That is the same point being made by Wells (2006:12) where he says: “the assumption that English is like your L1 thus leads to **interference** from the L1 as inappropriate elements are transferred”. In like manner, Welmers (1973) highlights the fact that African languages south of sahara are tone languages, minus a few. According to Welmers, tone is so important in these languages to the extent that non application or wrong application of tone renders the languages meaningless or uninteresting. Therefore he declares:

> To the native speaker of a tone language, tone is just as basic a part of his speech as consonants and vowels; if no progress is made in that aspect of pronunciation at first, here is no reason to expect that any improvement will be made later. (Welmers, 1973:78).

It suggests, therefore, that a tone language speaker engaged in speaking English is expected to either suffer negative transfer, or interference. However, a situation where the two language types share some universal elements as observed by Wells (2006) and which is the case in the two language types of study here, we expect positive transfer of such elements by the tone language speaker of English, even though interference cannot be ruled out either. All that is being said is that good performance in any L2 requires acquisition of relevant phonetic knowledge of the L2 plus identifying the universal elements which the L1 and L2 share. This will act as a check for negative transfer or interference.

2.2.2 The Acoustics of speech production

Speech as sound results from vibration of vocal chords. Vibration itself is a factor of some source of energy that generates it. But for any sound to be audible, Clark, *et al* (2007:205) note three criteria that must be satisfied, viz. “propagating medium”, “frequency” and “amplitude”. The

propagating medium, such as air, metal, liquid is what the sound will travel through. The medium provides a link between the source and destination of sound. If there is no lead, no sound can be heard as sound does not travel in a vacuum.

Besides the medium is the property of sound that is relative to the sensitivity of sound to the ear. This property is frequency of vibration which ranges from very rapid to very slow. As explained by the trio, "The ear detects only a certain range of these frequencies, commonly down to about 20 vibrations per second and up to about 20,000 vibrations, per second…" Ashby and Maidment (2005) describe frequency as the number of repetitions (or cycles) per second which was formerly expressed in cycles per second, (c.p.s or c/s), but presently, the unit used is Hertz (Hz), hence, 1Hz = 1 c.p.s. The range of vibration differs from individual to individual, and is affected by age. Clark, *et al* in (5) above give the *Fo* range for English speakers which they observe were challenged by different scholars as Peterson and Barney (1952). They further observe that inconsistency in phonation from cycle to cycle to cycle may have some effect on vocal quality. Hence, they say, "all speakers seem to exhibit some inconsistency in duration from cycle to cycle of phonation" (Clark, *et al* 2007:235). This inconsistency gives rise to what they call "pitch jitter" which is at its highest at the start of phonation before a voiceless consonant, and then decreases in the syllable peak. Fundamental frequency, therefore, is the frequency of vibration of the larynx in phonation which is measured from the speech waveform.

The waveforms of voiced and voiceless words are different. Ashby and Maidment (2005) explain that when the pattern of a wave repeats regularly in time, the wave is regarded as a "periodic wave". According to them, a period runs from one clearly identifiable point on the wave to the next place where the point occurs. Therefore, "one period of a simple (sine) wave contains one upwards-and-over excursion, and one downwards-and-up-again excursion, returning to the zero line. The length of one period is the periodic time, *T"* (Ashby and Maidment, 2005:28).

Apart from frequency, another phonetically important property of sound waves is duration. Durations being considered may be as small as a fraction of one cycle of a periodic waveform, or may be one complete period of vibration, or may even be far longer. All this depends on the nature of utterance being considered. Hence, Clark *et al.* (2007:223) state, "in some instances, we want to determine the duration of a whole word or utterance, or even the duration of a silence such as may occur in the closure of voiceless stop".

Ashby and Maidment (2005:127) contend, "…the duration of a particular sort of segment is not fixed. It will vary depending on the context in which the segment appears". As they illustrate, the last sound or two at the end of an utterance are generally prolonged. This effect they call "pre-pausal lengthening". Also, a syllable-initial consonant is known to be generally longer when alone, but considerably shorter when preceding another consonant in a cluster. Still illustrating the effect of context on duration of a sound, Ashby and Maidment further posit that the duration of a vowel may be influenced by whether the following consonant is voiced or voiceless. They observe this phenomenon in most English accents and refer to it as "pre-fortis clipping". Other factors that influence segment duration that they observe include "overall speech rate, and degree of stress placed on syllable".

The duration of the utterance is measured using reference markers on the waveform that have meaningful relationship to the phonetic structure that is being measured. This often involves displaying the waveform on a computer screen using speech editing and analysis package. Clark *et al* (2007) argue that this process does not really yield a reliable and consistent means of measuring larger durations and determining the appropriate threshold of intensity that mark the start and the end of the speech to be measured. Hence they say, that a combined process of "time and frequency domain information" was advocated.

The third criterion that needs to be satisfied for a sound to be audible is amplitude. The amplitude of a wave, according to Ashby and Maidment (2005) is a measure of the size of the

pressure variations (or eardrum movements). Variation in air pressure results to loudness of a sound, and they describe loudness as the auditory property that is correlated with amplitude. Hence, Ladefoged (2003) says that loudness of a sound can fairly well be determined by reference to intensity (a measure of acoustic energy) which he says, is its acoustic counterpart. In other words, Ladefoged is imputing that intensity depends on the amplitude of the sound wave. Therefore, to measure the intensity of a sound requires taking the amplitude of the waveform at each moment in time during a window, squaring same and finding the mean of all the points in the window, and then taking the square root of the mean. So, he concludes, "The power of a sound is the square of this mean" (Ladefoged, 2003:90). The implication of the aforesaid is that large movement of the source of sound produces a loud sound which invariably necessitates exertion of more energy during production of the sound.

The amplitude of the sound refers to the strength of each peak of pressure; the rate at which the peak (of pressure) occurs refers to pitch. The difference between high and low-pitched sounds, according to Ladefoged (1962), is that the higher-pitched sound is making a greater number of variations per second than the lower-pitched sound. Consequently, Ladefoged (1962:18) states, "the variation in air pressure in any sound that has a definite pitch must form a pattern which is repeated at regular intervals".

For Pickett (1980), the air molecule motions of speech sounds can be described aptly by considering them to be made up of simple oscillation. He illustrates the motions of speech sounds with the motions of a pendulum and says,

> An ideal pendulum, one that has no friction, moves back and forth in simple harmonic motion. The pendulum motion is very regular in time, but still it involves many different speeds and positions. … this motion can be represented very simply by relating it to uniform motion on a circle. (Pickett, 1980:14)

Pickett's illustration above presents two-value parameters for the description of pendulum motion, thus, one giving the 'rate of motion', and another giving the 'size of the motion'. These two values can be translated for our purpose here to 'the time, or period', for one complete revolution, and 'the size of the motion'. A plot in time of the motion of ideal pendulum, he continues, gives a wave called "sine wave"; and its sinuous form is the form in time for all simple harmonic motions. He, (Pickett, 1980:16) goes on, "The period of a sine wave is the time for one complete cycle … related to the rate of oscillation, is often given as the *frequency of repetition* of the cycle in a unit of time". The frequency which he simply describes as "the reciprocal of the period" is given in cycles per second or Hertz (Hz). Also, "the distance or amount of sine wave motion is called the *amplitude*", which corresponds to the extent of the oscillation from the resting position.

For any sound wave, asserts Pickett, the extent or amplitude of the motion and its form describe the sound completely, in physical terms. Therefore, when a sound is received by the ear, the extent of air molecule motion "determines the loudness of the sound heard", while the form of the motion "determines the timbre or quality of the sound heard". Pickett (1980:18) therefore declares, "The heard quality and loudness of sound are very important to us and these are very neatly represented by means of the frequency and amplitude of simple harmonic motions".

Pickett (1980) earlier declares the "propagating medium" (see Clark *et al* 2007) for sound transmission to the ear as air, and describes sinusoidal sounds (the form in time for all simple harmonic motions) as pure tones. He, however, notes that vibrations (of sounds) can be affected by "damping" forces which make the vibrations to die out gradually in time; and he explains the forces of "damping" to include air friction. All these acoustic variables of sound, viz. medium, frequency and amplitude with their correlates contribute to the relevance of any sound, both in production and perception.

2.2.3 The Acoustic analysis of speech

Until very recently, analysis of speech perception had depended solely on ear perception. For instance, Harley (1995:42) notes,

> …we are rather good at recognising speech. The process is automatic; when you hear speech, you cannot make yourself not to understand it. Most of the time it happens effortlessly and with little apparent difficulty. Speech perception is fast.

Similarly, Westerman and Ward (1990:134) assert, "the best way to investigate tones is for someone with a trained ear and some experience to note down the tones of isolated words and sentences in some graphic fashion which is easy to interpret". Further arguing in favour of ear perception of speech, Laver (1994) observes that pitch is a perceptual concept. It is also argued that in linguistics, perceptual analysis precedes other approaches used to analyse speech before the actual linguistic elements that may require to be investigated acoustically or instrumentally (experimentally). Perceptual analysis has also been described to have the advantage of general accessibility and acceptability. Hence, Laver (1994) again asserts that the auditory system and sense of hearing are a resource common to both the speaker and the listener and equally makes it more convenient for a researcher to dare an investigation or a reinvestigation of earlier studies. A striking campaign in favour of speech analysis via ear perception is the availability of the findings of the study to a wider readership than that of an acoustic or instrumental analysis which, it claims, would first require a good knowledge of acoustic phonetics, availability of phonetic laboratory, ability to operate the equipment relevant in any such research and even availability of power supply.

Nonetheless, modern researches have stiffly challenged and possibly overthrown sole dependence on ear perception for speech analysis. Clark *et al* (2007:230) aptly state, "It is characteristic of human perception that the sensations we experience in response to stimuli rarely correspond directly with the values we derive from measurement of those stimuli". For instance, they observe that the human auditory system is capable of responding to an enormous range of sound

intensities with the upper end ranging more than a million times greater than the lowest perceivable intensity; and they opine,

> Not only does this lead to some very inconvenient numerical values, but, given the nature of perception, the figures do not relate very well to the perceptual effects of differences in intensity…The relation between perceived loudness and acoustic intensity is more nearly logarithmic…Hence the most convenient way to express intensity so that it relates to perceived loudness is as a logarithmic ratio,… (Clark, *et al* 2007:230).

Similarly, pitch, as the perceived period or frequency of a sound wave is determined largely by the fundamental frequency of the sound. But, the relationship between pitch and fundamental frequency is nonlinear and varies with the frequency involved. Besides, Clark, *et al* contend that our sensitivity to changes in the frequency of a sinusoidal tone, that is, our pitch discrimination, varies as we move up the audible frequency scale. To check this, they proclaim that a perceptual unit called the MEL was devised to represent equal increments of pitch and relate them to frequency. This is one great advantage of acoustic approach to speech analysis. Secondly, Clark, *et al* (2007) record that prior to the establishment of spectrograph in the 1940s, it was a very tedious task to embark on an acoustic analysis of speech. To dare this was also restricted by dearth of equipment. Hence, formant structure and its auditory qualities of speech sounds were very little explored in natural speech, they opinionated. Consequently, they declare, "Given the problems of providing a reliable auditory description of vowel quality …, the availability of an ostensibly objective technique of acoustic analysis, free from the bias of the human observer was an important step in phonetic and phonological description" (Clark, *et al* 2007:264).

Presently, sound waves are mainly handled in digitised form. Thus, Ashby and Maidment (2005:29) say,

> The tracks on a CD, or **wav files** in a computer, are simply long strings

> of numbers representing waveforms sampled at regular intervals. The **sampling rate** controls what frequencies will be preserved when the wave is reconstructed. Basically you have to sample at a rate that is at least twice the highest frequency you need to show.

Explaining the exercise further, they say that a CD works at a sample rate of more than 40 kHz, enabling it to provide a 'hi-fi' sound to 20 kHz or so. They report that in many of the waves they illustrated in their book they used slower sampling rates of 16 kHz or even 10 kHz which give them much smaller wav files.

Two major advantages of digital analysis of speech signals have been identified by Clark, *el al.* (2007:258). In their words,

> The first is that once the signal has been digitally encoded and stored, it can be edited, processed, measured, manipulated and filed with far efficiency than is possible with analogue instruments and an ordinary tape recorder. The second is that the analysis itself can be more easily varied to give optimum time- and frequency-resolution properties.

The contentions by these scholars have given great credence to the option of acoustic analysis of speech perception over mere ear perception. The advantages enumerated above equally necessitate the option of this approach for this study.

In deciding what to record requires basically acquiring the knowledge of phonetics and understanding the phonology of the language of study. When looking for sets of words, Ladefoged (2003:3) suggests, "…another major source is the speakers of the language themselves". This, according to him, is to confirm that the words exist. He also recommends the Digital Audio Tape (DAT) for recording of data. He further lists four properties to be looked for in a good recorder to include a good frequency response (the range of pitches the system can record); a good signal/noise

ratio (the range of loudness); reliability and user-friendliness; and the possibility of using the recorder for a long time. He claims that all these properties can be found in DAT.

2.3 Empirical Studies

Several studies have been conducted on tone and its relationship with intonation. Notable among such studies include Emenanjo (1978) and Ikekeonwu (1993) which study the nature of tone in Igbo and assert that Igbo has features of downstep and downdrift. They insinuate that these features behave like the intonation especially in connected speech. Hence, Emenanjo (1978:16) says, "When, therefore, in an utterance there is a series of alternate high and low tones, because of the automatic lowerings of these tones, the utterance progresses down like the levels of a terrace…" So, although he recognises this feature of intonation in Igbo, this has not adequately proved the relationship between tone and intonation, especially as it concerns frequency and duration. Similarly, Ikekeonwu contends that downstep and downdrift occur in Igbo utterances and that they display some characteristics of intonation. She therefore concludes that 'downdrift' and 'downstep' are aspects of declination. She asserts that downstep is bound to focus/emphasis where she states,

> We reanalyse the downstep as an intonation feature arising as a result of the need for focus or emphasis on particular syllables within what describe here as a Tonal Intonation Group (TIG). The TIG would generally correspond to an intonatinal phrase or sentence. (Ikekeonwu, 1993:93).

She demonstrates this claim with some Igbo utterances which she analyses their *Fo* contours. Other scholars like Mbah and Mbah (2010) also carried out studies in tone and tonal behaviour in Igbo. But their study does not extend to finding their relatedness to intonation nor does it use acoustic approach.

However, these works have actually made very insightful contribution in discussing tone and functions of tone in Igbo. Also, these earlier efforts were limited at tone and patterns of tone in Igbo. They do not explore the area of interface of tone and intonation; neither do they conduct acoustic investigation of the suprasegments.

Uguru (2000), on her own part made an attempt to investigate the interaction of tone and intonation. But she limits her study on Ika, a dialect of Igbo. The study focuses on differences in tonal occurrences and patterns of Ika and standard Igbo. She also investigates whether intonation in Ika bears attitudinal and discoursal meanings in the dialect. She further attempts to determine whether Ika dialect shares intonational or tonal features with accent languages. Her study reveals, among others, that Ika can neither be classified as predominantly tonal nor intonational considering the fact that it manifests many features of both. Also, the research discovers that features of tone and intonation interact to some degree. This is indeed a great step in the phonological study of Igbo language. But the study is only limited to a dialect of Igbo. Besides, no pure intonation language was used as control to the Ika (intonation) dialect in this study. Similarly, Uzochukwu (2001) makes a comparison between stress-timed and syllable-timed languages (with illustrations in Igbo), and thereby contributes to the contentious debate about the presence of stress in Igbo. Although Uzochukwu's work is on Igbo traditional funeral poetry, his contribution is no less an important one as it dwells on rhythm which is an important variable in considering the melody of speech. Nonetheless, this great stride in speech analysis has not provided the required answer for the question of point of relatedness between tone and intonation.

Uguru (2006) however, introduces acoustic analysis in her study of English and Ika intonation patterns. This work is geared towards proving the existence of intonation in an otherwise predominantly tone language. Again, it aims to carry out detailed investigation of the nature of intonation in Ika as well as compare some tune patterns already observed in Ika with similar tune patterns in English. This is, no doubt, advancement in the analysis of Igbo tone. It

has also provided a step further in phonetic studies by introducing acoustic analysis which will enhance the phonology of the language (Igbo) and its understanding by scholars.

All the same, the study is carried out with only one dialect of Igbo which speakers constitute a very small fraction of the entire Igbo race, and which has not fully developed as to provide wide spectrum of features of tone or intonation required to compare with such an international language as English. Also, the size of the population for the research is not representative enough for the speakers of Igbo language. This is perhaps because the study is limited to a dialect.

With these observations, therefore, it becomes necessary to investigate further on tone and tone patterns in standard Igbo (tone language) and English (intonation language). Furthermore, it is important to conduct a further acoustic study on tone in Igbo and intonation in English to confirm or dispute the claims by the earlier studies.

2.4 Summary

From the myriad of literatures reviewed, it has become clear that only very little has been written on African languages in general, and in Igbo language in particular. This has therefore left a yawning gap between the available studies in other world languages and those of African languages, especially Igbo language which is one of the languages of our study. Also, not much has been discussed on the interface of stress and tone, especially using native speakers of non-stressed languages. This has therefore led to dearth of linguistic data on the topic of discussion.

This study therefore hopes to contribute its own quota toward bridging the existing gap and then provide further materials for linguistic studies. It also appears to be a debut in the investigation of correlates of stress in the English language spoken by Igbo language (tone) speakers, using the acoustic analysis.

CHAPTER THREE

METHODOLOGY

This chapter presents the research design, the sources of data, the population as well as method of selecting the population and the size of the population. The instruments used in the study and method used to collect the data used are also presented. Finally, the method used in presenting the data and in analysing same is shown in this chapter.

3.1 Research Design

This research sets out to analyse speech patterns of native English speakers and Igbo speakers of English with a view to ascertaining the points of relatedness of tone and intonation. For accurate account of acoustic correlates of pitch patterns, the research relies on acoustic phonetics which is experimental analysis, thus confirming its relevance in speech synthesis. The experimental approach used involves the use of audio recorder for speech recording of selected consultants, and computer analysis of their speech.

3.2 Sources of Data

Since the research is set to investigate interface of tone and intonation in the speech of Igbo speakers of English, Igbo native speakers were selected. Also, Roach's and Wells' recordings (which present accurate reflection of intonation in English) comprise the English native speakers since they were the best available RP English.

3.3 Sample Technique and Size

The prospective sampling was used to select the respondents. This is because the researcher believes that to draw conclusions based on the sample to the study will produce the expected result of the interface of tone and intonation in the Igbo and English languages.

Also, considering the enormity of work to be done in processing the data to get the required result, and coupled with the fact that it may not be possible to work with all those that may meet with the required qualification, the researcher decided to use a sample population of a limited size. Consequently, eight consultants were used – four males and four females. Out of this number, the best four (two males and two females) that met the requirements of the purpose of the study were used. These consultants are people who speak Igbo as their native language as well as command high competence in the English language speech. They are predominantly teachers and scholars of language and linguistics

3.4 Instrumentation and Data Collection

To elicit the required information from the consultants first required gaining the confidence of the selected consultants and obtaining their permission to be subjected to the task of speaking for a considerable length of time, and also to tape-record their speech. When this was achieved, selected excerpts (words, phrases and sentences) were used. The essence of the diversification of text types was to provide wider spectrum for testing the research questions. The respondents were then asked to read the selected disyllabic words, two-word phrases and sentences.

To collect their speech, high powered Zoom digital recorder and a 4GB secure digital card were used. It was believed that live speeches of the respondents would be most useful for the purpose of the research. The recordings were done in a conducive and relaxed condition in order not to adversely affect the performance of the consultants. Furthermore, the recordings were, as much as possible, conducted in a noise-free environment so as to shut out noise from filtering into the speeches and therefore falsify the result.

For the RP English, the recordings of Roach (2000) and Wells (2006) were used. These serve as control group for the consultants.

3.5 Method of Data Presentation and Analysis

First, the data collected (recorded) with the Zoom digital recorder was transferred into an XP 2007 Windows Computer. The recordings were made at 16 kHz and saved as wav files in Praat and then subjected to pitch trackings and later analysed. The pitch variations on the speeches were interpreted acoustically based on the acoustic correlates of fundamental frequency and duration.

Secondly, the patterns of pitch variation that correlate in the tone and intonation languages of focus were identified and analysed. These are aimed at explaining whether stress is manifested in tone languages. If stress is manifested in tone languages, they will then provide its exponent(s) in the languages, as well as determine whether native speakers of tone languages encounter any problems in the articulation of intonation languages. Thereafter, conclusions were made based on the findings.

CHAPTER FOUR

PRESENTATION AND ANALYSIS OF DATA

In this chapter, the data collected are presented and analysed. This is done in compliance with the methodology earlier proposed in this study. The work is an instrumental analysis; hence, XP 2007 Windows computer and Praat speech analysis software package developed by Paul Boersma and David Weenink were used in the analysis.

4.1 Investigation of the acoustic correlates of intonation

Basically, the major acoustic correlates of intonation – the fundamental frequency and duration - are investigated in the data collected. These correlates present the pitch value and timing (duration) of the speech segments produced by the consultants which are in turn compared with the other consultants and then with the control. We then use the results to make categorical statements, which answer the research questions and justify the objectives upon which the study was based. The sound waves and acoustic correlates of the recorded utterances (words and sentences) are tagged Figures 1 to 16, while the spreadsheets of the results of the acoustic correlates of the intonation of the utterances are numbered Tables 1 to 4. These are followed by summaries/discussions of the analyses.

4.1.1 Stress in words (Lexical stress)

The sound waves and acoustic correlates of the utterances of word stress are presented in this section. The spreadsheet showing the result of the acoustic correlates of the utterances of word stress by the control group and consultants is also presented below as Table 1, and is then followed by the summary of the acoustic analysis.

4.1.1(a) Sound waves and acoustic correlates of utterances of word stress

In this sub section, three categories of word stress are investigated, viz. two syllable words, three-syllable words and compound words. Generally, the stressed syllable receives greater prominence and is as such perceived louder. Consequently, the syllable which receives less prominence is said to be unstressed. Secondly, whereas only one syllable receives greater prominence, the number of unstressed syllables depends on the number of syllables that constitute the word.

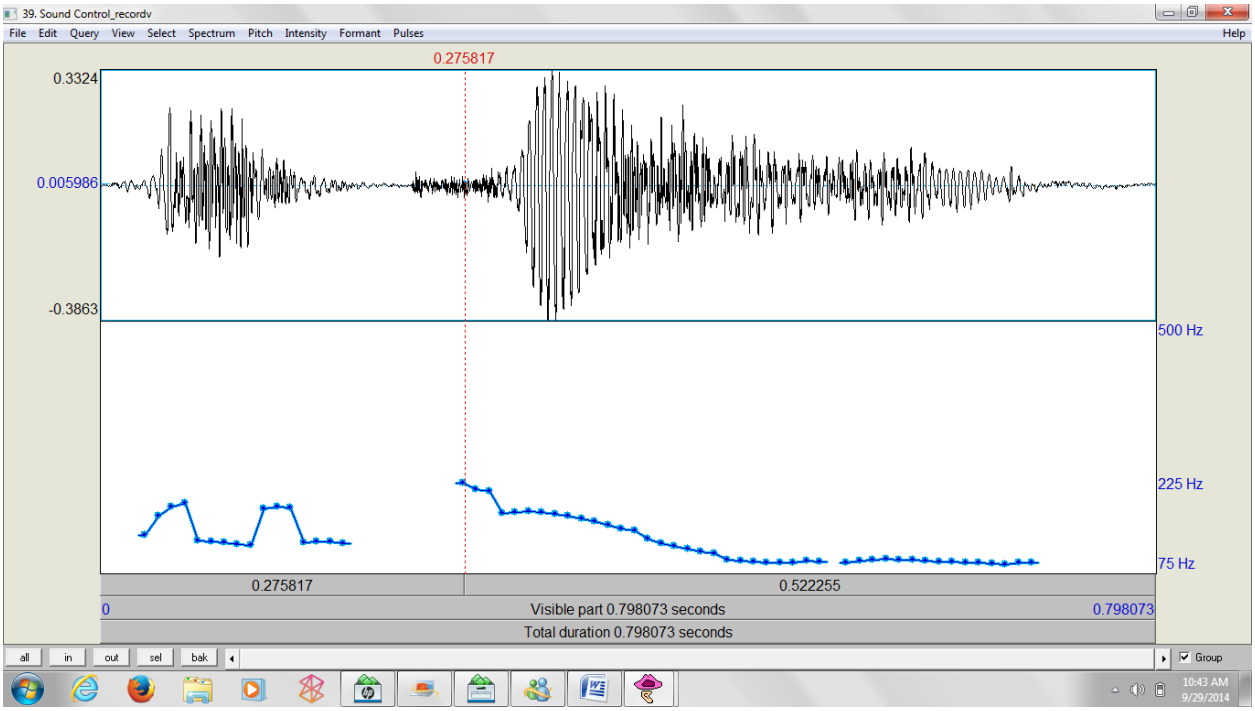


Figure 1(i): Sound wave and *Fo* and time duration of Control utterance of *record (v)* [rɪˈkɔ:d]

Record is a two-syllable word. The word is uttered by the Control native (British) English speaker at the periodic waveform cycles of 0.275817 seconds and 0.522255 seconds, thus giving a total duration of 0.798073 seconds. The (highest) pitch value within the pitch range of 75Hz and 500Hz is 225Hz. The stress is on the second syllable which is the heavy syllable, and this gives the pitch contour of the waveform activity as fall rise.

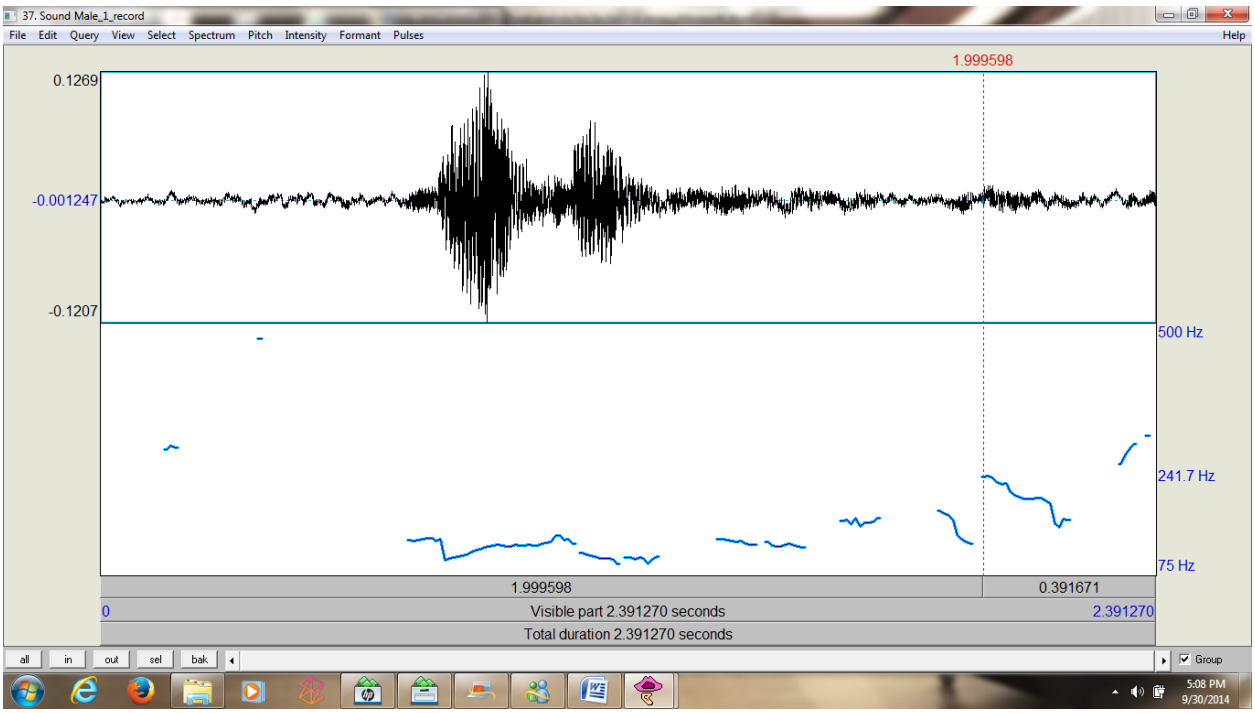


Figure 1(ii): Sound wave and *Fo* and time duration of Male 1 utterance of *record (v)* [rɪˈkɔ:d]

The Male 1 native Igbo speaker's utterance of *record (v)* was rendered at the periodic waveforms of 1.999598 seconds and 0.391671 seconds cycles, giving a total duration of 2.391270 seconds. The pitch value is 241.7Hz. The highest pitch is on the second syllable thereby producing the pitch contour as fall rise. But, the pitch contour portrays rather broken and wavy strings of utterance.

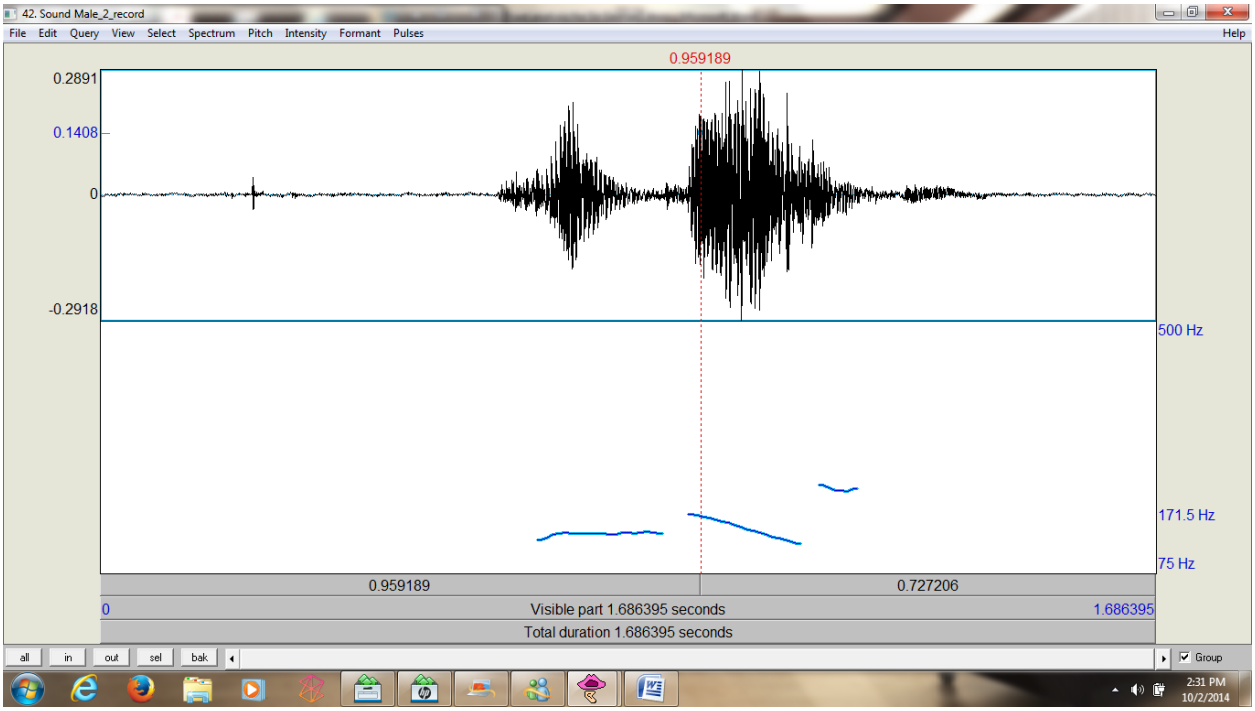


Figure 1(iii): Sound wave and *Fo* and time duration of Male 2 utterance of *record (v)* [rɪˈkɔ:d]

Male 2 native Igbo speaker's utterance of *record (v)* was rendered at the periodic waveform cycles of 0.959189 seconds and 0.727206 seconds. The total duration is 1.686395 seconds, and the pitch value is 171.5Hz. The pitch contour of the waveform activity also indicates a fall rise. Nonetheless, the waveform activity appears shorter than that of the Control.

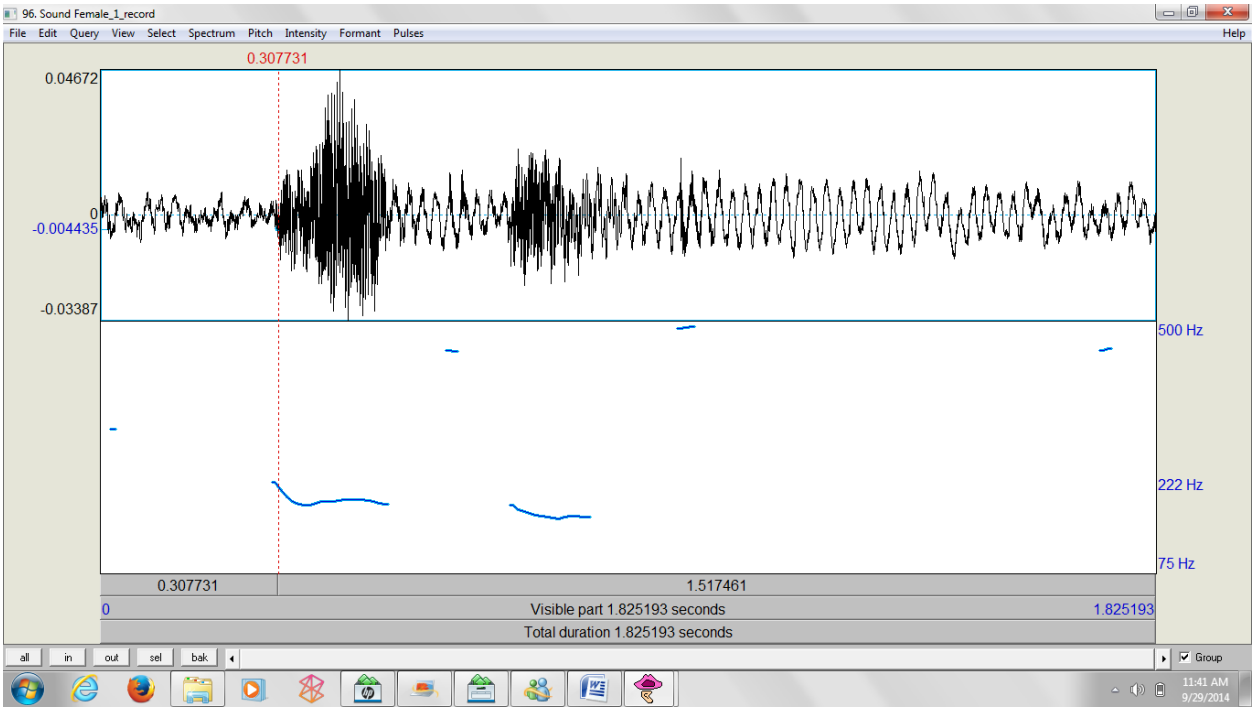


Figure 1(iv): Sound wave and *Fo* and time duration of Female 1 utterance of *record(v)*[rɪˈkɔ:d]

The utterance of Female 1 Igbo speaker was made at the periodic wave form cycles of 0.307731 secs.and 1.51746 secs. The total time duration is 1.825193 secs, and the pitch value is 222 Hz. Although the Fo contour and wave form activity appear short, the pitch contour presents the two-syllable structure of the word.

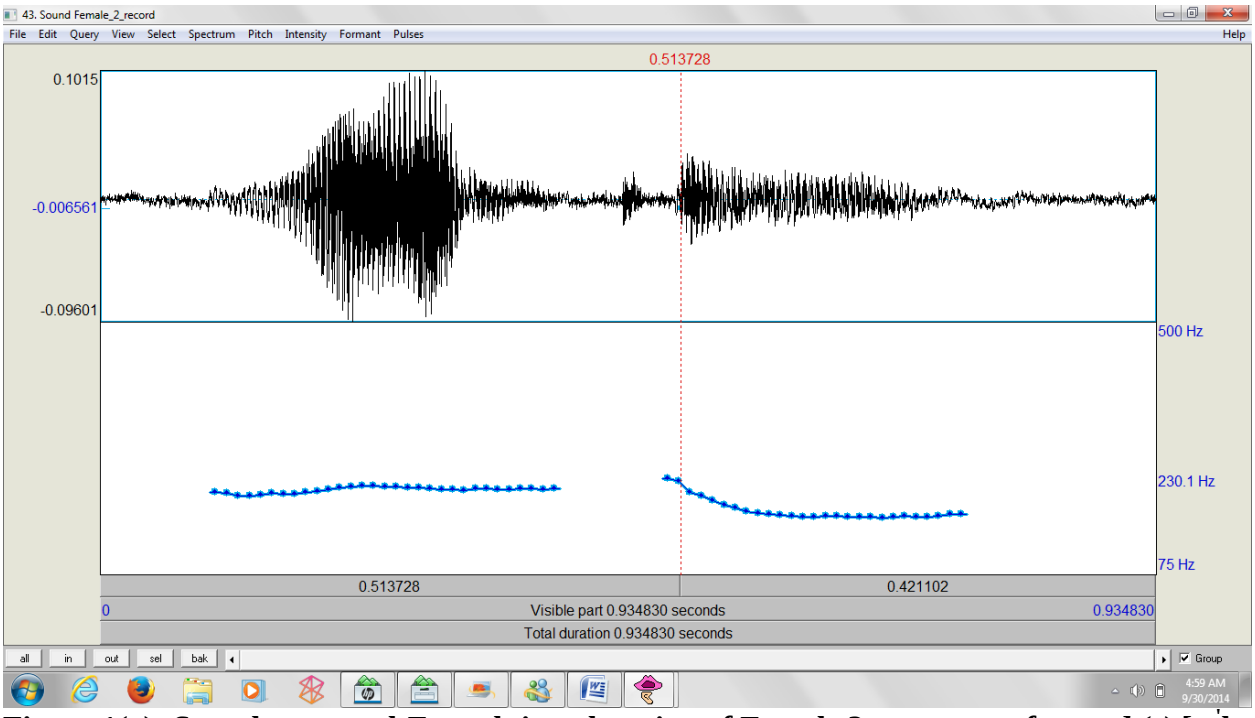


Figure 1(v): Sound wave and *Fo* and time duration of Female 2 utterance of *record (v)* [rɪˈkɔ:d]

The Female 2 native Igbo speaker's utterance of *record (v)* was rendered at the periodic wave form cycles of 0.513728 seconds, and 0.421102 seconds. The total time duration is 0.934830 seconds,

while the pitch value is 230.1 Hz. With the highest pitch value as 230 Hz and the declination that follows, the pitch contour can as well be described as fall rise.

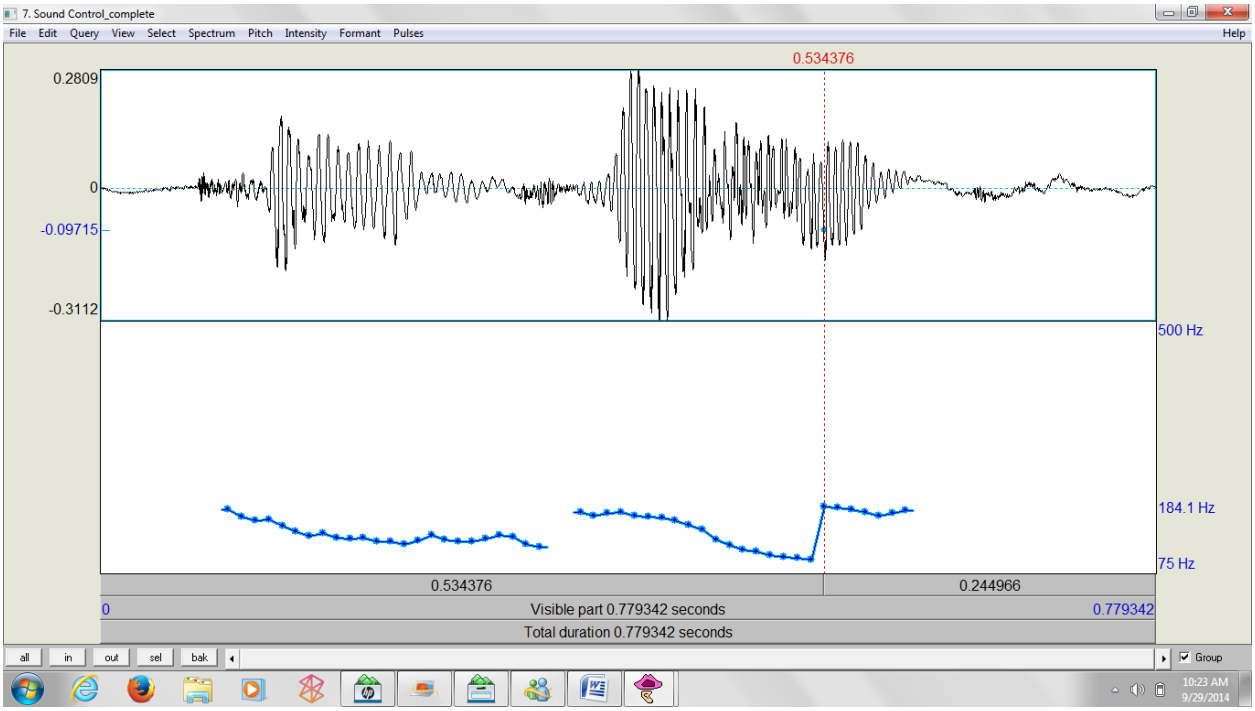


Figure 2(i) Sound wave and *Fo* and time duration of Control utterance of *complete (adj)* [kəmˈpli:t]

As an adjective, *complete* has its second syllable as the heavy syllable and is thus stressed. Hence, in the Control utterance of the word, the periodic wave form cycles are 0.534376 seconds, and 0.244966 seconds, giving total time duration of 0.779342 seconds. The highest pitch value is 184.1 Hz. The pitch contour portrays the two-syllable structure with rising pitch (primary stress) on the second syllable.

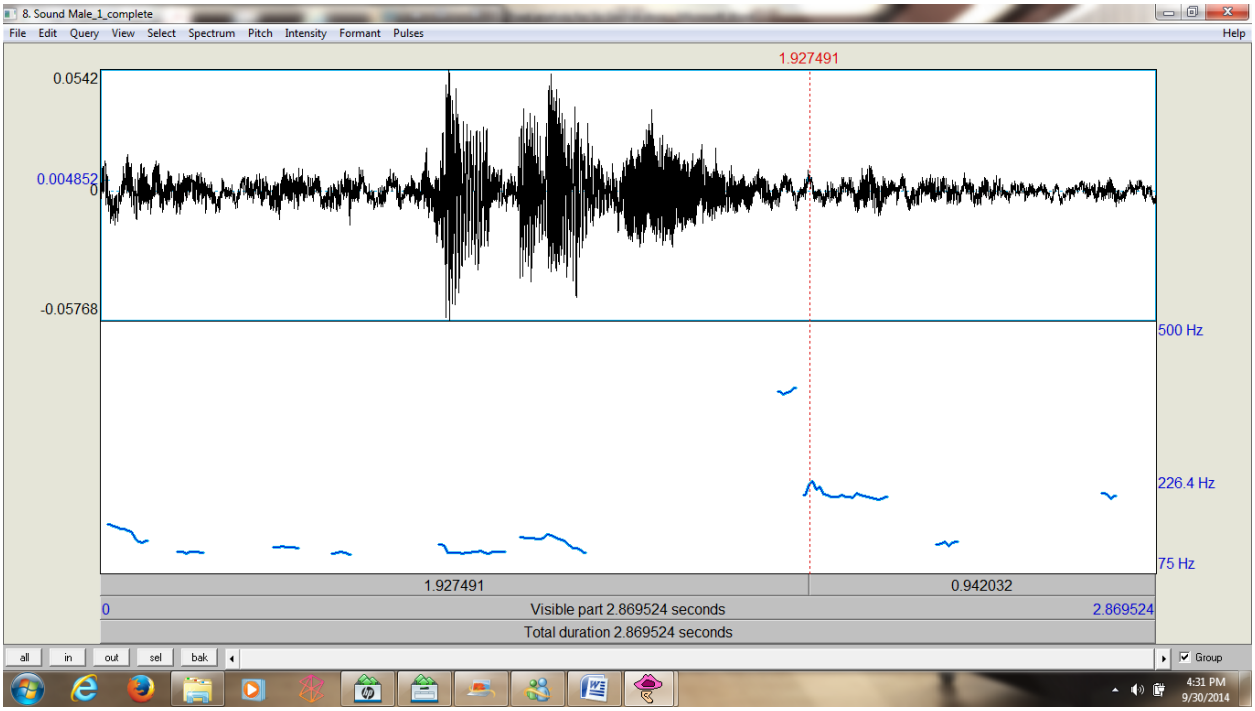


Figure 2(ii) Sound wave and *Fo* and time duration of Male 1 utterance of *complete (adj)* [kəmˈpli:t]

The Male 1 native Igbo speaker's utterance of *complete (adj)* shows the periodic wave form cycles as 1.927491 seconds and 0.942032 seconds, both giving the total time duration of 2.869524 seconds. The pitch value stands at 226.4 Hz. However, the Fo contour and wave form activity appear in broken strings, thus making the syllable structure unclear.

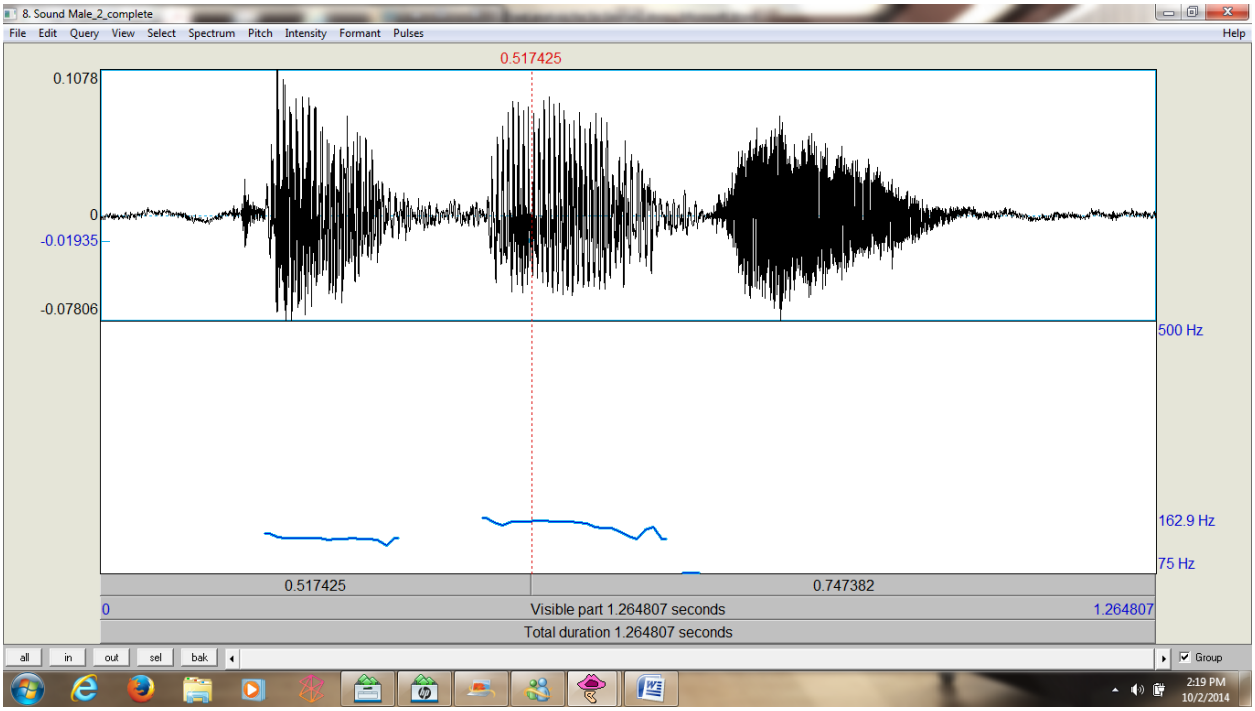


Figure 2(iii) Sound wave and *Fo* and time duration of Male 2 utterance of *complete (adj)* [kəmˈpli:t]

The Male 2 native Igbo speaker rendered the word *complete (adj)* in the periodic wave form cycles of 0.517425 secs and 0.747382 secs, both totalling 1.264807 secs of time duration. The pitch value measures 162.9 Hz. The pitch contour displays the two-syllable structure of the word, and the stress pattern of unstressed and stressed is clearly manifested.

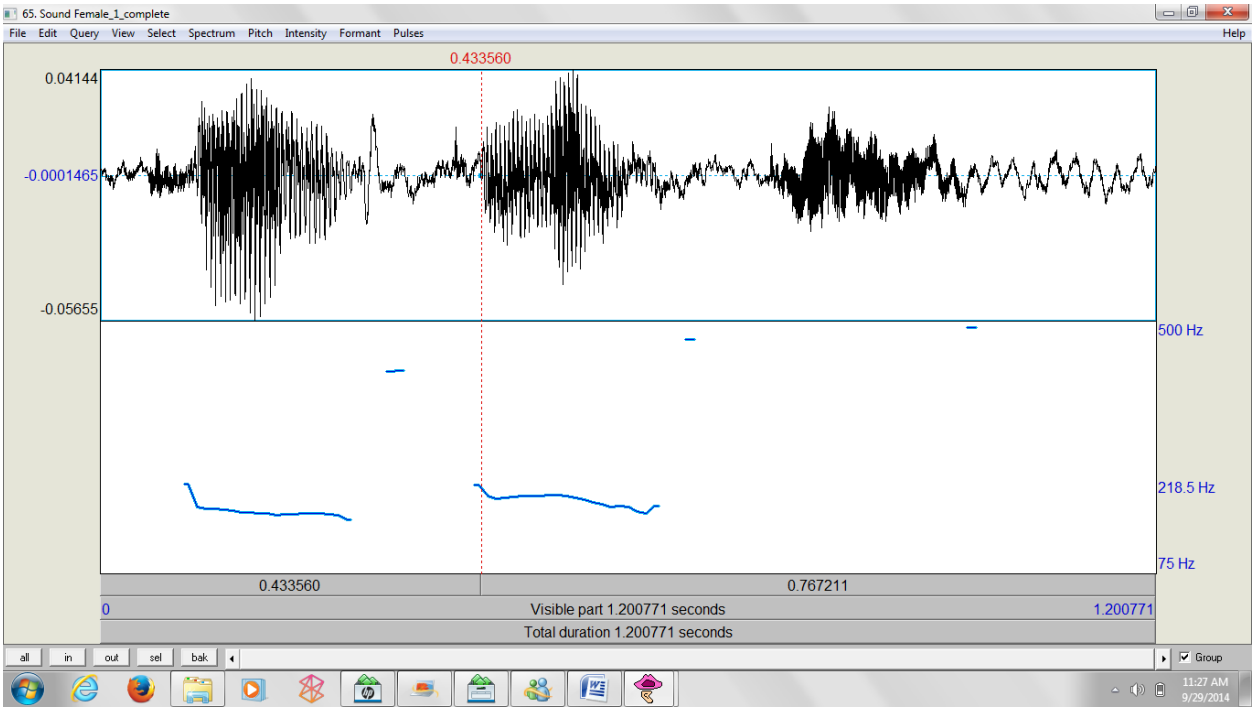


Figure 2(iv) Sound wave and *Fo* and time duration of Female 1 utterance of *complete (adj)* [kəmˈpli:t]

The Female 1 native Igbo speaker rendered *complete (adj)* at the periodic wave form cycles of 0.433560 secs. and 0.767211 secs. The total time duration is 1.200771 secs, while the pitch value measures 218.5 Hz. Again, the Fo contour indicates the two-syllable structure and stress pattern of the Control utterance of the word.

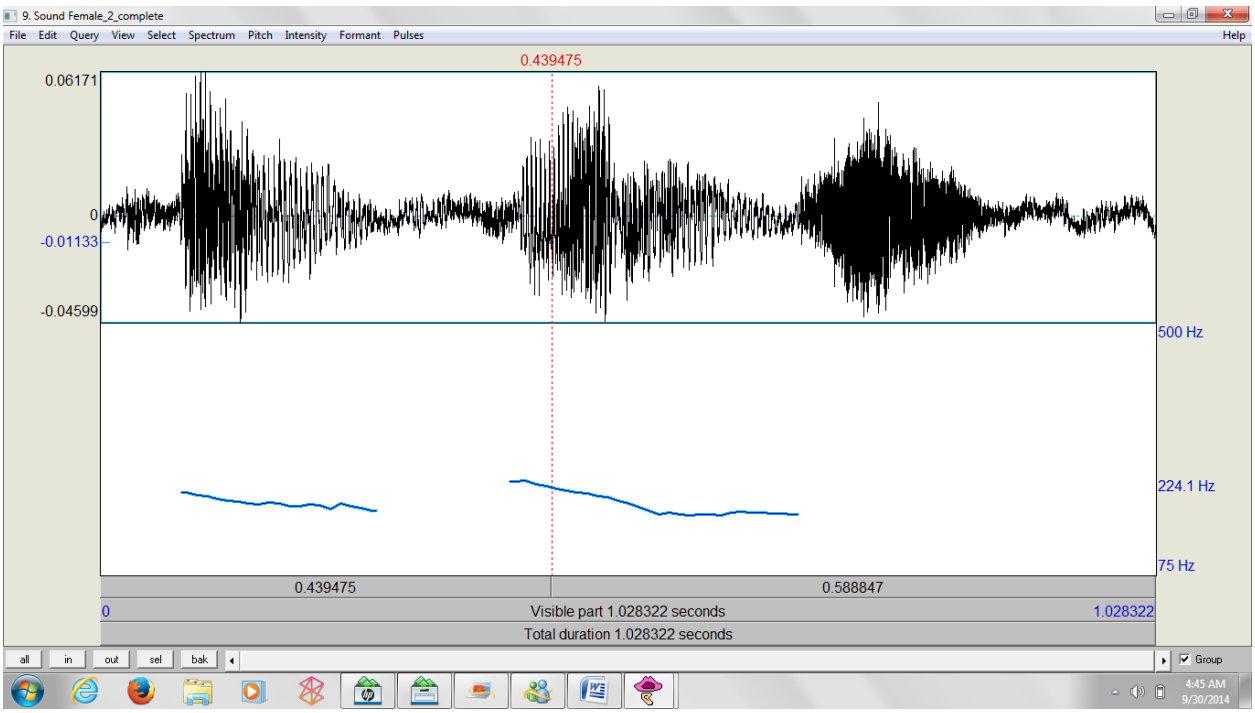


Figure 2(v) Sound wave and *Fo* and time duration of Female 2 utterance of *complete (adj)* [kəmˈpli:t]

The Female 2 native Igbo speaker uttered the word *complete (adj)* at the periodic cycles of 0.439475 secs and 0.588847 secs respectively. The total time duration stands at 1.028322 secs, and the pitch value is 224.1 Hz. The pitch contour clearly presents the two-syllable structure of the word and the stress pattern of the word which gives the first syllable as unstressed and the second stressed.

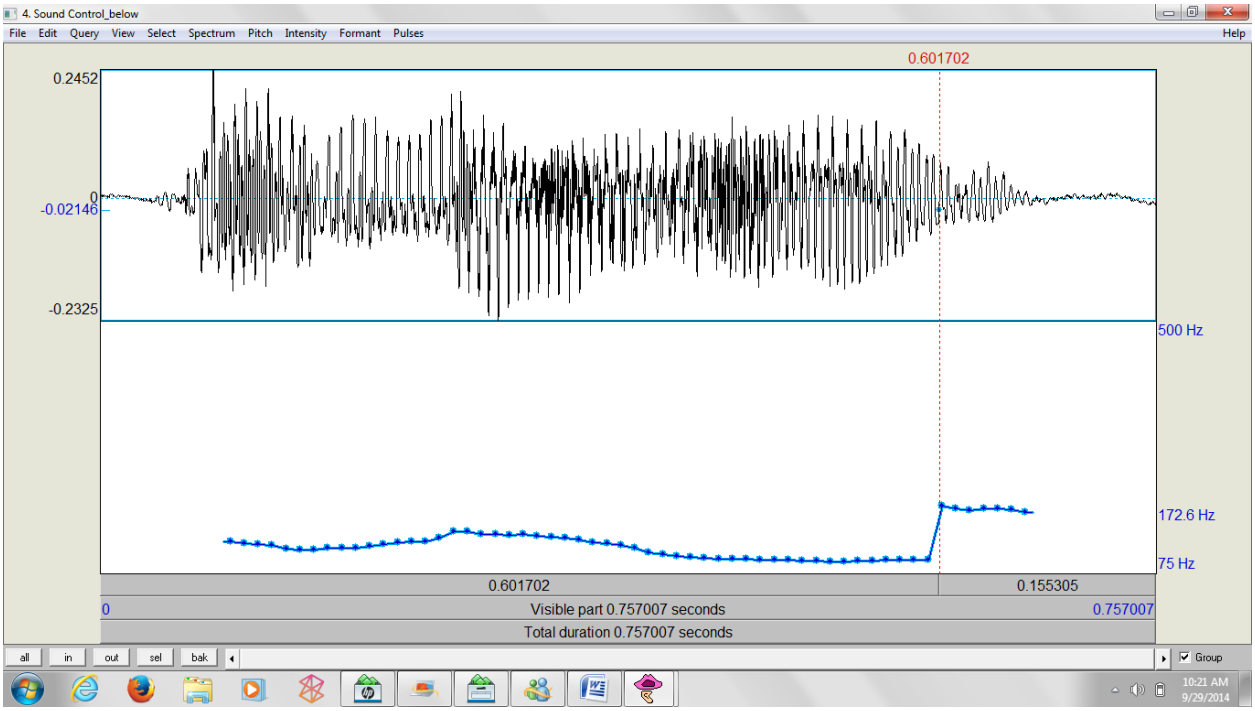
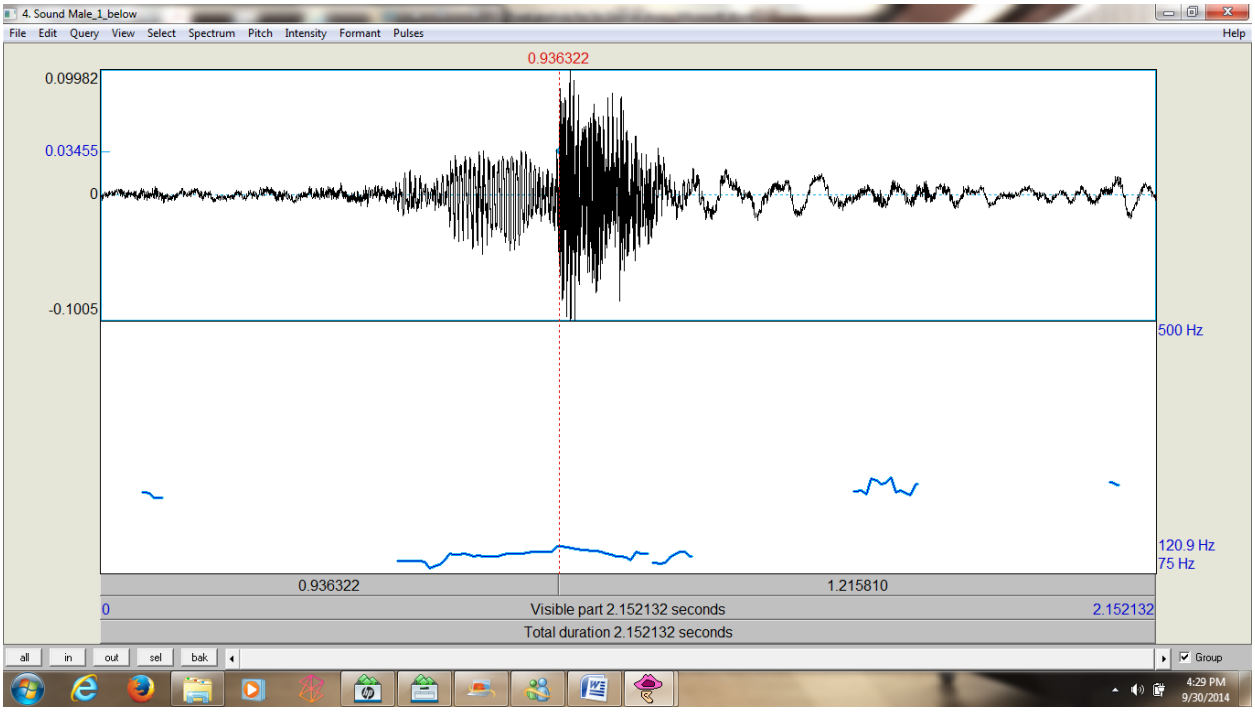


Figure 3(i) Sound wave and *Fo* and time duration of Control utterance of *below (adj)* [bɪˈləʊ]

Below, here, is an adjective and as such the second syllable is the heavy syllable and therefore bears the stress. The Control utterance of the word is made in wave form cycles of 0.601702 secs and 0.155305 secs. The total time duration for the production of the word is 0.757007 secs, and the highest pitch value is 172.6 Hz. The pitch contour of the utterance shows the second syllable as stressed.



Figiure 3(ii) Sound wave and *Fo* and time duration of Male 1 utterance of *below (adj)* [bɪˈləʊ]

The periodic cycles for the Male 1 native Igbo speaker's utterance of *below (adj)* are 0.936322 secs and 1.215810 secs while the total time duration is 2.152132 secs. The pitch value is 120.9 Hz. The pitch contour is short, perhaps the speaker does not observe the lengthening introduced by the diphthong /əʊ/.

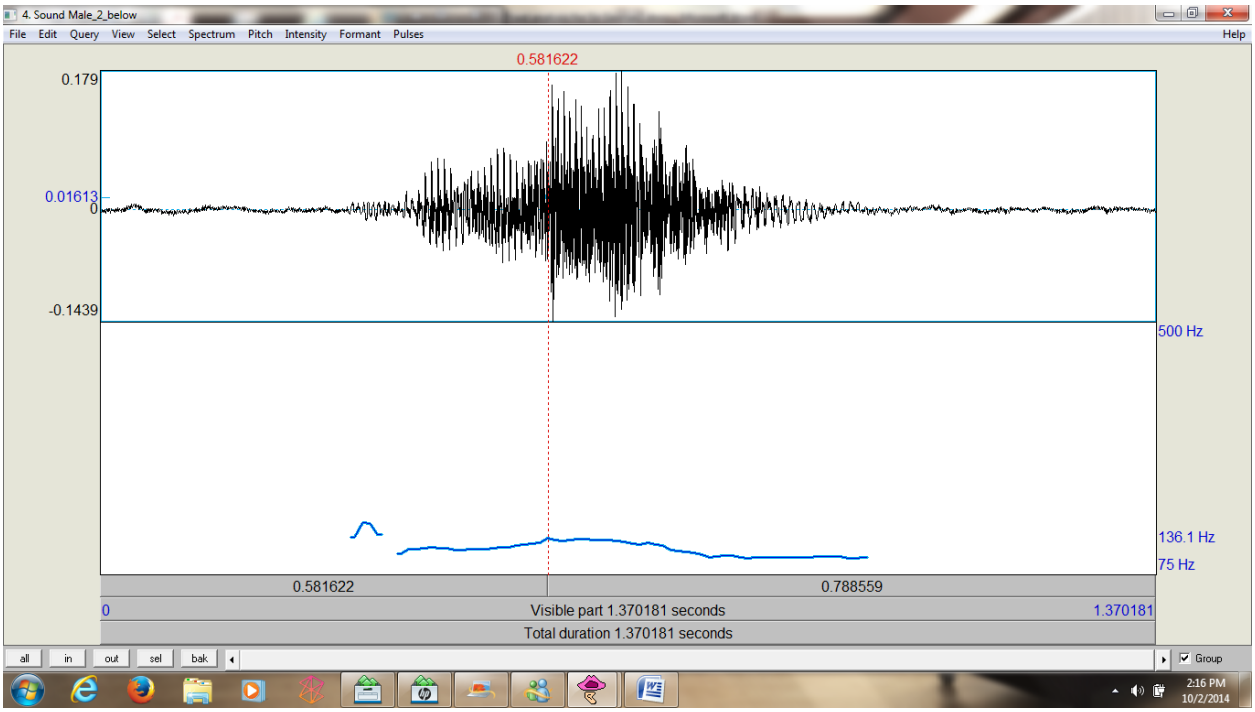


Figure 3(iii) Sound wave and *Fo* and time duration of Male 2 utterance of *below (adj)* [bɪˈləʊ]

The wave form cycles used by Male 2 native Igbo speaker in the production of the word *below (adj)* are 0.581622 secs and 0.788559 secs. The total time duration is 1.370181 secs, and the pitch value is 136.1 Hz. The pitch contour shows a pause at the inception of the utterance, perhaps, to signify the two-syllable structure of the word. But, the stress pattern appears wrongly placed, with the first syllable being assigned the primary stress.

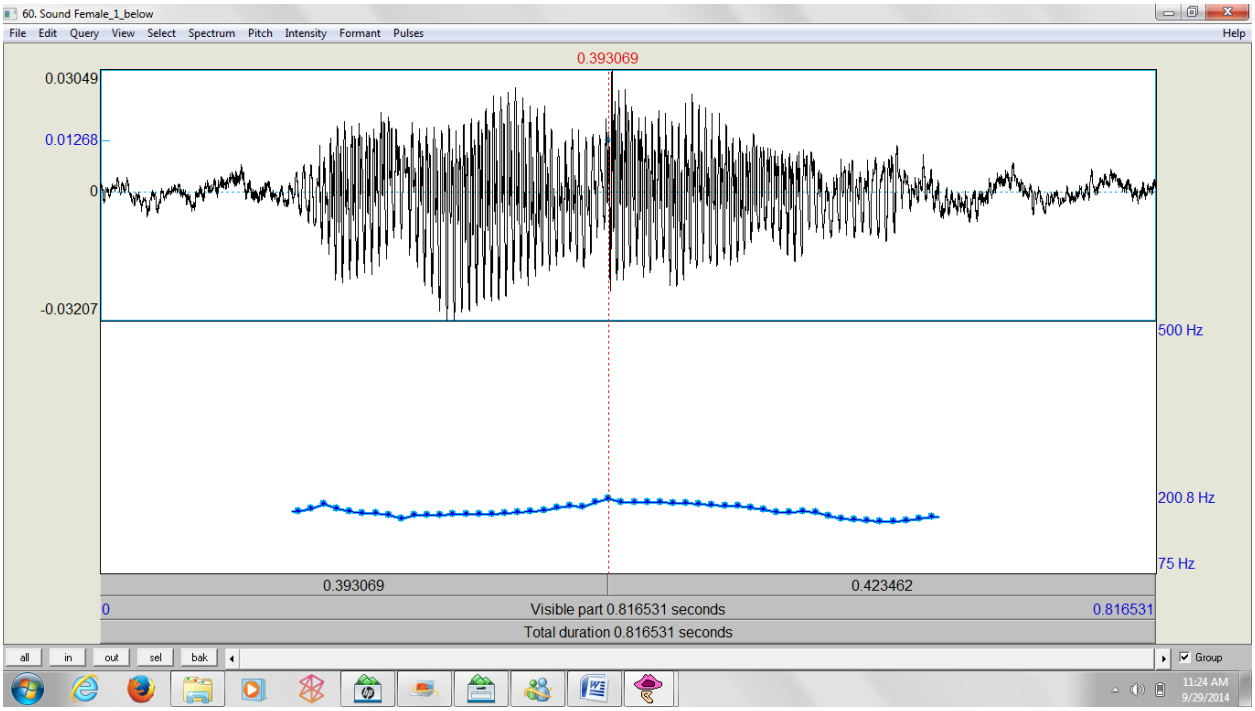


Figure 3(iv) Sound wave and *Fo* and time duration of Female 1 utterance of *below (adj)* [bɪˈləʊ]

The wave form cycles used by Female 1 here to produce *below (adj)* are 0.393069 secs and 0.423462 secs, while the total time duration is 0.816531 secs. The pitch value is 200.8 Hz. The pitch contour shows the stress pattern as unstressed in the first syllable and stressed in the second.

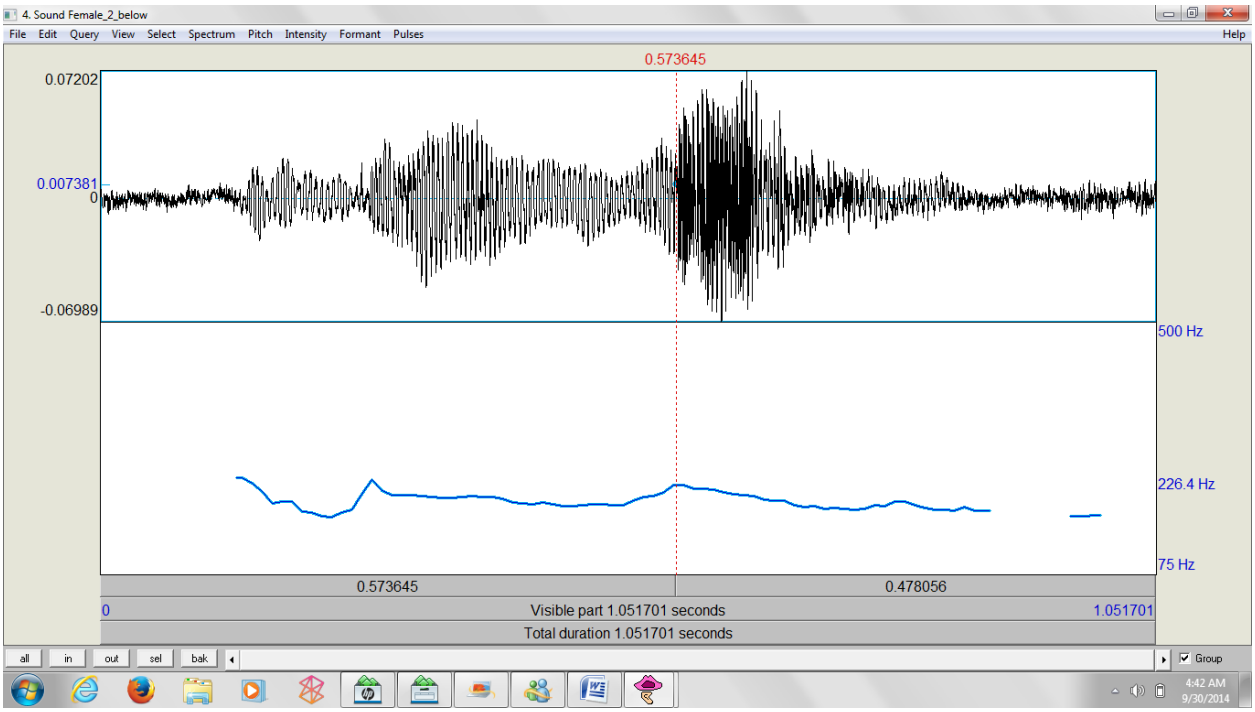


Figure 3(v) Sound wave and *Fo* and time duration of Female 2 utterance of *below (adj)* [bɪˈləʊ]

The periodic cycles for the production of *below (adj)* by the Female 2 native Igbo speaker are 0.573643 secs and 0.478056 secs. the total time duration is 1.051701 secs, and the pitch value is 228. 4 Hz. The pitch contour indicates a syllable structure and stress pattern, which has the last syllable stressed.

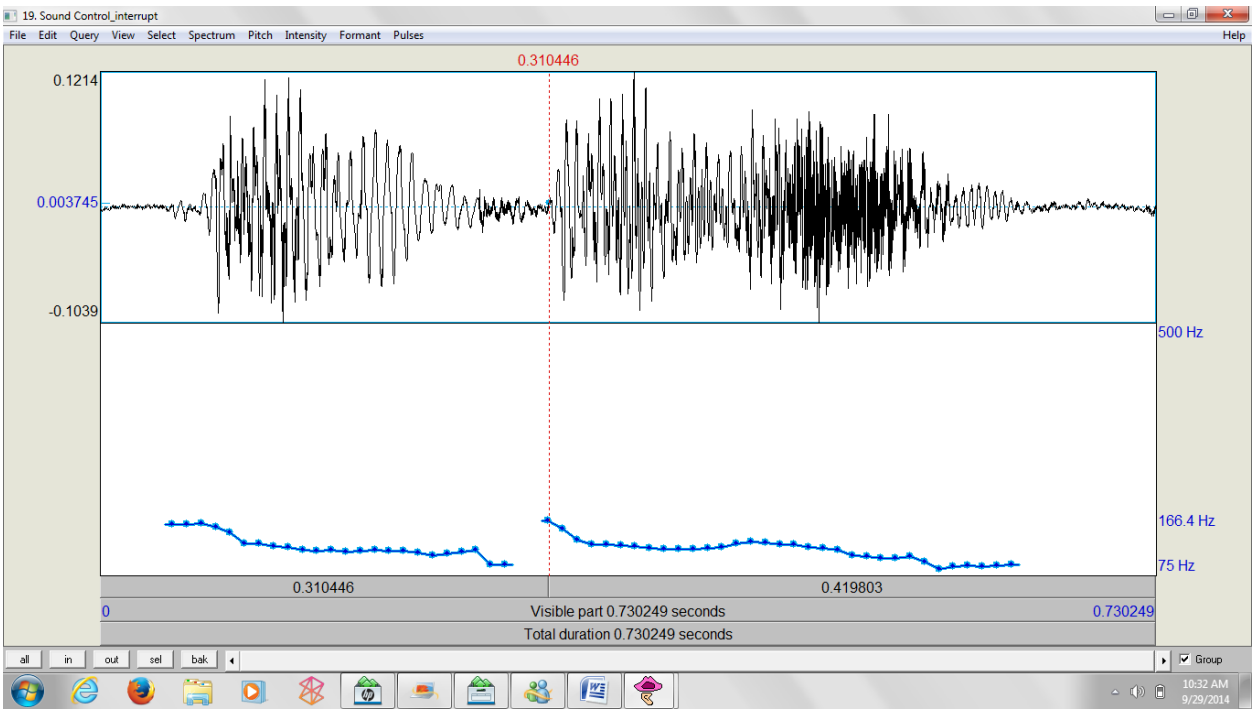


Figure 4(i) Sound wave and *Fo* and time duration of Control utterance of *interrupt (v)* [ˌɪntəˈrʌpt]

Interrupt (v) is a three-syllable word that has the third syllable as the heavy syllable. Thus, it bears the primary stress, while the first syllable bears the secondary stress and the third syllable is unstressed. In the Control utterance of the word, the wave form cycles of 0.310446 secs and 0.419803 secs giving total time duration of 0.730249 secs. The pitch value is 166.4 Hz. The pitch contour actually indicates that the first syllable bears secondary stress, the second unstressed, while the third syllable which indicates the highest pitch value bears the primary stress.

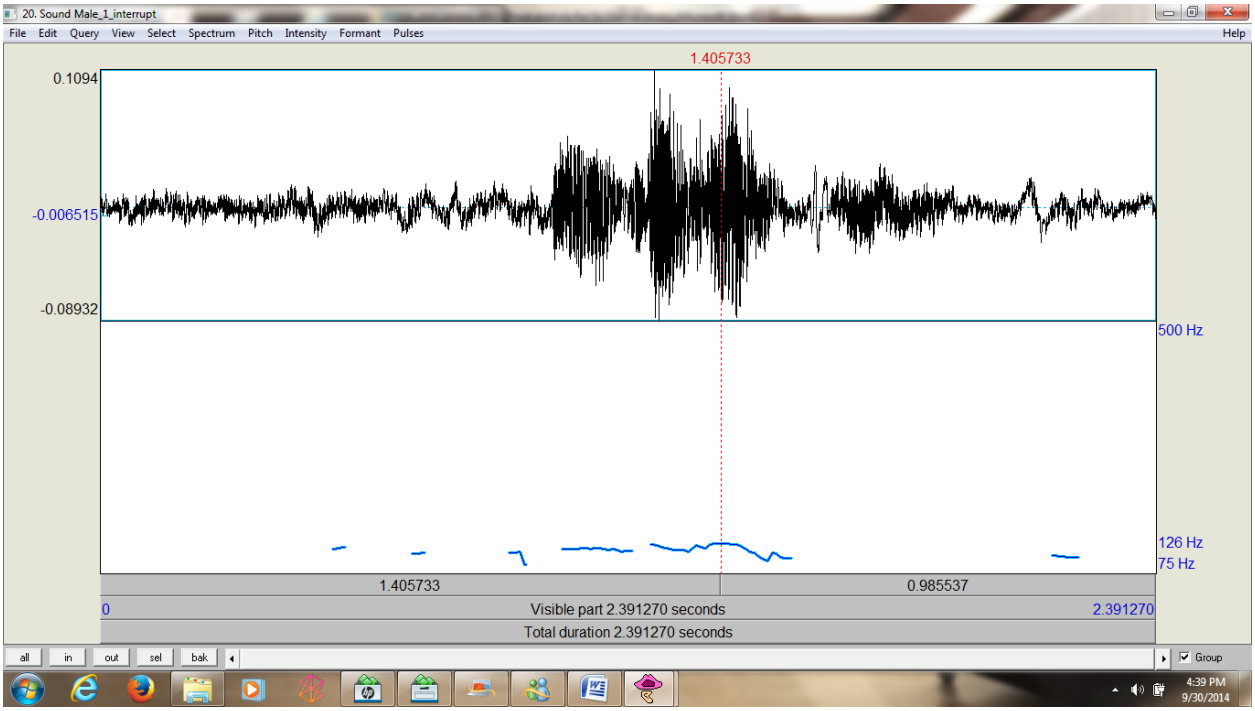


Figure 4(ii) Sound wave and *Fo* and time duration of Male 1 utterance of *interrupt (v)* [ˌɪntəˈrʌpt]

The Male 1 native Igbo speaker rendered *interrupt (v)* in the periodic cycles of 1.405733 secs and 0.985537 secs with a total time duration of 2.391270 secs. The pitch value for the utterance is 126 Hz. However, the pitch contour is not clear enough to determine the observance of the syllable structure of the word. This may have been caused by the speech style of the individual.

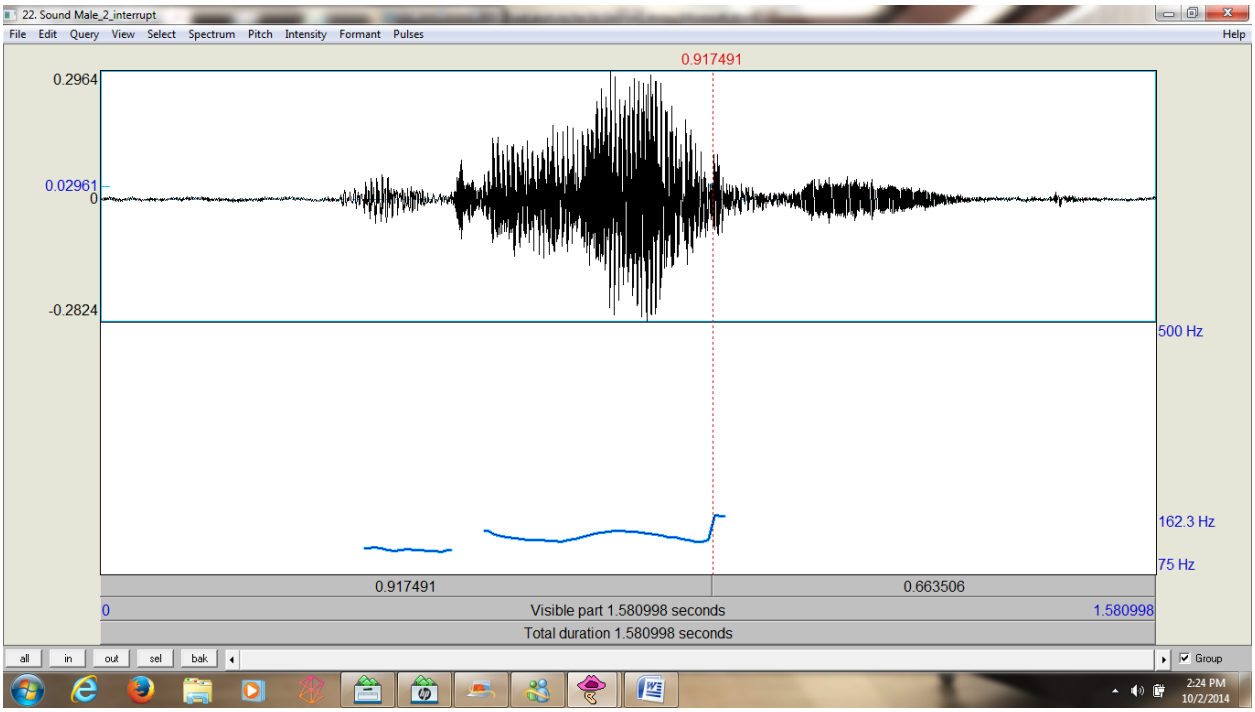


Figure 4(iii) Sound wave and *Fo* and time duration of Male 2 utterance of *interrupt (v)* [ˌɪntəˈrʌpt]

Male 2 utterance of *interrupt (v)* was done in the periodic cycles of 0.917491 secs and 0.663506 secs. The total time duration is 1.663506 secs, and the pitch value is 162.3 Hz. The pitch contour shows clearly that the word consists of three syllables with secondary stress, unstressed and primary stress.

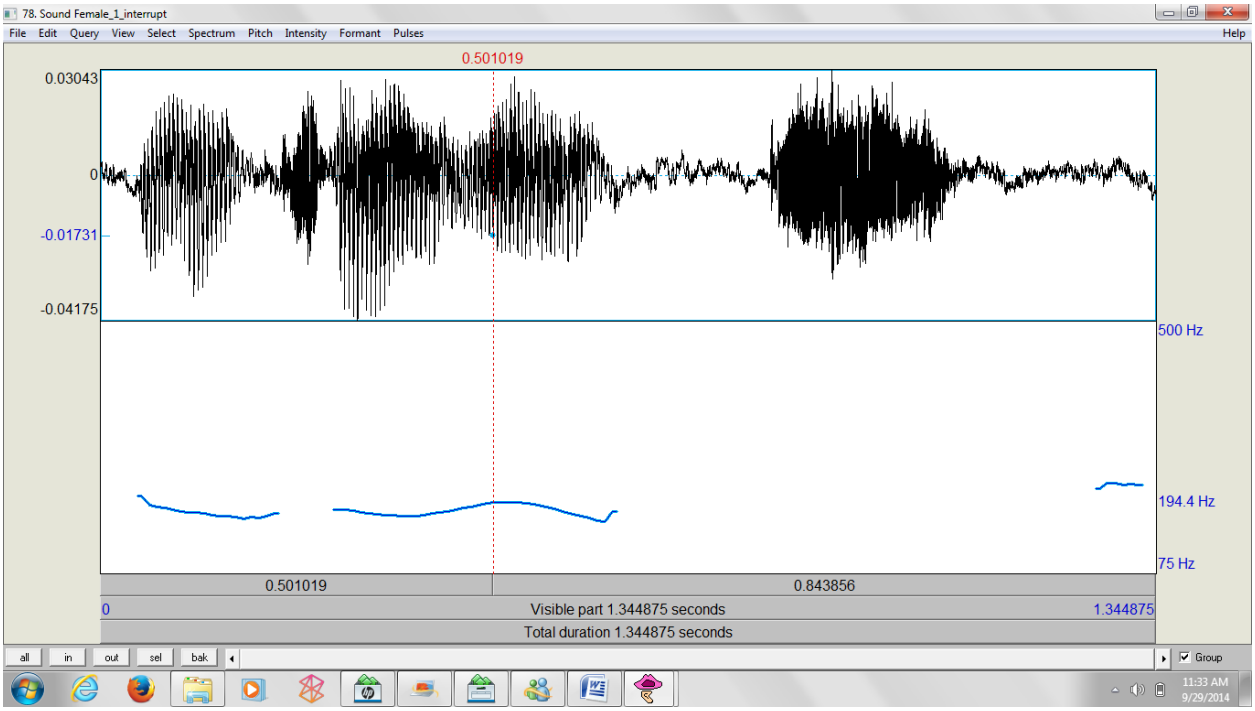


Figure 4(iv) Sound wave and *Fo* and time duration of Female 1 utterance of *interrupt (v)* [ˌɪntəˈrʌt]

Female 1 native Igbo speaker renders *interrupt (v)* in the periodic cycles of 0.501019 secs and 0.843856 secs. The total time duration is 1.344875 secs, and the pitch value is 194.4 Hz. Again, the syllable structure is indicated by the pitch contour. However, a trace showing delayed release of the last segment, the alveolar stop /t/ can be seen in the pitch contour.

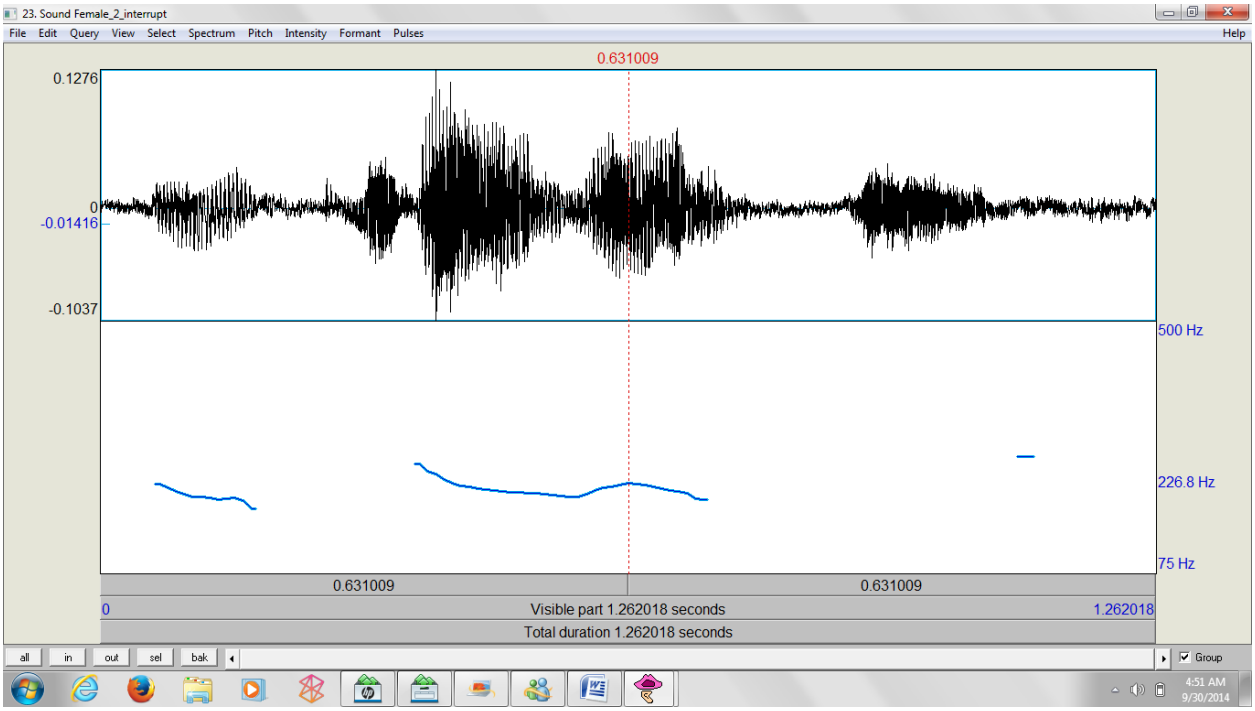


Figure 4(v) Sound wave and *Fo* and time duration of Female 2 utterance of *interrupt (v)* [ˌɪntəˈrʌpt]

The Female 2 utterance of *interrupt (v)* is presented in the periodic cycles of 0.631009 secs and 0.631009 secs, thus giving total time duration of 1.262018 secs. The pitch value stands at 226.8 Hz. Again, a trace of delayed release of the last segment in the word /t/ can be seen even where the pitch contour indicates the three-syllable structure and stress pattern of the word.

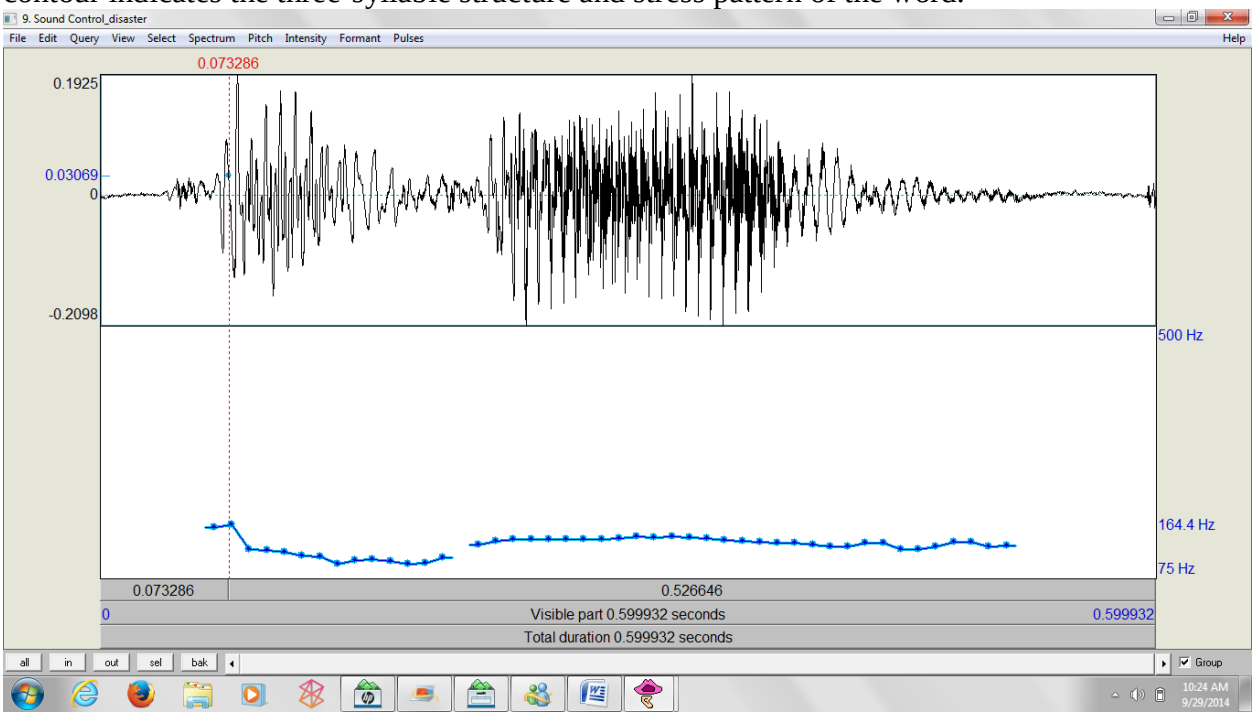


Figure 5(i) Sound wave and *Fo* and *time* duration of Control utterance of *disaster (n)* [dɪˈzaːstə(r)]

Disaster (n) is a three-syllable word which has its stress pattern as unstressed, primary stress and unstressed. The Control utterance of the word is made in periodic cycles of 0.7073286 secs and 0.526646 secs, giving total time duration of 0.59932 secs. The highest pitch value measures 164.4 Hz. The pitch contour seems to show an initial near zero prominence followed by an increase in prominence signalled by pitch rise introduced by the stressed syllable and a gradual drop in prominence introduced by the following unstressed syllable.

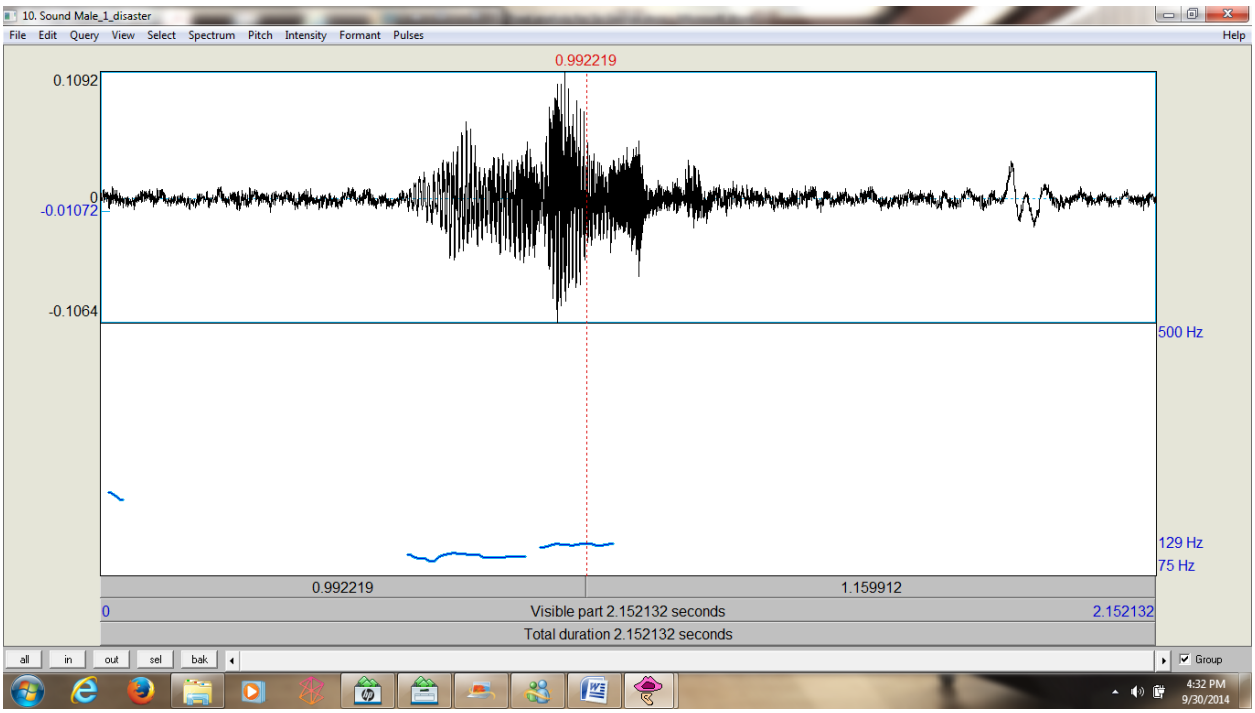


Figure 5(ii) Sound wave and *Fo* and time duration of Male 1 utterance of *disaster (n)* [dɪˈzɑːstə(r)]

The Male 1 utterance of *utterance (n)* was rendered in periodic cycles of 0.992219 secs and 1.59915 secs, giving total time duration of 2.152132 secs. The pitch value is 129 Hz. However, the pitch contour seems to portray the unstressed initial syllable with a pause before the next string of the utterance, which appears to have been given the greatest prominence.

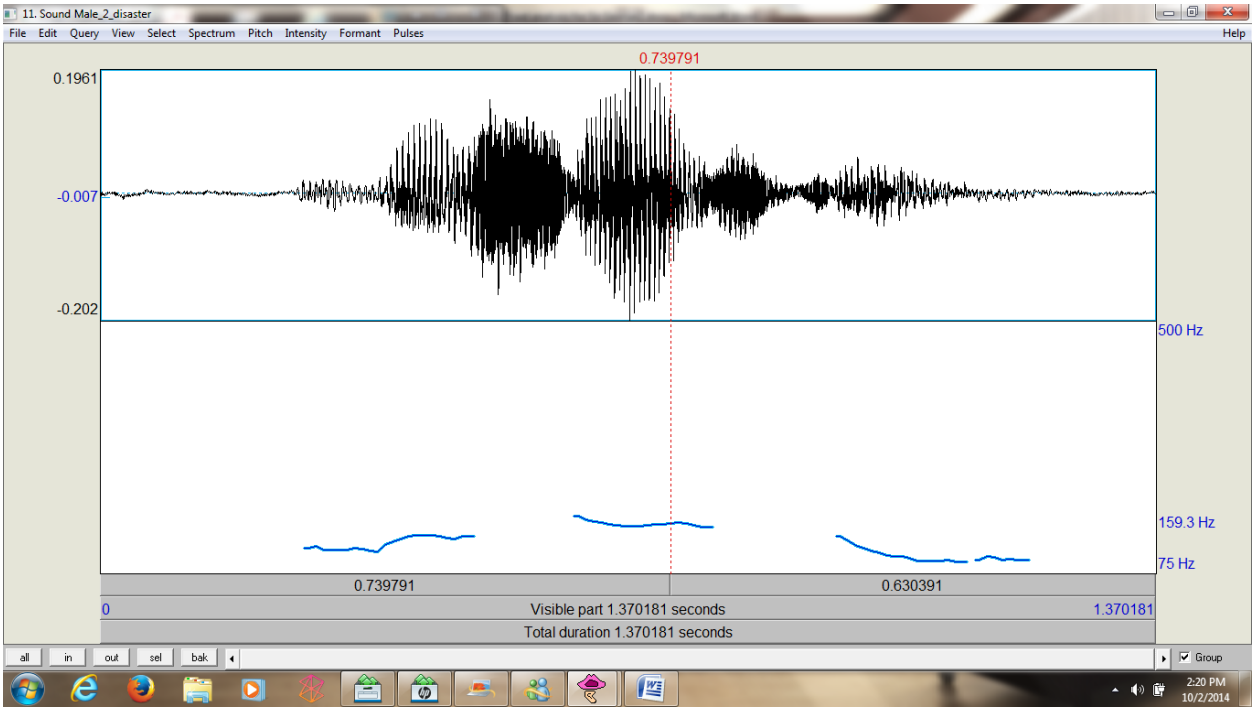


Figure 5(iii) Sound wave and *Fo* and time duration of Male 2 utterance of *disaster (n)* [dɪˈzɑːstə(r)]

Male 2 utterance of *disaster (n)* was rendered in periodic cycles of 0.739791 secs and 0.630391 secs, giving total time duration of 1.370181 secs. The pitch value is 159.3 Hz. It is further observed that the pitch contour appears to represent the three syllable structure of the word with the second syllable showing greater prominence and the last receiving no prominence. Also, it is obvious that the utterance is made in three beats with intermittent pauses.

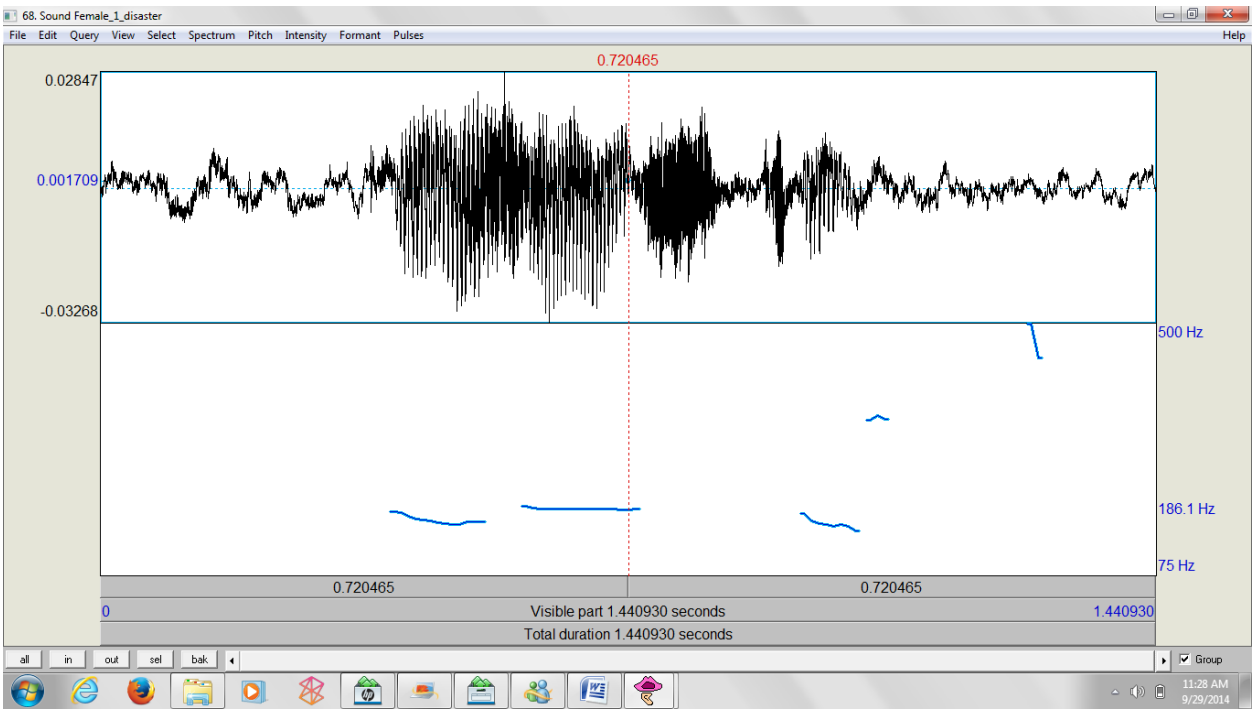


Figure 5(iv) Sound wave and *Fo* and time duration of Female 1 utterance of *disaster (n)* [dɪˈzɑːstə(r)]

Female 1 produced *disaster (n)* in periodic cycles of 0.302069 secs and 1.170312 secs. The total time duration is 1.472381 secs, and the highest pitch value is 223.5 Hz. Clearly, the pitch contour presents the word in three syllables with the third syllable remarkably low (unstressed). Also, the utterance is interspersed by pauses.

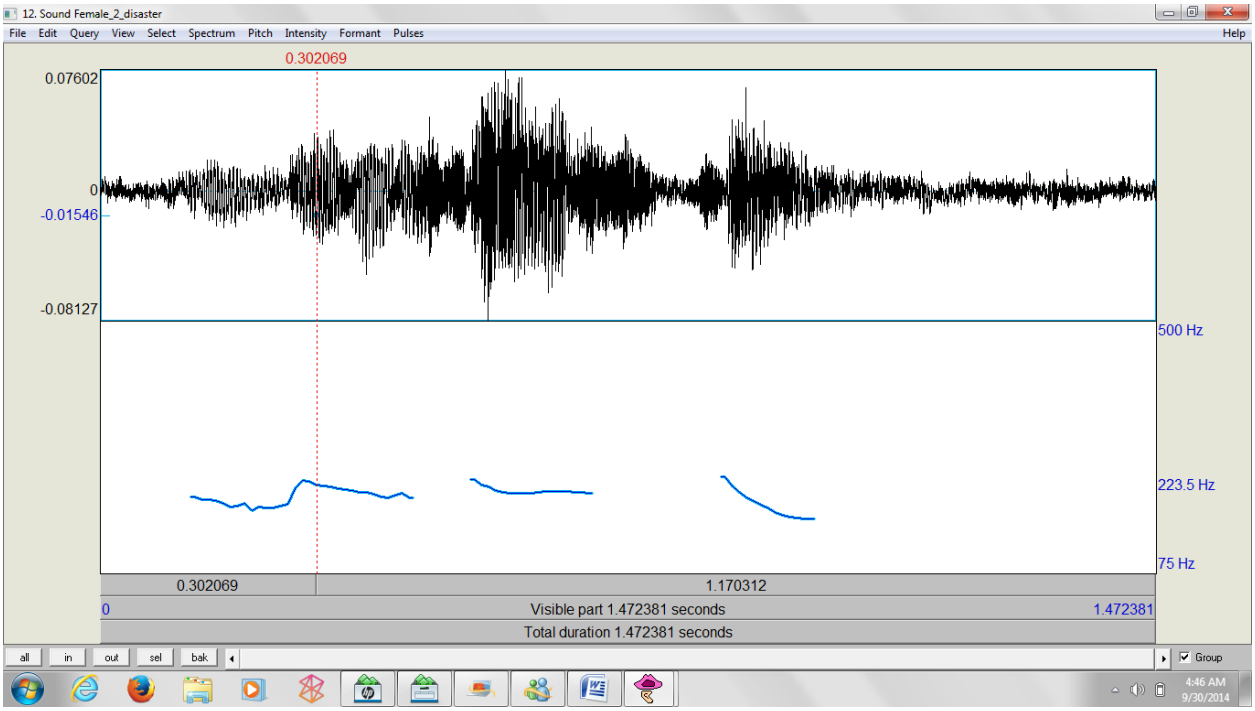


Figure 5(v) Sound wave and *Fo* and time duration of Female 2 utterance of *disaster (n)* [dɪˈza:stə(r)]

Female 2 uttered *disaster (n)* in periodic cycles of 0.302069 secs and 1.170312 secs. The total time duration is 1.472381 secs, and the highest pitch value is 223.5 Hz. The pitch contour presents the word in three syllables interspersed by pauses, and the third syllable is a fall (unstressed).

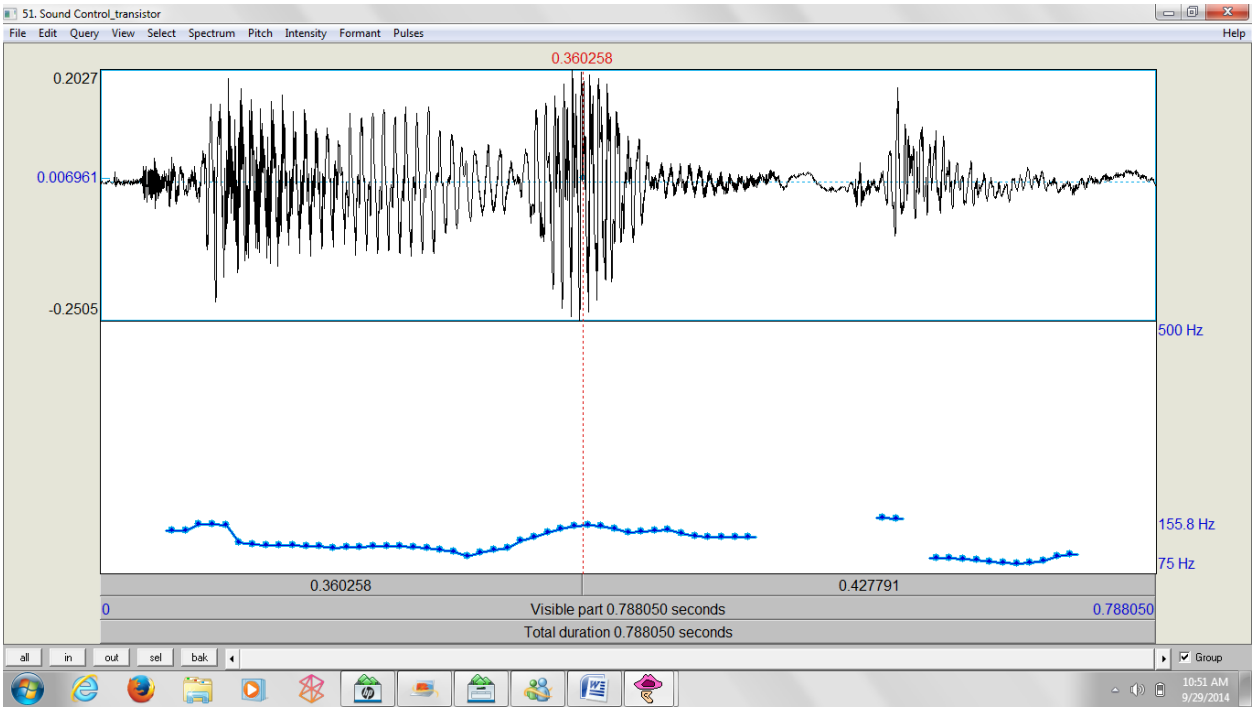


Figure 6(i) Sound wave and *Fo* and time duration of Control utterance of *transistor (n)* [trænˈzɪs.tə(r)]

The Control utterance of the three-syllable word *transistor (n)* was made in periodic cycles of 0.360258 secs and 0.427791 yielding a total time of 0.788050 secs. The highest pitch value is 155.8 Hz. The pitch contour presents the three syllables of the word, with the first syllable as bearing the secondary stress, the second syllable as the heavy stress and therefore bearing the primary stress. The third syllable is unstressed.

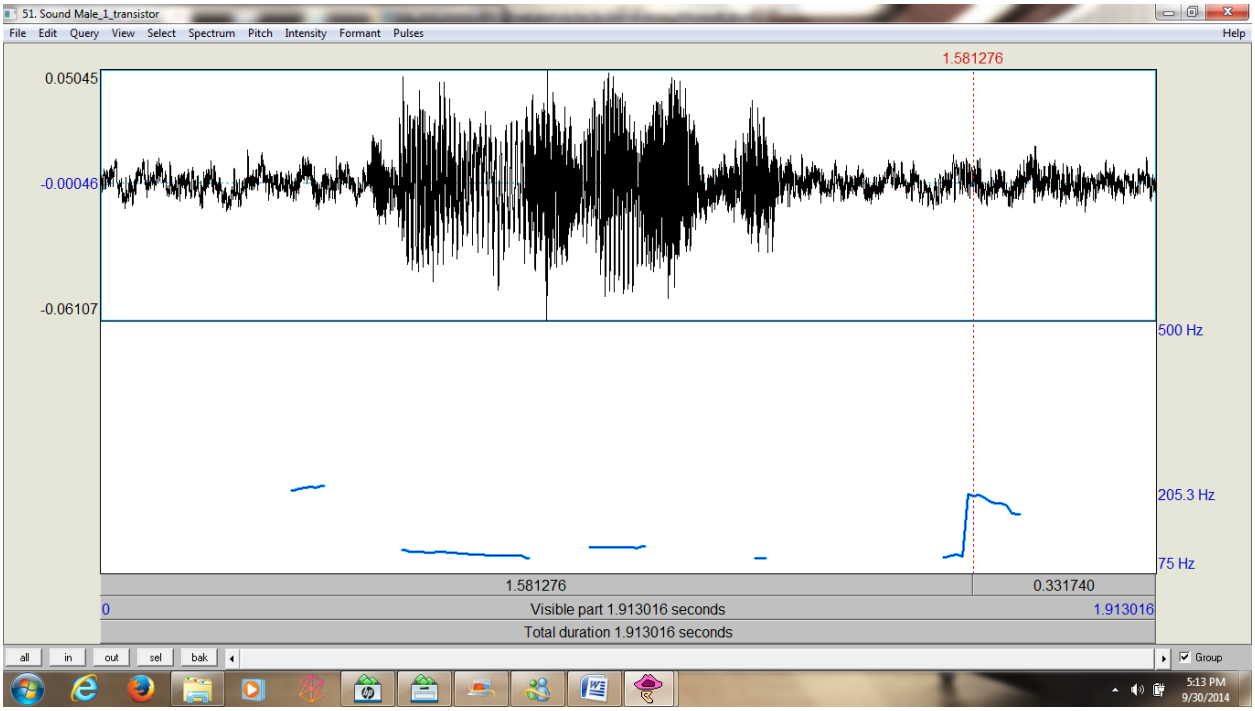


Figure 6(ii) Sound wave and Fo **and time duration of Male 1 utterance of *transistor (n)* [trænˈzɪs.tə(r)]**

Male 1 utterance of *transistor (n)* is presented in periodic cycles of 1.581276 secs and 0.331740 secs. The total time duration is 1.9313016 secs, and the pitch value is 205.3 Hz. The pitch contour appears in several stretches interspersed by pauses. There is evidence of rise in pitch immediately followed by a fall at the end of the word.

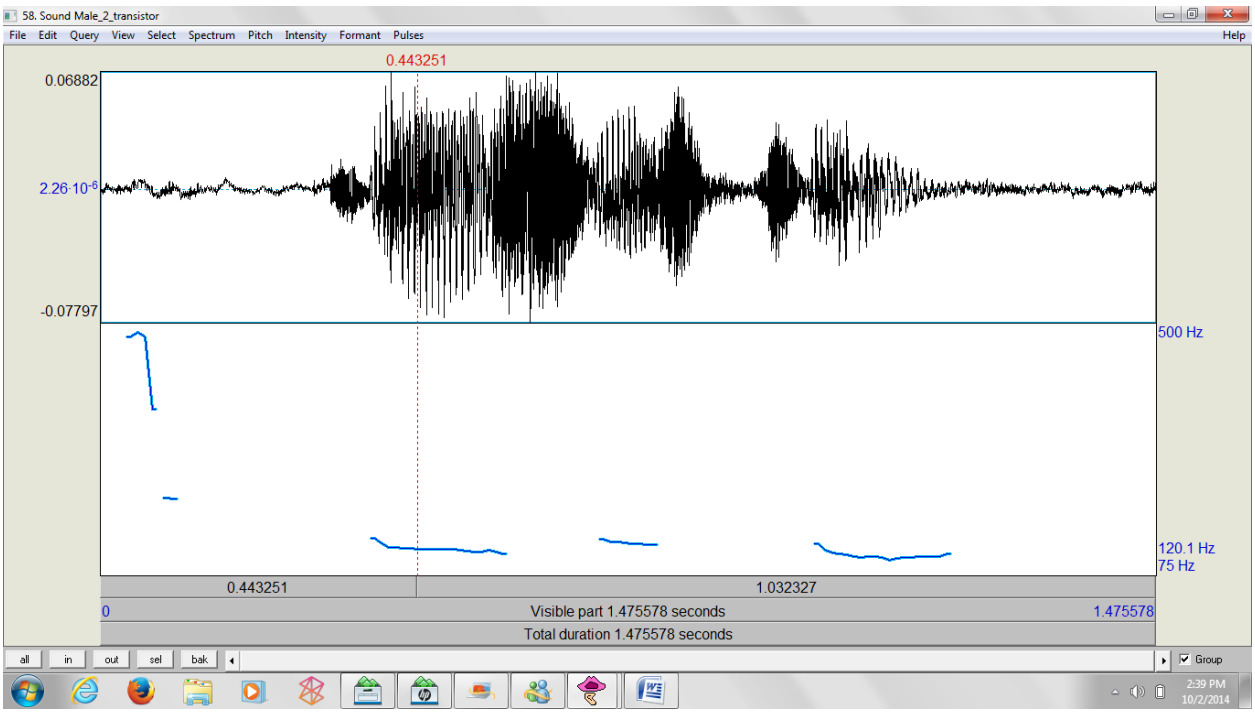


Figure 6(iii) Sound wave and *Fo* and time duration of Male 2 utterance of *transistor (n)* [trænˈzɪs.tə(r)]

The Male 2 utterance of *transistor (n)* is presented in periodic cycles of 0.443251 secs and 1.032327 secs. The total time duration is 1.475578 secs while the highest pitch value is 120.1 Hz. The pitch contour presents the word in three stretches (syllables) with the second syllable appearing slightly more prominent.

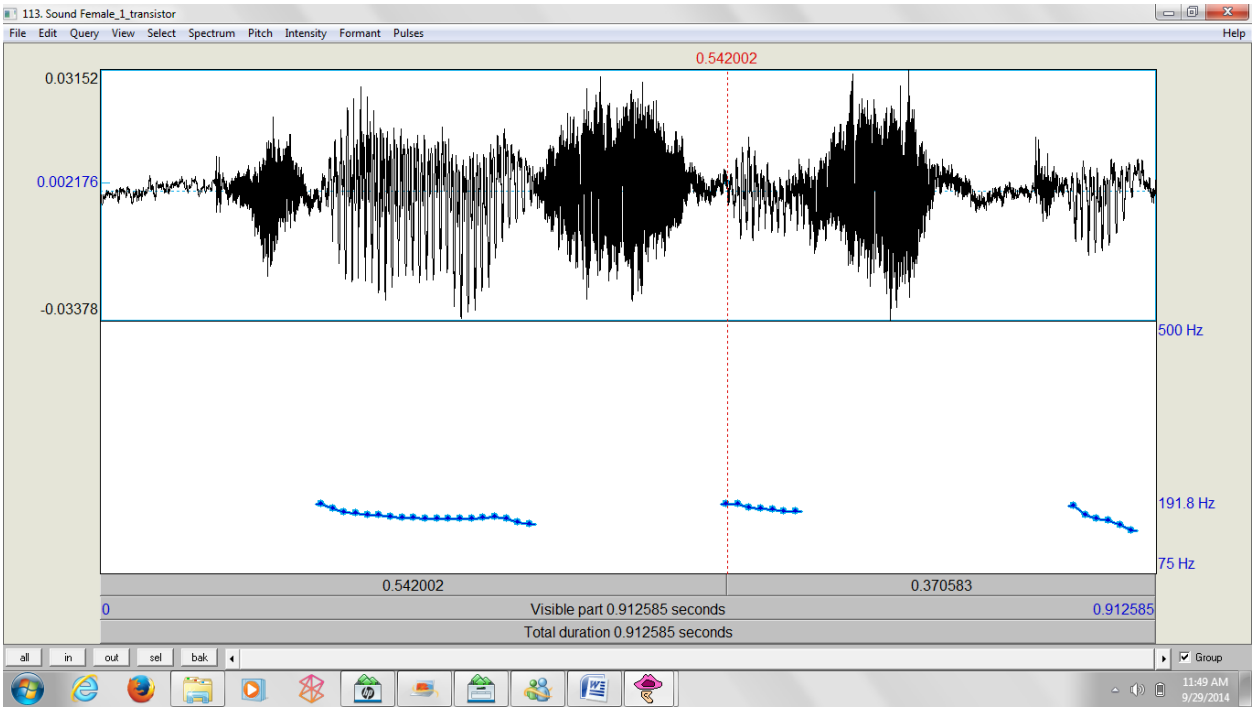


Figure 6(iv) Sound wave and *Fo* and time duration of Female 1 utterance of *transistor (n)* [trænˈzɪs.tə(r)]

The Female 1 utterance of *transistor (n)* is presented in waveform cycles of 0.542002 secs and 0.370583 secs, giving total time duration of 0.912585 secs. The pitch value is 191.8 Hz. Here, the pitch contour clearly presents the three syllables of the word with the second syllable bearing the primary stress; the first syllable shows less prominence while the pitch drops in the third syllable; thus, it is unstressed.

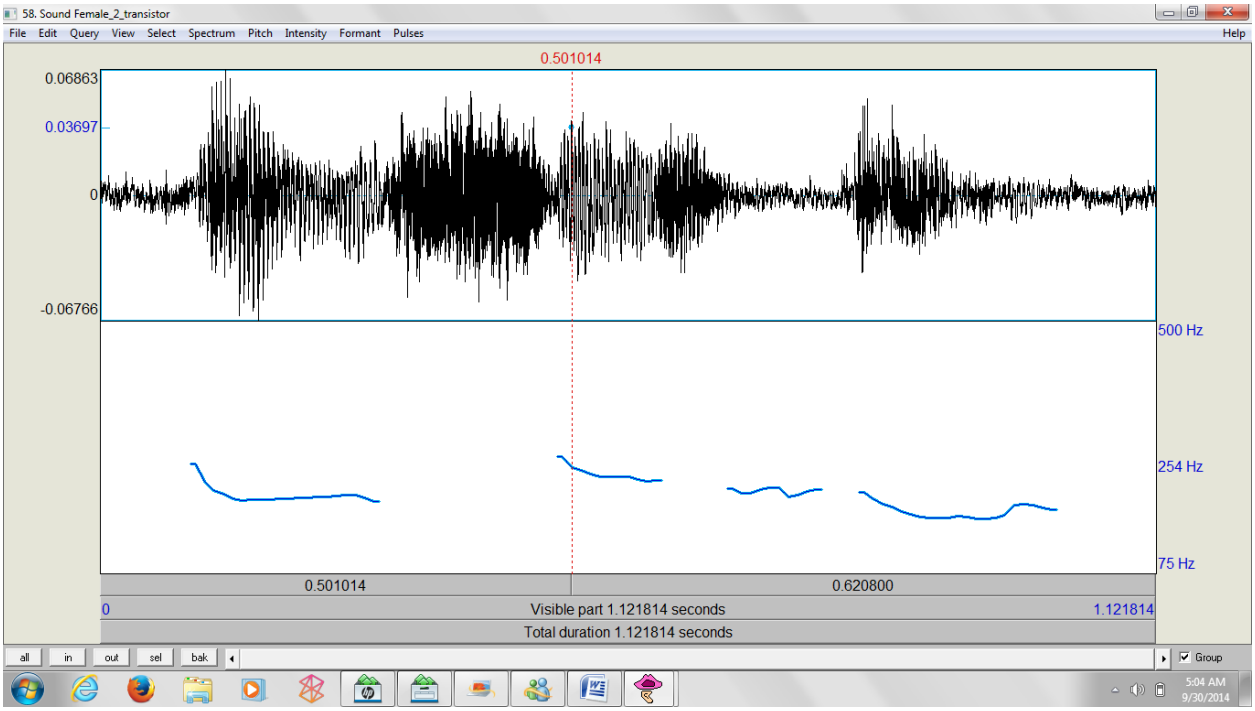


Figure 6(v) Sound wave and *Fo* and time duration of Female 2 utterance of *transistor (n)* [trænˈzɪs.tə(r)]

Female 2 utterance of *transistor (n).*This is rendered in periodic cycles of 0.501014 secs and 0.620800 secs, giving the total time duration as 10121814 secs. The pitch value is 254 Hz. The pitch contour here appears in four stretches. However, the second stretch carries the greatest prominence; thus, it bears the primary stress, while the following stretches drop; they are unstressed.

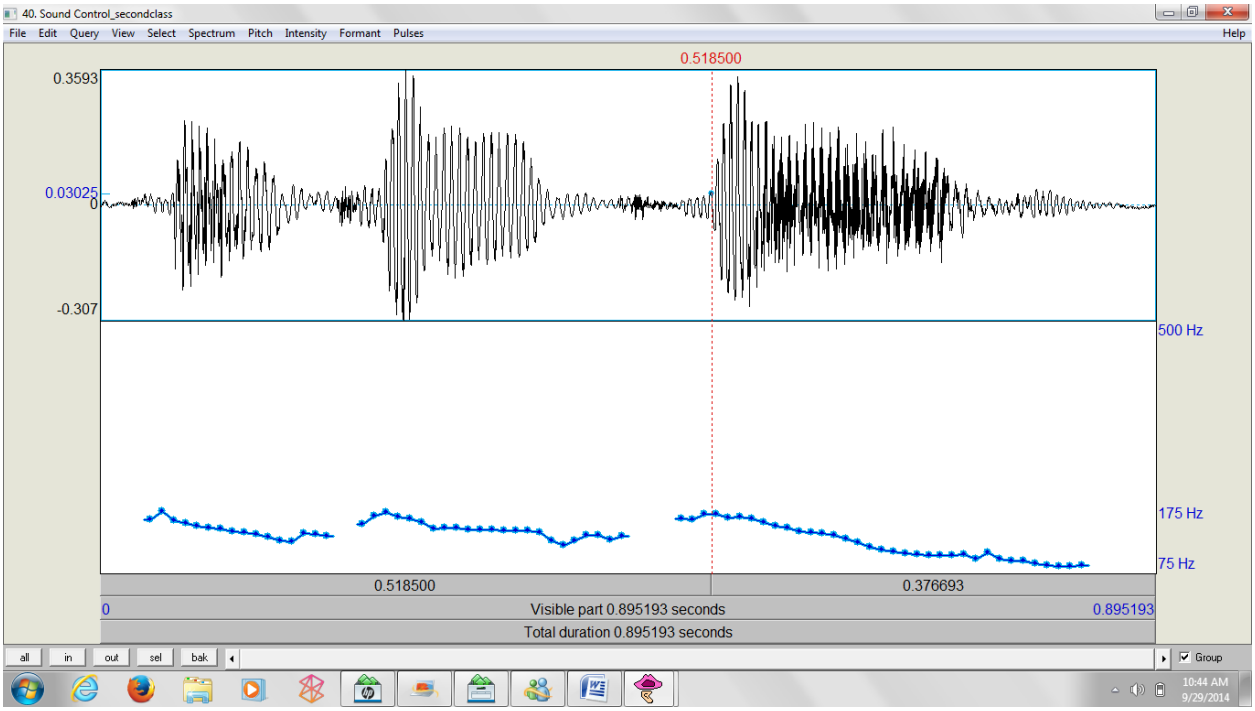


Figure 7(i) Sound wave and *Fo* and time duration of Control utterance of *second-class* [ˌsek.ᵊ*nd*ˈkla:s]

A compound word has its stress either on the first or second element depending on the word class of the elements that constitute it.

Second-class as a compound word is an adjective (and number). The primary stress is to be on the second element, *class.* The Control utterance of the word is in periodic cycles of 0.518500 secs and 0.376693 secs, giving the total time duration as 0.895193 secs. The highest pitch value is 175 Hz. The pitch contour indicates the three-syllable structure of the compound with the third syllable as the heavy syllable and therefore stressed.

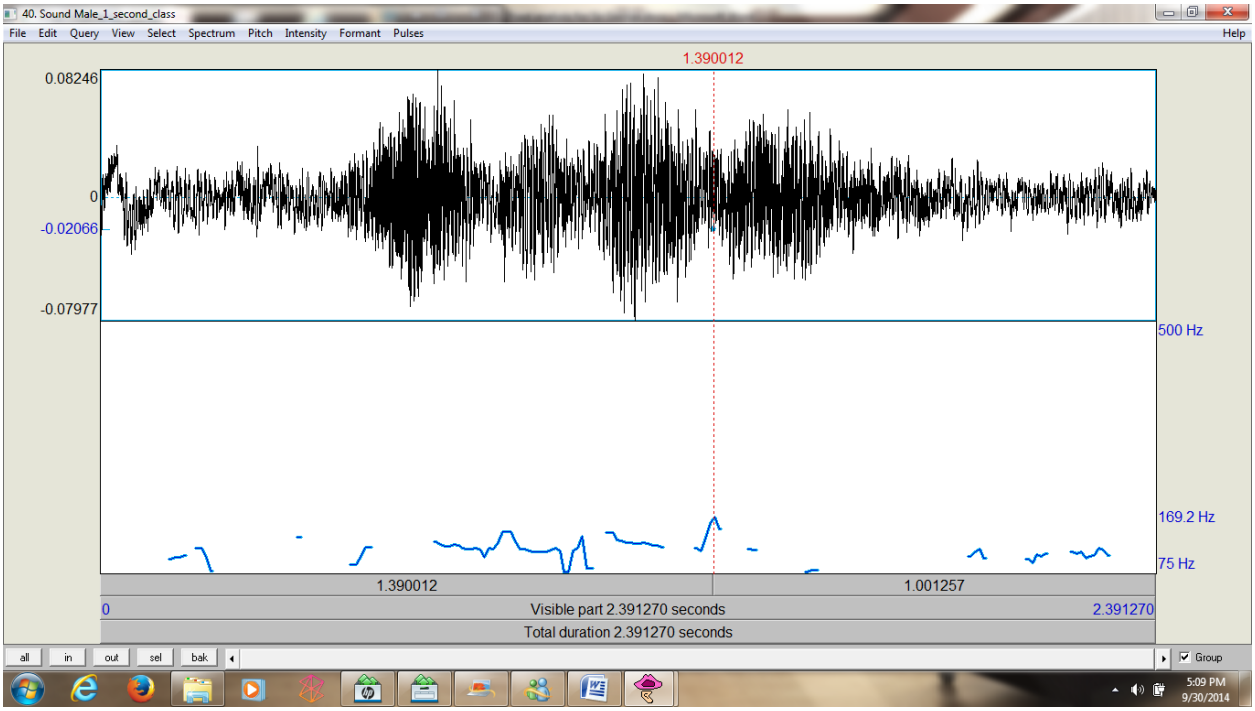


Figure 7(ii) Sound wave and *Fo* and time duration of Male 1 utterance of *second-class* [ˌsek.ˀ*nd*ˈkla:s]

Male 1 uttered *second-class* in periodic cycles of 1.390012 secs and 1.00257 secs. However, the pitch contour presents an unclear picture of the syllable structure and stress pattern resembling the pitch contour of an aperiodic waveform.

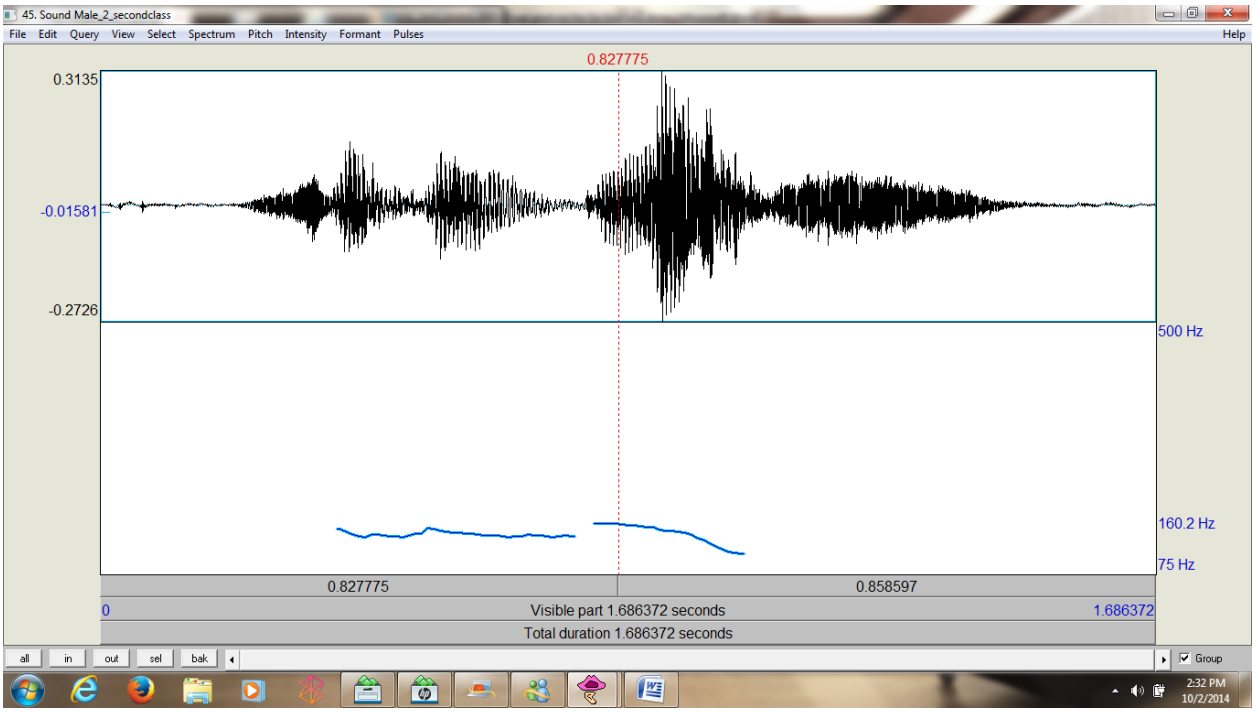


Figure 7(iii) Sound wave and *Fo* and time duration of Male 2 utterance of *second-class* [ˌsek.ˀ*nd*ˈkla:s]

Male 2 utterance is in periodic cycles of 0.827775 secs and 0.858597 secs, giving the total time duration as 1.686372 secs and the pitch value as 160.2 Hz. The pitch contour appears to indicate that the first two syllables in the first element *second* are glossed together. However, the stress is rightly placed on the second element, *class*.

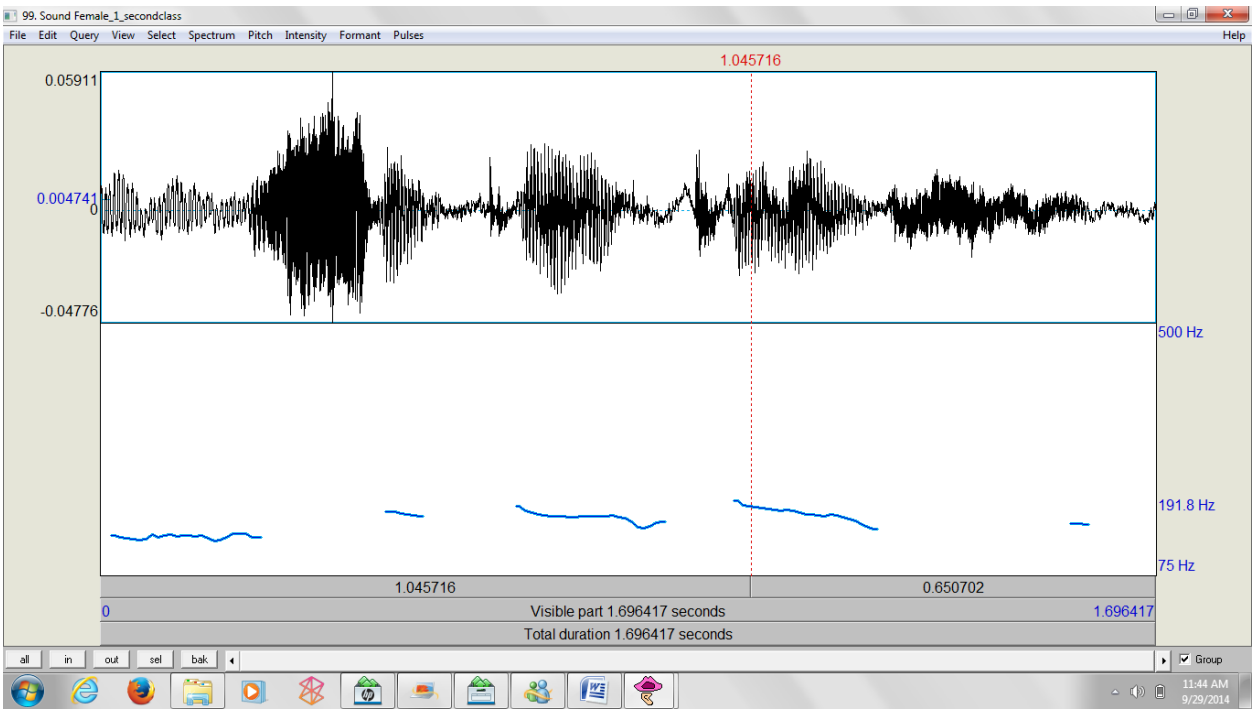


Figure 7(iv) Sound wave and *Fo* and time duration of Female 1 utterance of *second-class* [ˌsek.ˀ*nd*ˈkla:s]

Female 1 produced the compound in periodic cycles of 1.045716 secs and 0.650702 secs, with the total time duration as 1.696417 secs. The pitch value is191.8 Hz. The pitch contour for the utterance is in several stretches, which do not correspond to the syllable structure and stress pattern of the word.

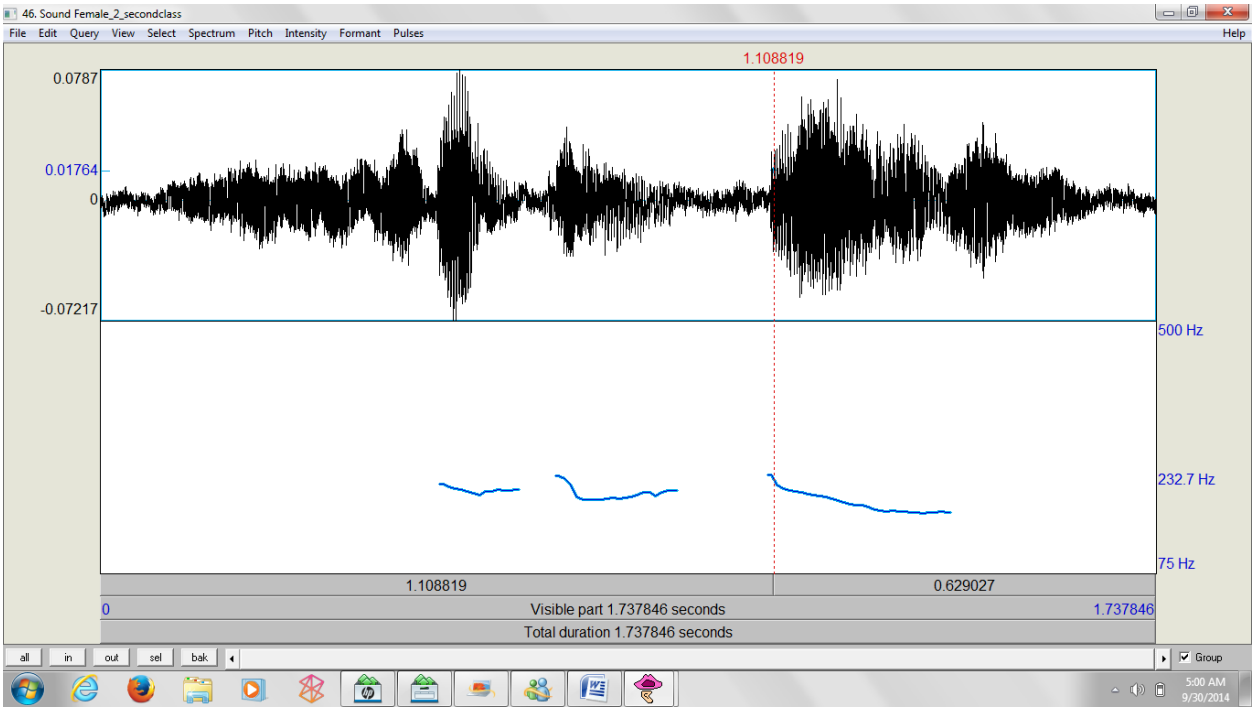


Figure 7(v) Sound wave and *Fo* and time duration of Female 2 utterance of *second-class* [ˌsek.ˀndˈkla:s]

The Female 2 utterance of *second-class* was produced in periodic cycles 1.108819 secs and 0.629027 secs, totalling to 1.737846 secs. The highest pitch value is 232.7 Hz. The pitch contour presents the three-syllable structure of the compound, while the third syllable bears the primary stress as the heavy syllable.

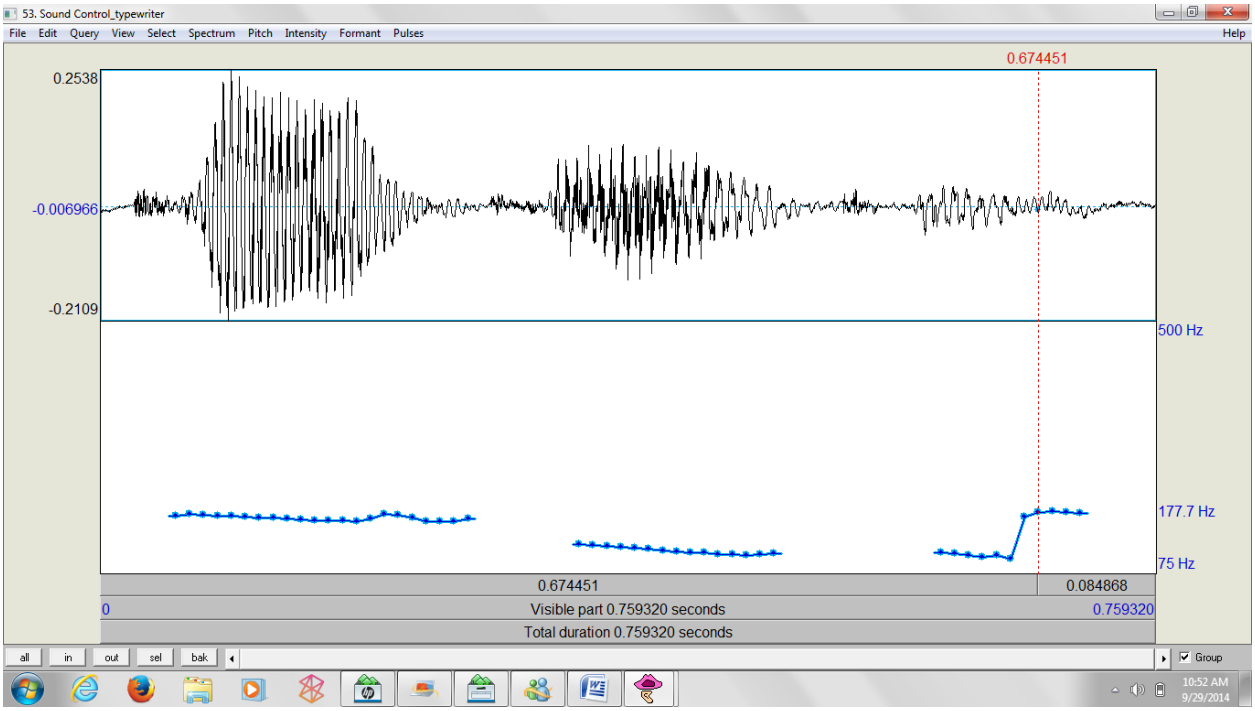


Figure 8(i) Sound wave and *Fo* and time duration of Control utterance of *typewriter* ['taɪpˌraɪ.tə(r)]

This is a compound word made up of two nominals with the stress normally on the first element. Here, it is produced by the Control at the periodic cycles of 0.674451 secs and 0.084868 secs, both summing up to 0.759320 secs. The highest pitch value is 117.7 Hz. The pitch presents the three syllables of the compound. But the highest pitch value appears to be on the third syllable.

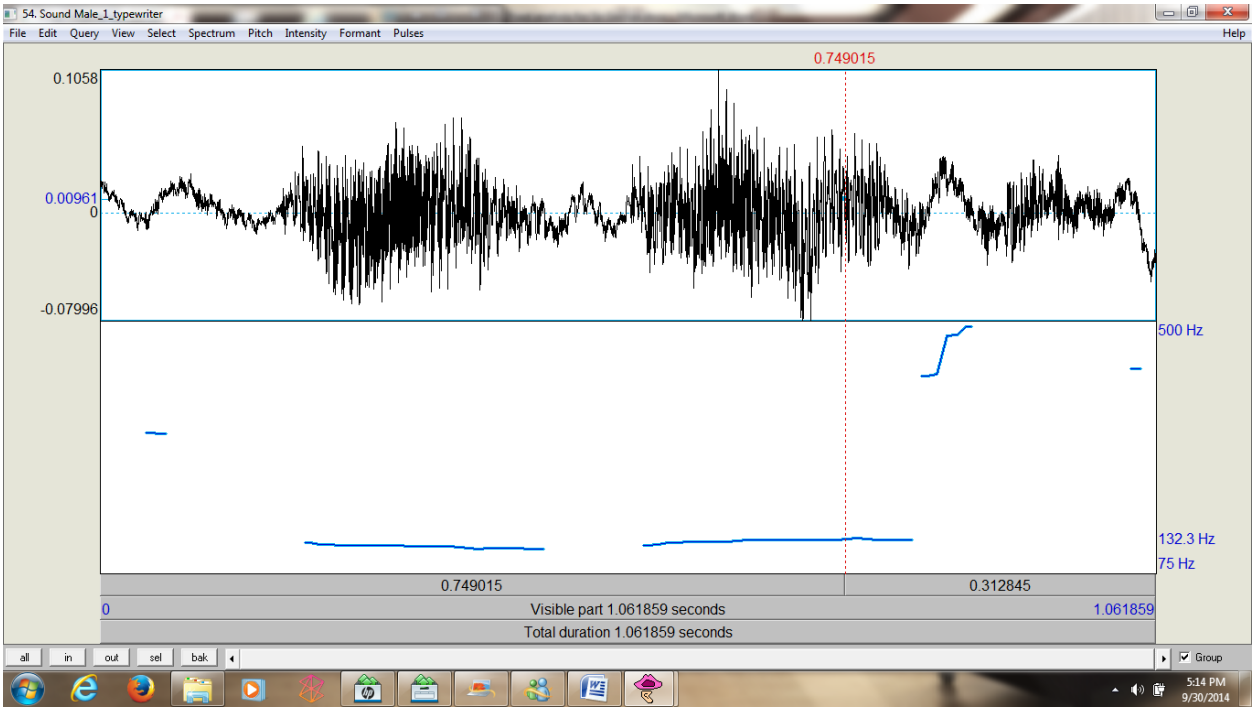


Figure 8(ii) Sound wave and *Fo* and time duration of Male 1 utterance of *typewriter* ['taɪpˌraɪ.tə(r)]

Male 1 utterance of *typewriter* is rendered in periodic cycles of 0.749015 secs and 0.312845 secs, giving total time duration of 1.061859 secs. The highest pitch value is 132.3 Hz. The pitch contour does not clearly present the three-syllable structure of the word and the utterance appears to assign the heavy syllable to the third syllable that otherwise should be unstressed.

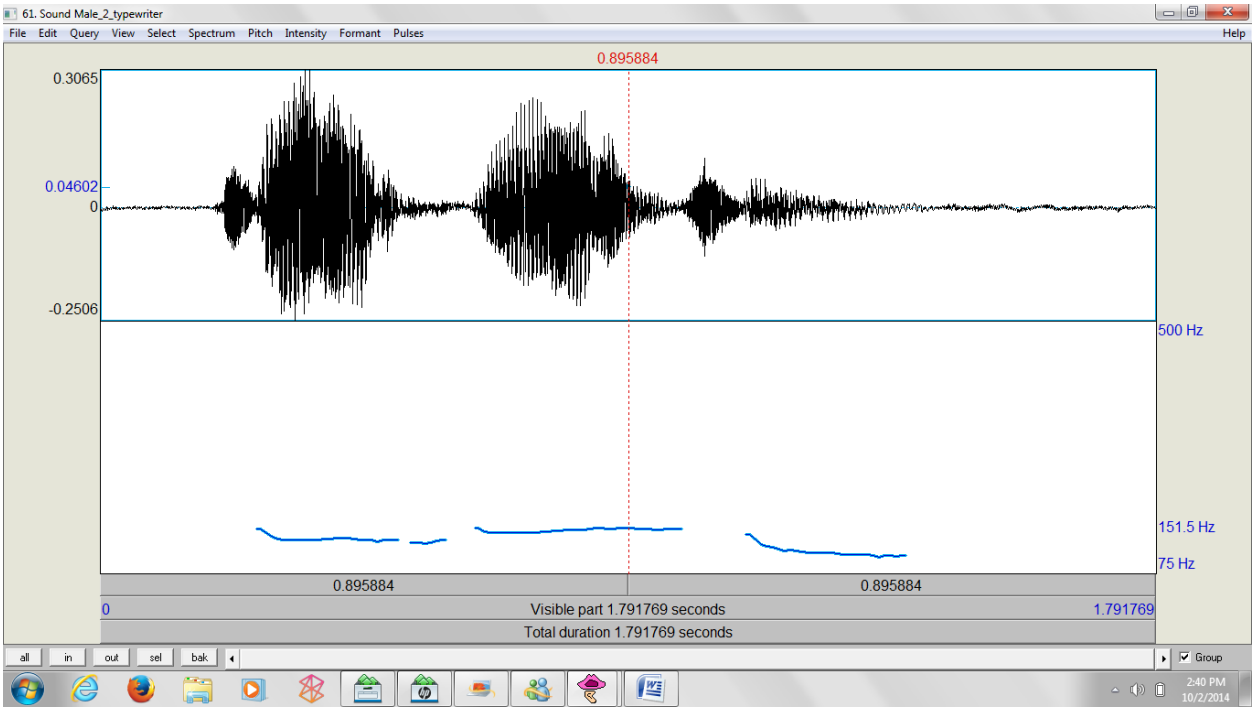


Figure 8(iii) Sound wave and *Fo* and time duration of Male 2 utterance of *typewriter* ['taɪp ˌraɪ.tə(r)]

Male 2 utterance of *typewriter* made in periodic cycles of 0.895884 secs and 0.895884 summing up to 1.791769 secs. The pitch value measures 151.5 Hz. There is evidence of delayed release of the voiceless bilabial plosive /p/, while the second syllable appears to have been assigned the primary stress.

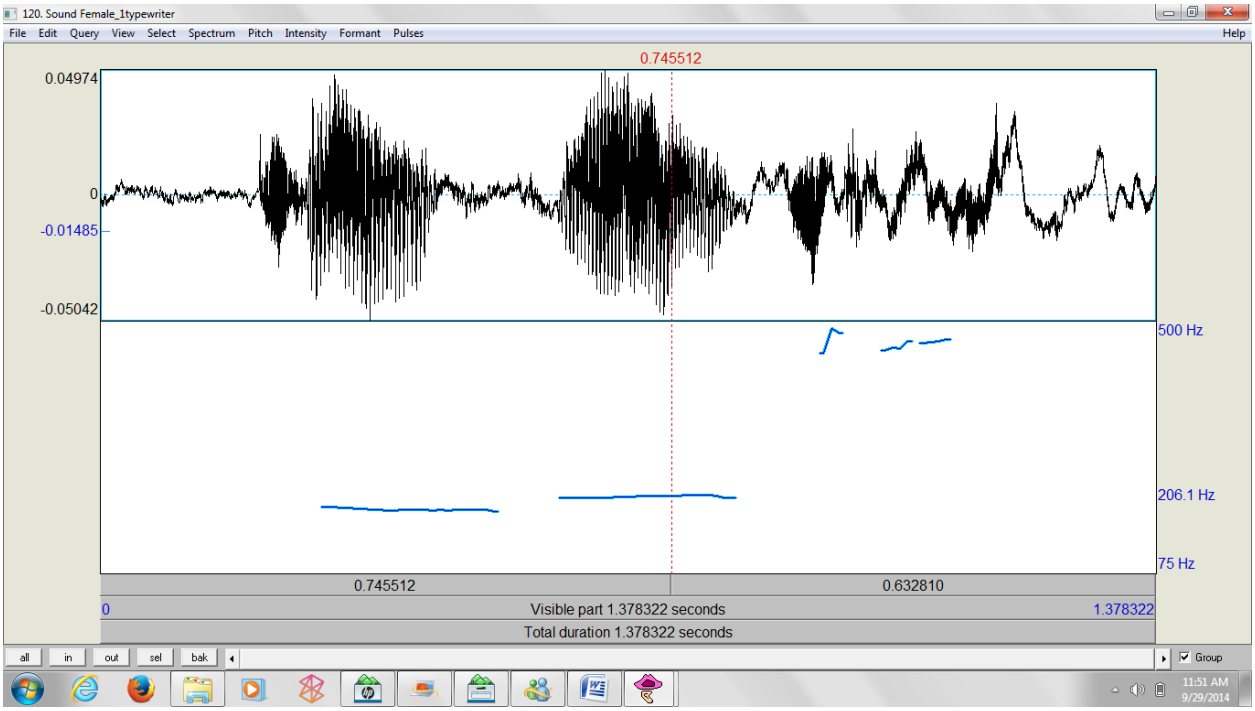


Figure 8(iv) Sound wave and *Fo* and time duration of Female 1 utterance of *typewriter* ['taɪp ˌraɪ.tə(r)]

Female 1 uttered *typewriter* in periodic cycles of 0.745512 secs and 0.632810 secs, thus giving a total time duration of 1.378322 secs. The highest pitch value is 206.1 Hz. The pitch contour signifies that the utterance is rendered in two syllables as against three syllables. Also, the second syllable carries the primary stress here.

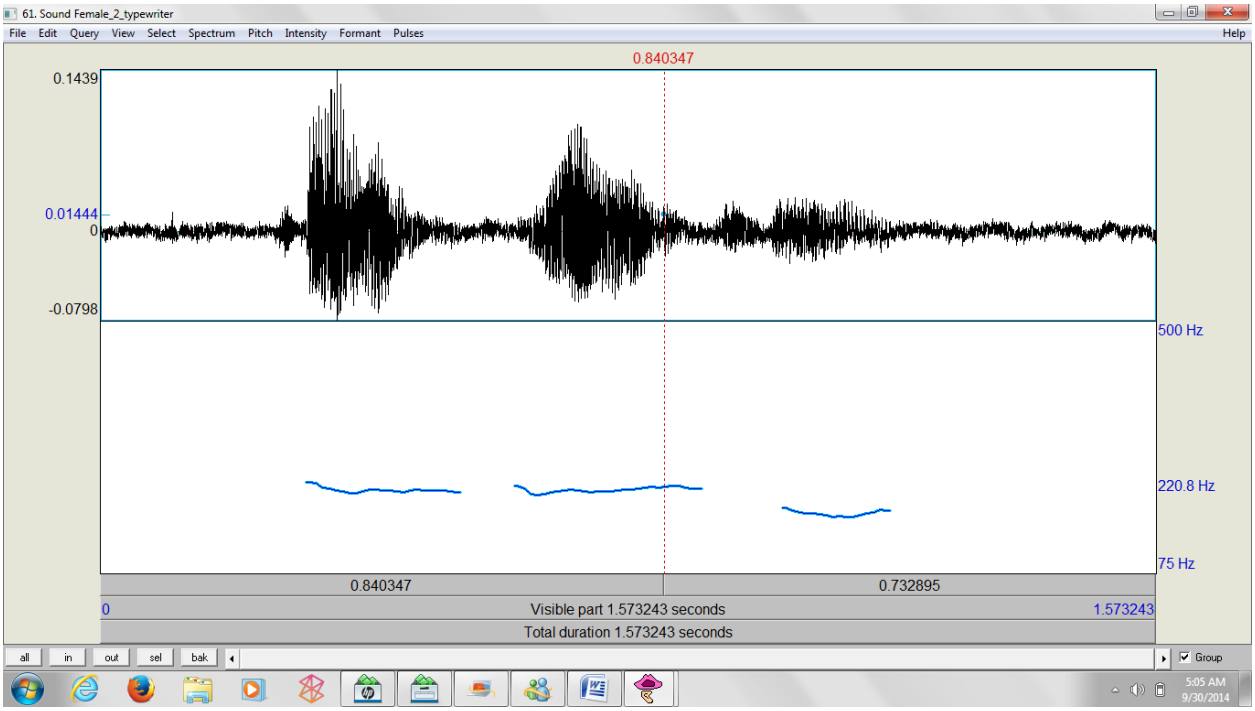


Figure 8(v) Sound wave and *Fo* and time duration of Female 2 utterance of *typewriter* ['taɪpˌraɪ.tə(r)]

Female 2 Consultant utters *typewriter* in the periodic cycles of 0.840347 secs and 0.732895 secs, both totalling 1.573243 secs. The highest pitch value is 220.8 Hz. Although the word is presented in three syllables, the first two syllables receive more prominence than the third syllable with the primary stress on the first syllable.

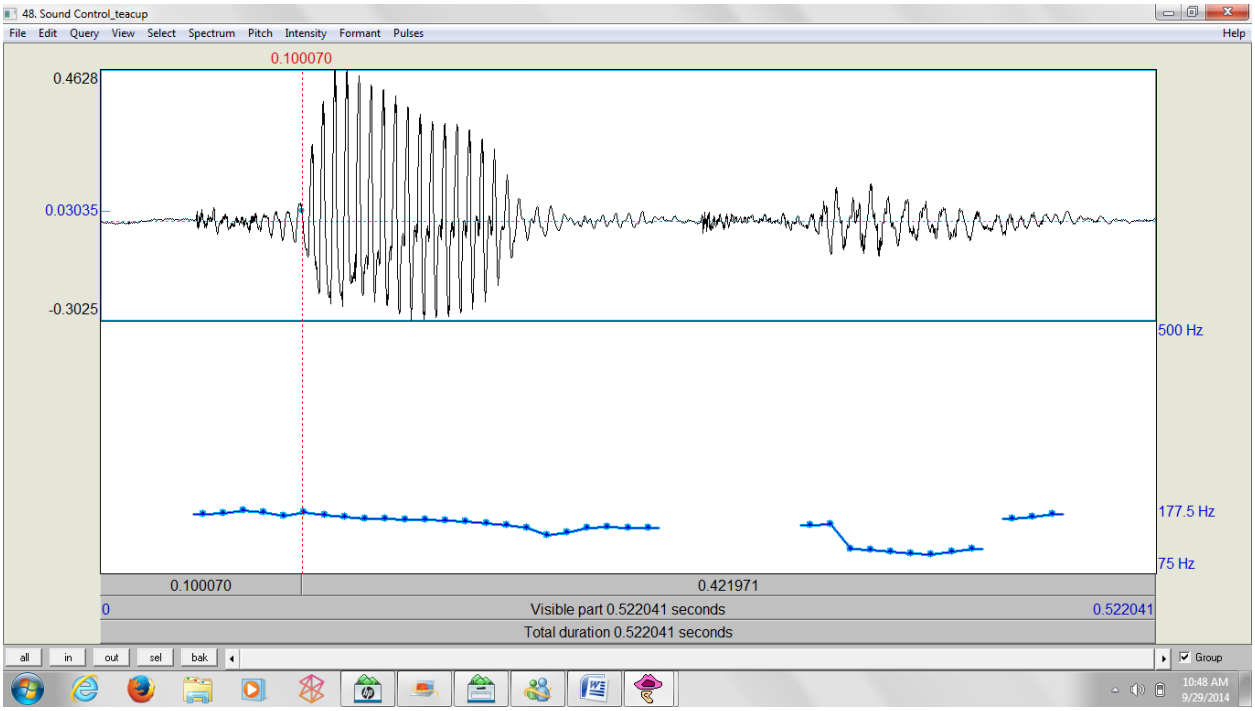


Figure 9(i) Sound wave and *Fo* and time duration of Control utterance of *tea-cup* [ˈti:kʌp]

The compound *tea-cup* consists of two nominals and is a two-syllable word. The first syllable is the heavy syllable and therefore bears the primary stress. In the Control utterance of the word, the stressed syllable is clearly more prominent, while there appears to be the delayed release of the last segment /p/. The utterance is produced in the periodic cycles of 0.100070 secs and 0.421971 secs. The total time duration is 0.522041 secs and the pitch value is 177.5 Hz.

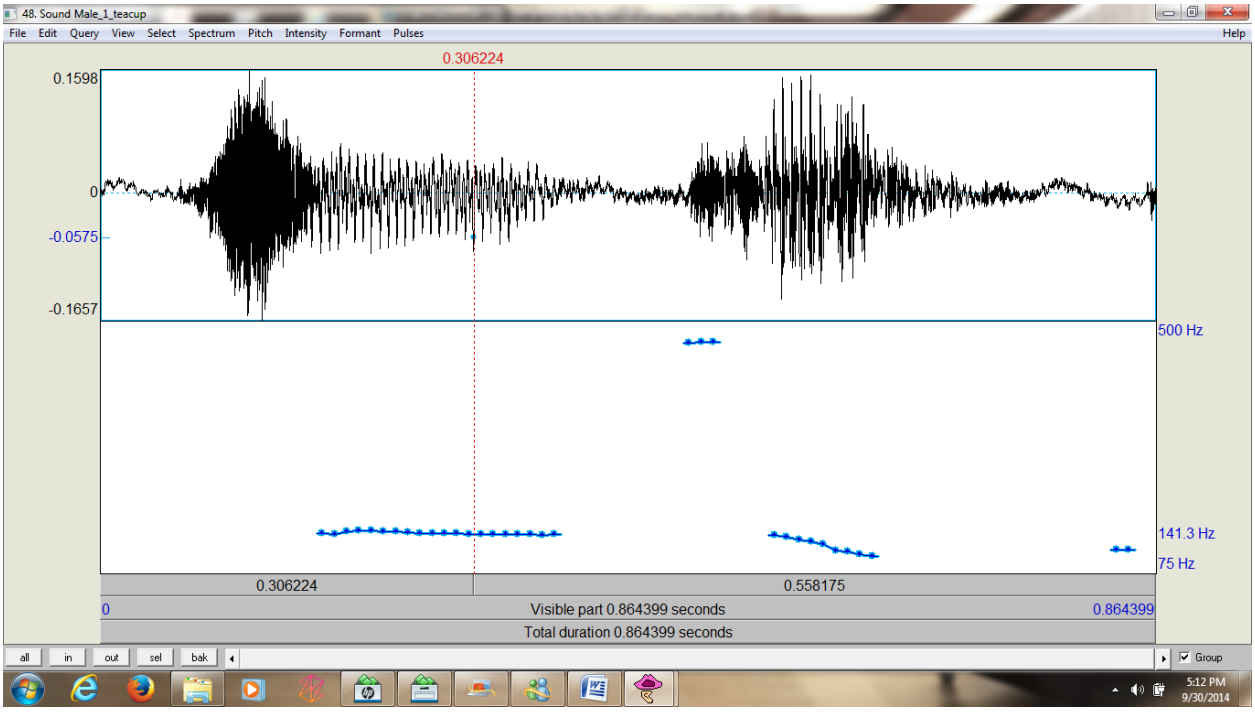


Figure 9(ii) Sound wave and *Fo* and time duration of Male 1 utterance of *tea-cup* [ˈti:kʌp]

Male 1 utterance of *tea-cup* was done in periodic cycles of 0.306224 secs and 0.5558175 secs. The total time duration is 0.864399 secs; and the pitch value is141.3 Hz. The pitch contour presents the first and second syllables as the primary and secondary stress respectively. There is also evidence of the delayed release of the final segment, /p/.

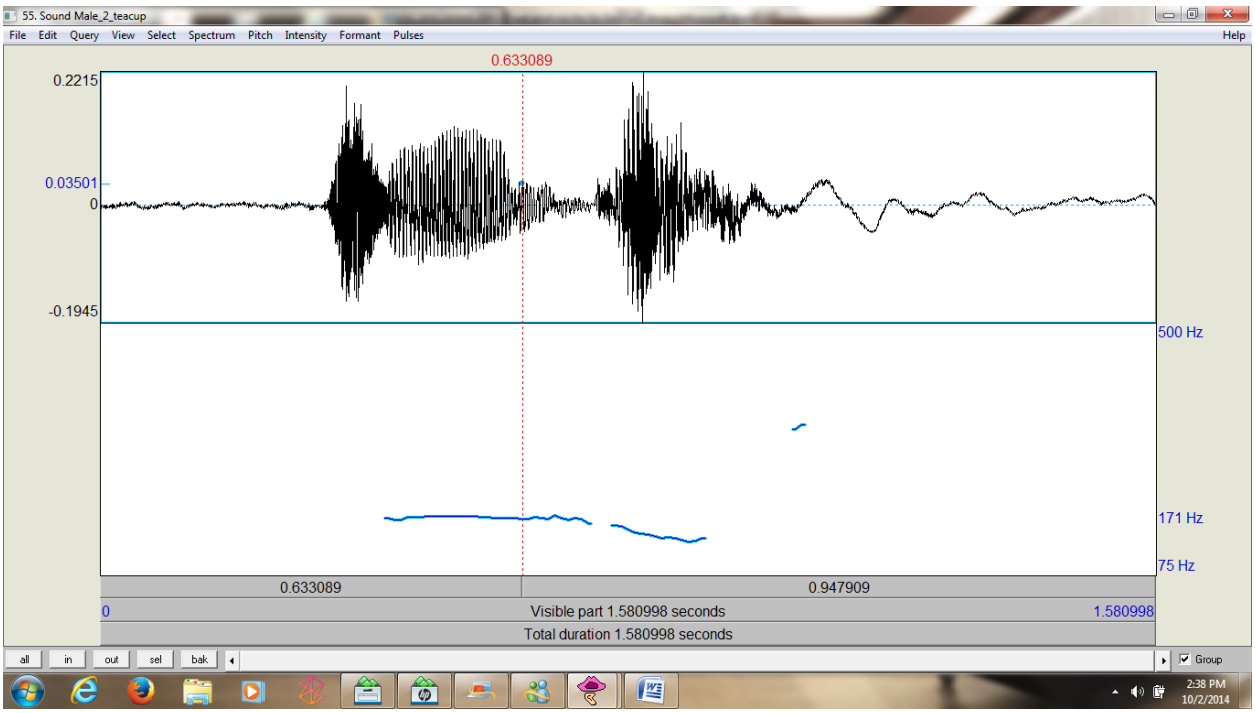


Figure 9(iii) Sound wave and *Fo* and time duration of Male 2 utterance of *tea-cup* [ˈtiːkʌp]

Male 2 presented *tea-cup* in periodic cycles of 0.633089 secs and 0.947909 secs, giving a total time duration of 1.580998 secs. The pitch value is 171 Hz. The pitch contour indicates the two-sllable structure of the compound, with the first syllable receiving more prominence and bearing the primary stress.

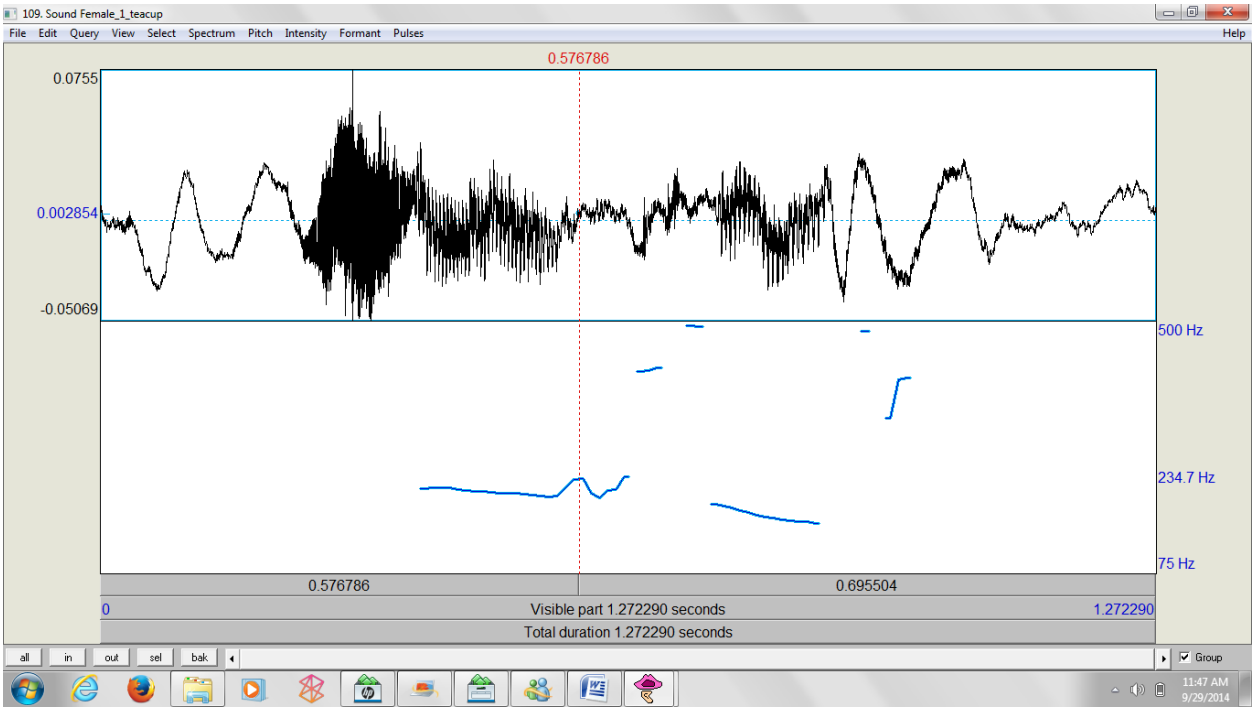


Figure 9(iv) Sound wave and *Fo* and time duration of Female 1 utterance of *tea-cup* [ˈti:kʌp]

Female 1 presents *tea-cup* in periodic cycles of 0.576786 secs and 0.695504 secs. The total time duration is 1.272290 secs, and the pitch value is 234.7 Hz. The pitch contour also indicates the two-syllable structure and stress pattern of the compound, which assign primary stress to the first syllable.

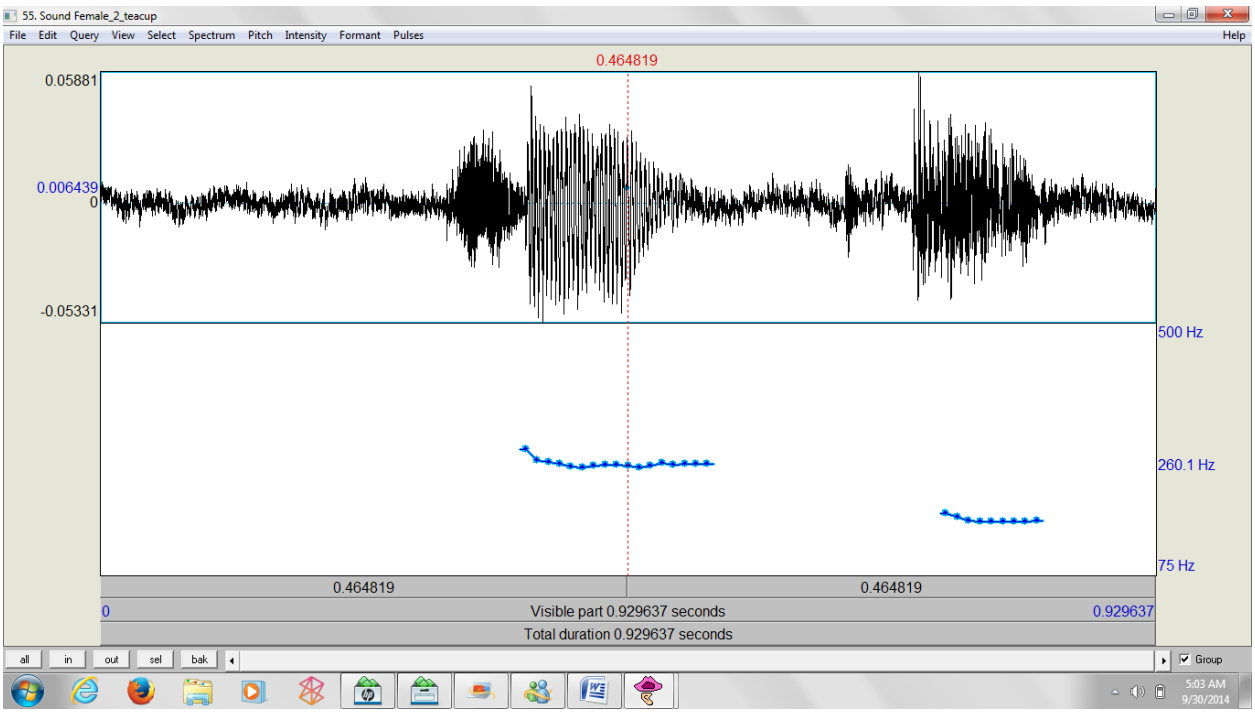


Figure 9(v) Sound wave and *Fo* and time duration of Female 2 utterance of *tea-cup* [ˈti:kʌp]

Female 2 uttered the compound word in periodic cycles of 0.464819 secs and 0.464819 secs summing up to 0.929637 secs. The pitch value is 260.1 Hz. Again, the pitch contour clearly presents the two-syllable structure of the compound, with the primary stress on the first syllable.

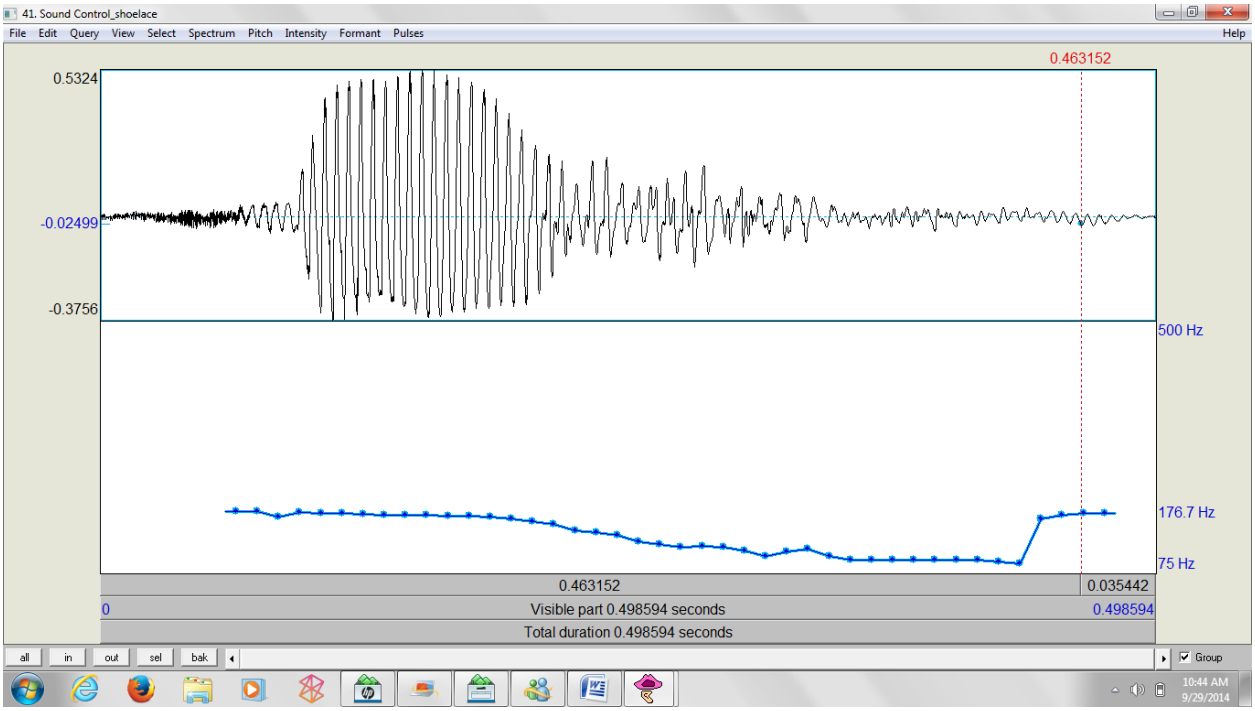


Figure 10(i) Sound wave and *Fo* and time duration of Control utterance of *shoelace* ['ʃu:leɪs]

The compound word, *shoelace* is produced in periodic cycles of 0.463152 secs and 0.035440 secs, with a total time duration of 0.498594 secs. The highest pitch value is 176.7 Hz. Clearly, the first syllable carries a greater prominence than the second with the contour indicating a rise on the second element.

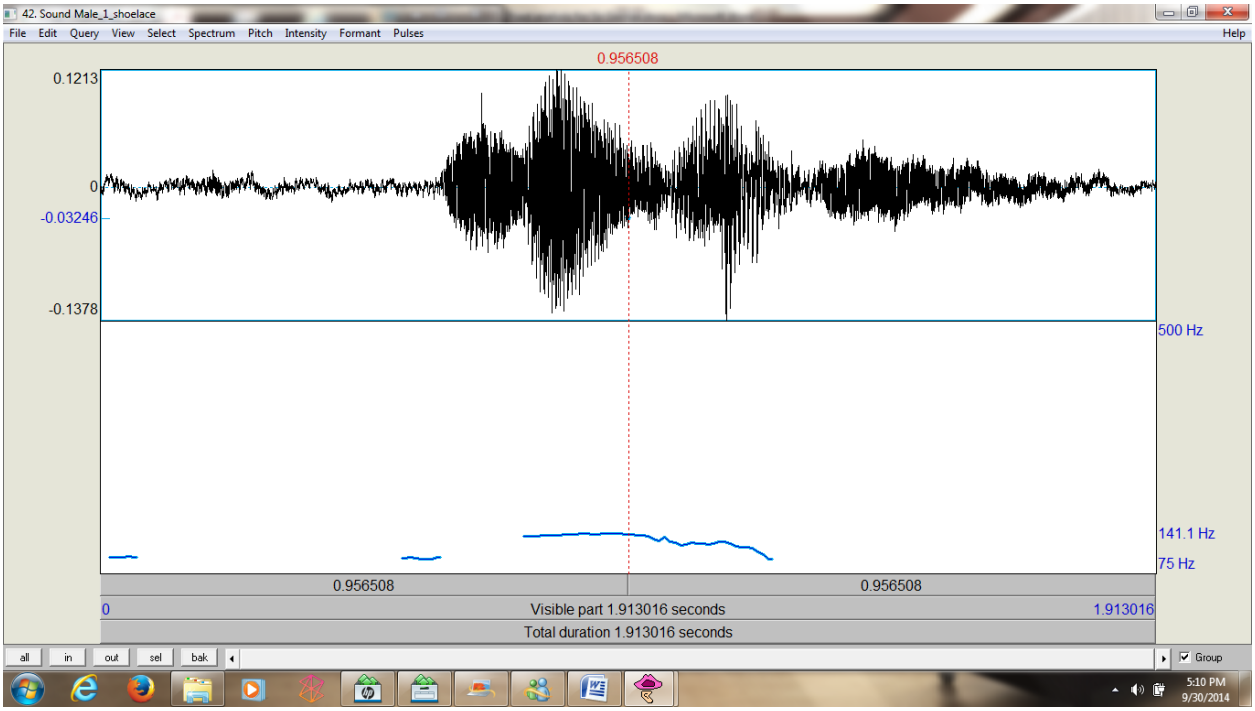


Figure 10(ii) Sound wave and *Fo* and time duration of Male 1 utterance of *shoelace* ['ʃu:leɪs]

The Male 1 utterance of the compound was done in periodic cycles of 0.956508 secs and 0.956508 secs. The total time duration is 1.913016 secs, while the pitch value is 141.1 Hz. The pitch contour, however, presents more prominence on the first element and a downward movement on the second element indicating non observance of the fall-rise signalled by the diphthong at the end of the word.

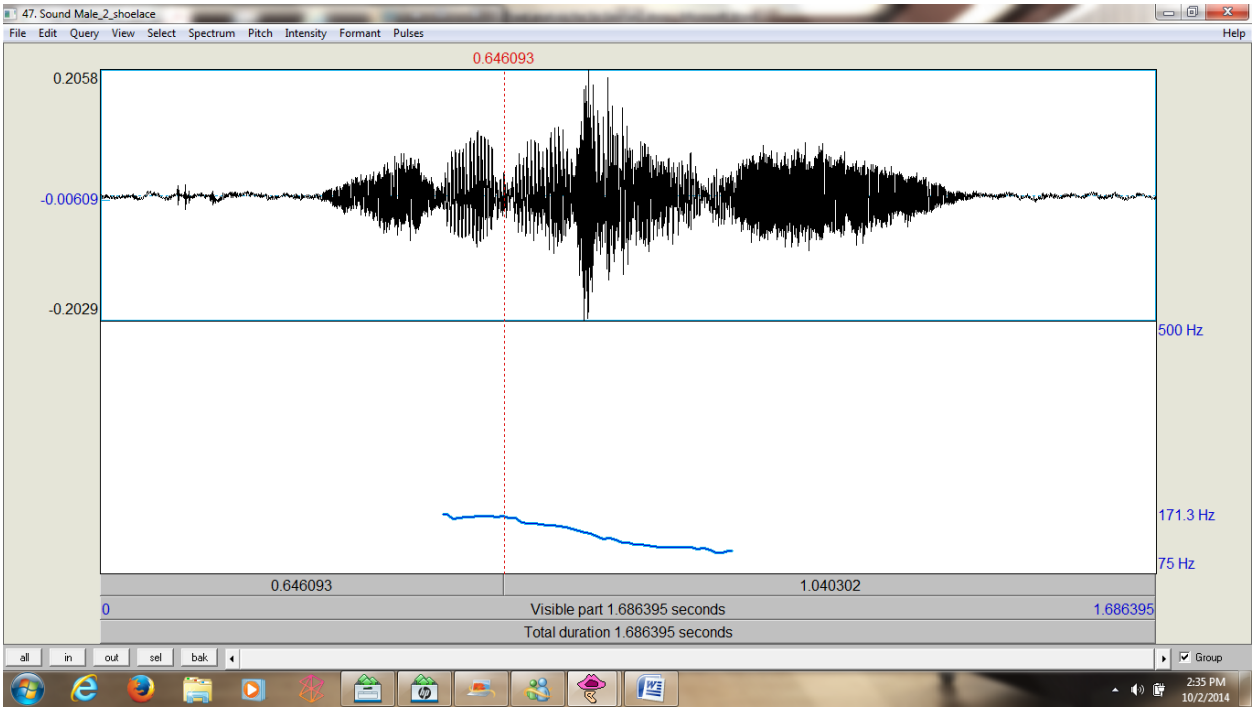


Figure 10(iii) Sound wave and *Fo* and time duration of Male 2 utterance of *shoelace* ['ʃu:leɪs]

The periodic cycles used in uttering *shoelace* are 0.646093 secs and 1.040302 secs. The total time 1.686395 secs; and the pitch value is 171.3 Hz. The pitch contour rather indicates a glossed articulation of the syllables, although there seems to be a portrayal of stress at the inception of the utterance and a downward movement of the pitch at the end.

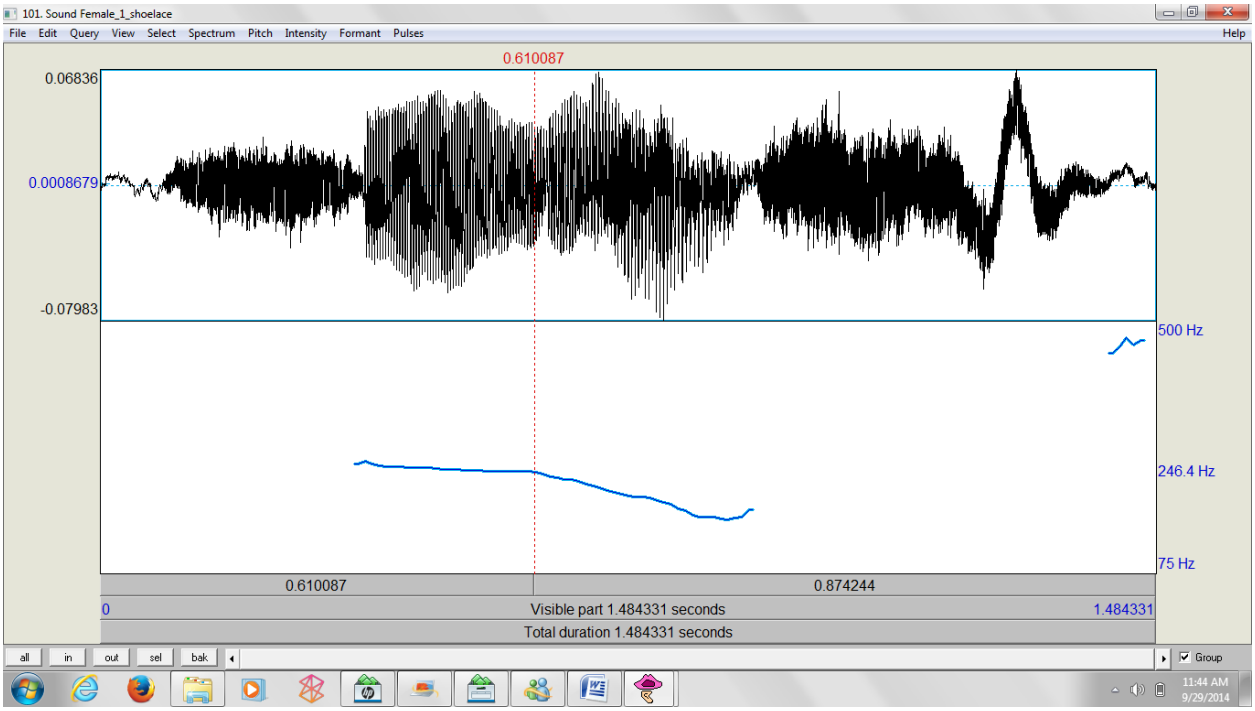


Figure 10(iv) Sound wave and *Fo* and time duration of Female 1 utterance of *shoelace* ['ʃu:leɪs]

Female 1 uttered *shoelace* in periodic cycles of 0.610087 secs and 0.874244 secs. The total time duration is 1.0484331 secs, while the pitch value is 246.4 Hz. As in other Consultants, the pitch contour shows one stretch articulation, and it also begins on a high pitch which drops and rises perhaps in observance of the diphthong.

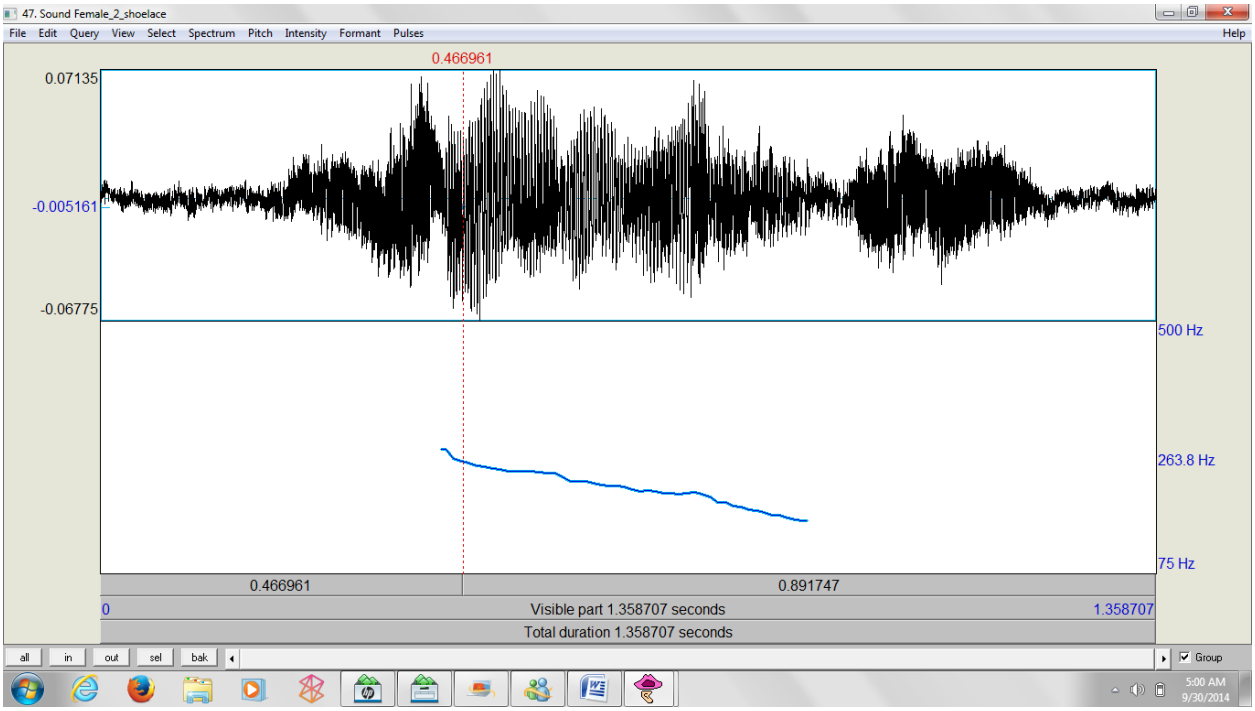


Figure 10(v) Sound wave and *Fo* and time duration of Female 2 utterance of *shoelace* ['ʃuːleɪs]

Female 2 presented *shoelace* in periodic cycles of 0.466961 secs and 0.891747 secs. The total time duration is 1.358707 secs. The highest pitch value measures 263.8 Hz. The pitch contour presents a straight string of utterance, which begins on a high pitch and gradually drops which actually indicates MT interference as Igbo does not have diphthongs.

Table 1: Spreadsheet of result of acoustic correlates of utterances of word stress

	CONTROL		CONSULTANT							
			MALE 1		MALE 2		FEMALE 1		FEMALE 2	
	Pitch value (Hz)	Duration (ms)	Pitch value (Hz)	Duration (ms)	Pitch value (Hz)	Duration (ms)	Pitch value (Hz)	Duration (ms)	Pitch value (Hz)	Duration (ms)
Two-syllable words										
i. record (v)	225	0.798013	241.7	2.391270	171.5	1.686395	222	1.825193	230.1	0.934830
ii. complete (adj.)	184.1	0.779342	132.1	0.197392	162.9	1.264807	218.5	1.200771	224	1.028322
iii. below (adj.)	172.6	0.757007	120.9	2.152132	136.9	1.370181	200.8	0.816531	226.4	1.051701
Three-syllable words										
i. interrupt (v)	166.4	0.730249	126	2.391270	162.3	1.580998	194.4	1.344875	226.8	1.262018
ii. disaster (n)	164.4	0.599932	142.8	2.630408	159.3	1.370181	186.1	1.440930	223.5	1.472381
iii. transistor (n)	155.8	0.788050	205.3	1.913016	120.1	1.475578	191.8	0.912585	254	1.121814
Compound words										
i. second-class	175	0.895193	169.2	2.391270	160.2	1.686372	191.8	1.696417	223.7	1.737846
ii. typewriter	177.7	0.759320	132.3	1.061859	151.5	1.791769	206.1	1.378322	220.8	1.573243
iii. tea-cup	177.5	0.522041	141.3	0.864399	171	1.580998	244.1	0.435193	260.1	0.929637
iv. shoelace	176.7	0.498594	141.1	1.913016	171.3	1.686395	246.4	1.484331	263.8	1.358707

4.1.1(b) Summary of acoustic analysis on stress in words

(i) In the two-syllable words, the Control and most Consultants' utterances indicate the two-syllable structure of the words. Specifically, Male 1, Male 2 and Female 2 utterances of *record* present fall-rise pitch contours reminiscent of the verb with the second syllable receiving more prominence. The total time duration spent by the Control and Female 2 are more closely related while the highest pitch value for the Control, Male 1, Female 1 and Female 2 fall within the same range of between 222 Hz to 241.7 Hz. However, only the pitch value of Male 2 measures below 200 Hz, but it still falls within the normal F*o* range of 80 Hz to 210 Hz highlighted by Ashby & Maidment (2005) and Clark, Yallop and Fletcher (2007) for adult male English speaker. Also, the pitch value of 241.7 Hz recorded by Male 1, and 231.1 Hz recorded by Female 2 defeat the claim that the females naturally have higher pitches.

The adjectives *complete* and *below* have similar pitch contour, which starts somewhere high to a very low pitch before an upward movement that ends somewhere at the mid. The pitch contour of all the Consultants, with the exception of Male 1 and Male 2 are similar to that of the Control. The same is the case with the utterances for *below* with very slight differences though. The pitch value for the Control, Male 1 and Male 2 each measures below 200 Hz. thus signifying attitudinal differences in individuals.

(ii) For the three syllable words, *interrupt (v), disaster (n)* and *transistor (n),* the pitch contour of the English and the Consultants resemble except the slight variations of rise at the end of the first syllable for Female 1 and rise at the end of Male 2 and Female 1 utterances of *interrupt.* Also, Female 1 and Female 2 produced the word at a higher pitch as reflected in their pitch values of 194.4 Hz and 226.8 Hz respectively as against 166.4 Hz, 126 Hz and 162.3 Hz for the Control, Male 1 and Male 2 respectively. Again, in *disaster* and *transistor,* the pitch values for Female 1 and Female 2 measure 186.1 Hz and 222.5 Hz; and 191.8Hz and 254 Hz respectively, while the Control, Male 1 and Male 2 have their pitch values measuring 166.4 Hz and 155.8 Hz; 142.8 Hz and 205.3 Hz and 159.3 Hz and 120.1 Hz respectively.

Again, there is a clear distinction between the male and female F*o* ranges in the data. However, the total time duration for the Control utterances of the words are shorter than those of the Consultants. The reason may be because of the nature of the Igbo syllable weight and length for H tone, which is usually long.

(iii) The English utterance for *second-class* and those for Male 2, Female 1 and Female 2 share a lot of similarities in their wave forms and pitch height, except that the pitch value for Female 2 is above 200 Hz and higher than those for Male 1 and Male 2, thus justifying the assertions that females have higher pitches than their male counterparts. Male 1 registers the longest time duration for producing the compound word. Similarly, the English utterance for *typewriter* and those for Male 2 and Female

2 resemble in syllable division, with Male 1 and Male 2 having close resemblance with the pitch value of the Control.

For *tea-cup,* only Male 1 shows evidence of delayed release of the voiceless aspirated bilabial plosive. The pitch value of both Female 1 and Female 2 again measure above 200 Hz as Male 1, Male 2 and the Control measure 141.3Hz, 171 Hz and 177.5 Hz respectively. The total time duration for all utterances falls below 1 second, except that for Male 2 that sums up to 1.580998 secs.

In *shoelace,* Male 2 and Female 1 utterances resemble that of the English in observance of fall-rise tone at the end of the utterance indicative of the diphthong /eɪ/. The other Consultants' utterances of the compound end on fall instead, perhaps, indicating evidence of interference of their native language, Igbo. Generally, they indicate the movement of the pitch contour starting somewhere up and down as is the case in the English utterance.

4.1.2 Stress placement in Sentences/Tonic syllable

Below are sound waves and analysis of stress placement in sentences. In this section, different types of sentence comprising statements and questions are presented. The different pitch patterns that characterize these types of sentence are identified and discussed both from the utterances of the Control and those of the Consultants.

4.1.2(a) Sound waves and pitch contour of sentences with definitive fall

A statement often has a fall as it indicates completeness and tends to signal finality. Usually, the starting point may be anywhere from mid to high, but the end point is low, often initiated by a step up in pitch. This study, however, concentrates on the nuclear tone, which, according to Ashby and Maidment (2005) and Wells (2006) is the most important accent in the IP which also bears the definitive fall and which is very important for the interpretation of an utterance.

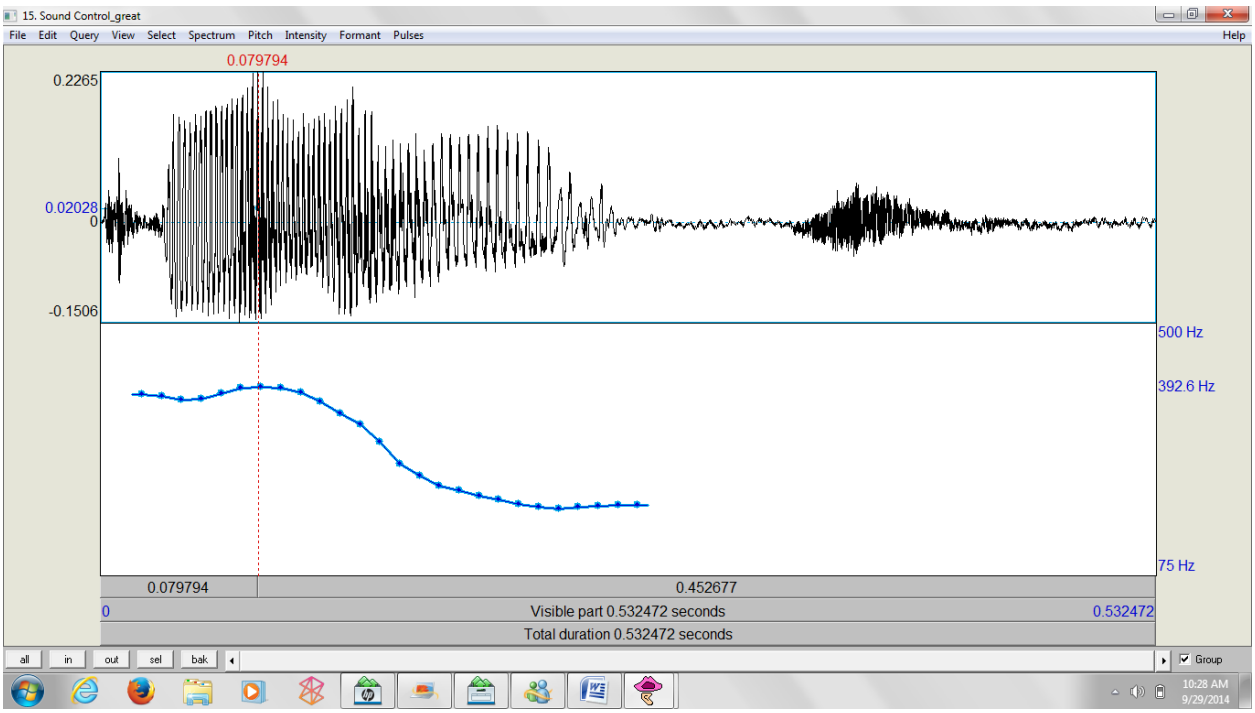


Figure 11(i) Sound wave and *Fo* and time duration of Control utterance of *I 'think it's* ˌgreat. [aɪ 'θɪŋk ɪts ˌgreɪt]

The Control rendition of the nuclear tone of this statement is made in the periodic wave forms of 0.079794 secs. and 0.452677 secs, both totalling 0.532472 secs. The highest pitch value is 392.6 Hz, with the pitch contour clearly indicating a fall at the end.

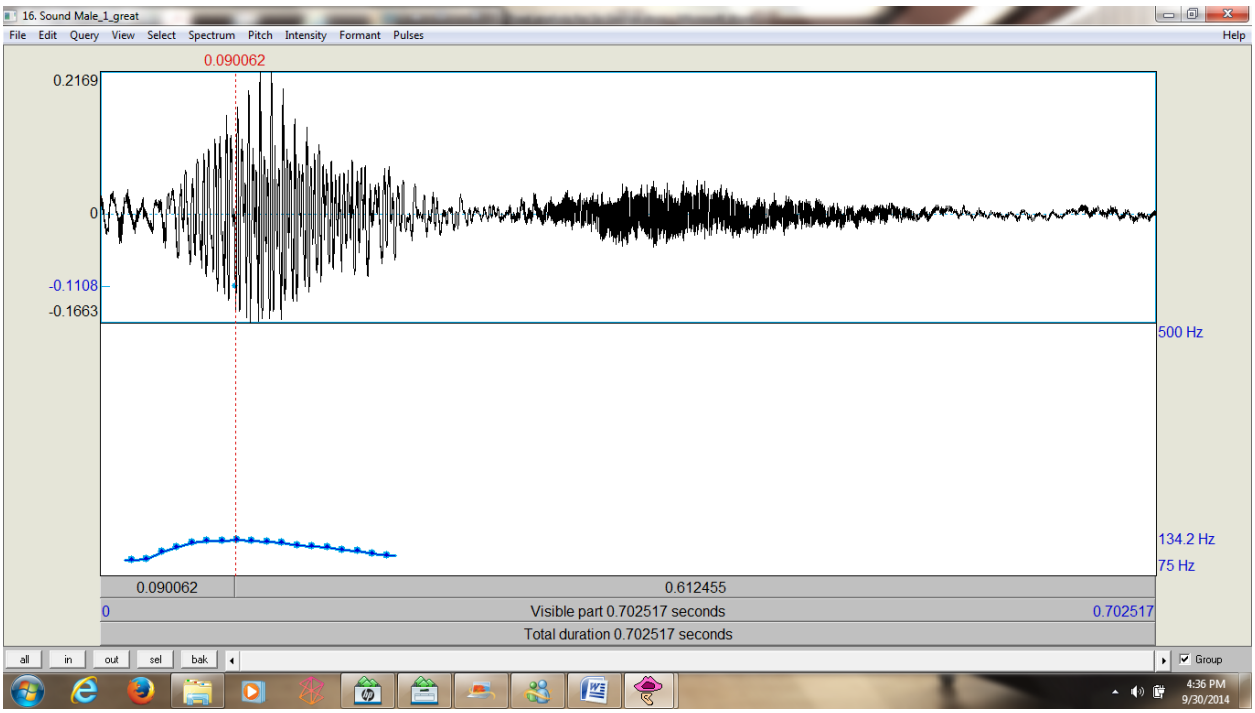


figure 11(ii) Sound wave and *Fo* and time duration of Male 1 utterance of *I 'think it's* ˌgreat. [aɪ 'θɪŋk ɪts ˌgreɪt]

The Male 1 native Igbo speaker produced the nucleus of the statement at the periodic wave form cycles of 0.090062 secs and 0.612455 secs, giving a total time duration of 0.702517 secs. The highest pitch value is 134.2 Hz. The pitch contour equally presents a step up to high, followed by a fall, although the pitch level is not as high as that of the Control.

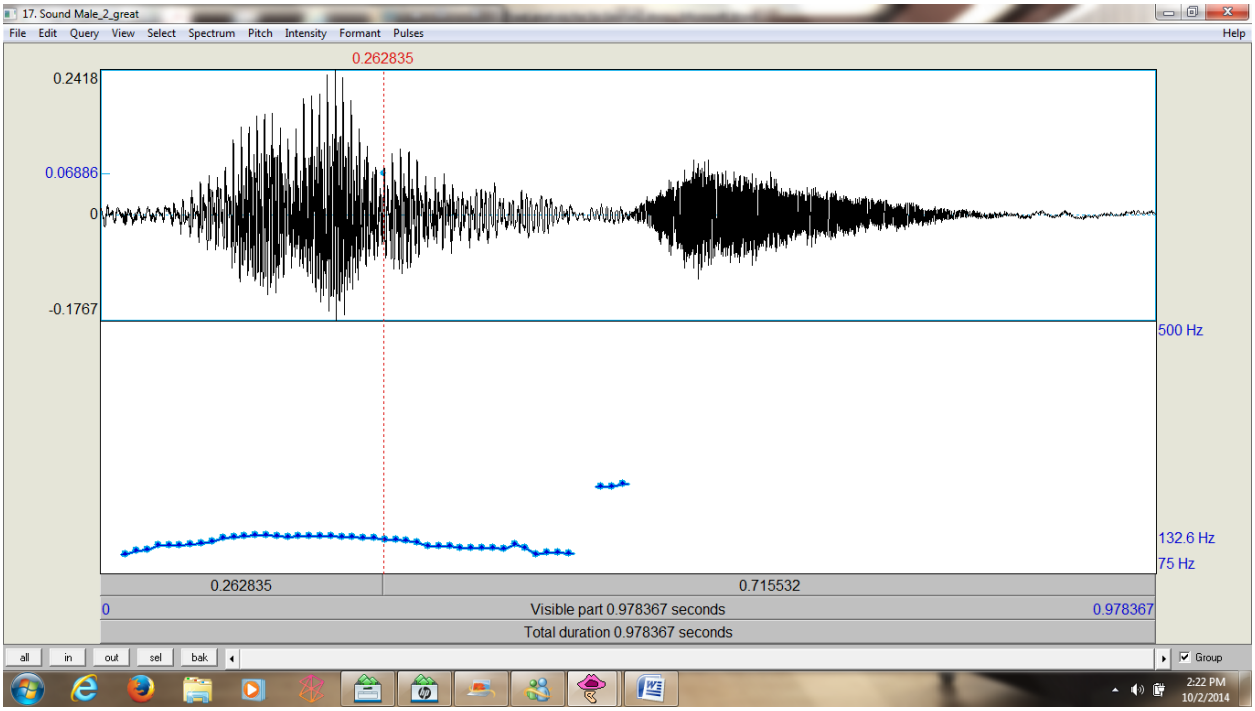


Figure 11(iii) Sound wave and *Fo* and time duration of Male 2 utterance of *I 'think it's* ˌgreat. [aɪ 'θɪŋk ɪts ˌgreɪt]

Male 2 presented the nuclear tone at the periodic wave forms of 0.262835 secs and 0.715532 secs. The total time duration for the production, therefore, is 0.978367 secs, and the pitch value is 132.6 Hz. The pitch contour presents a slight step up followed by a short downward and upward movement, although the starting point of the step up is not as high as that of the Control.

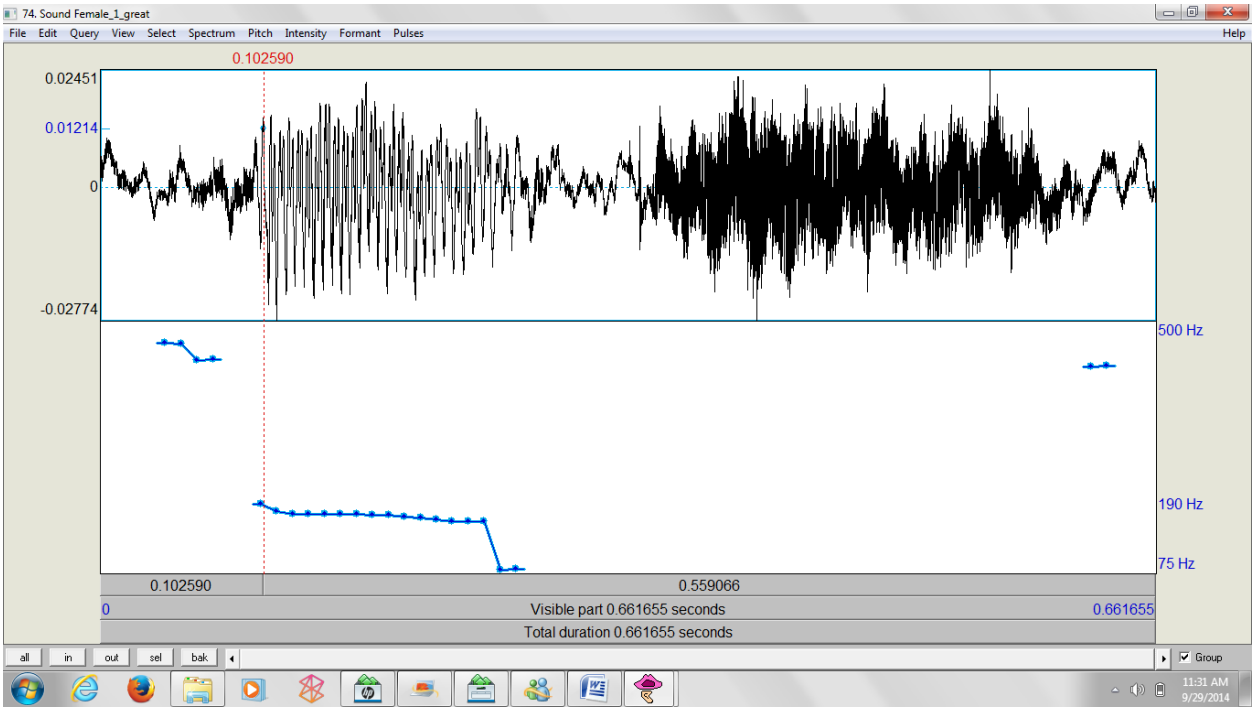
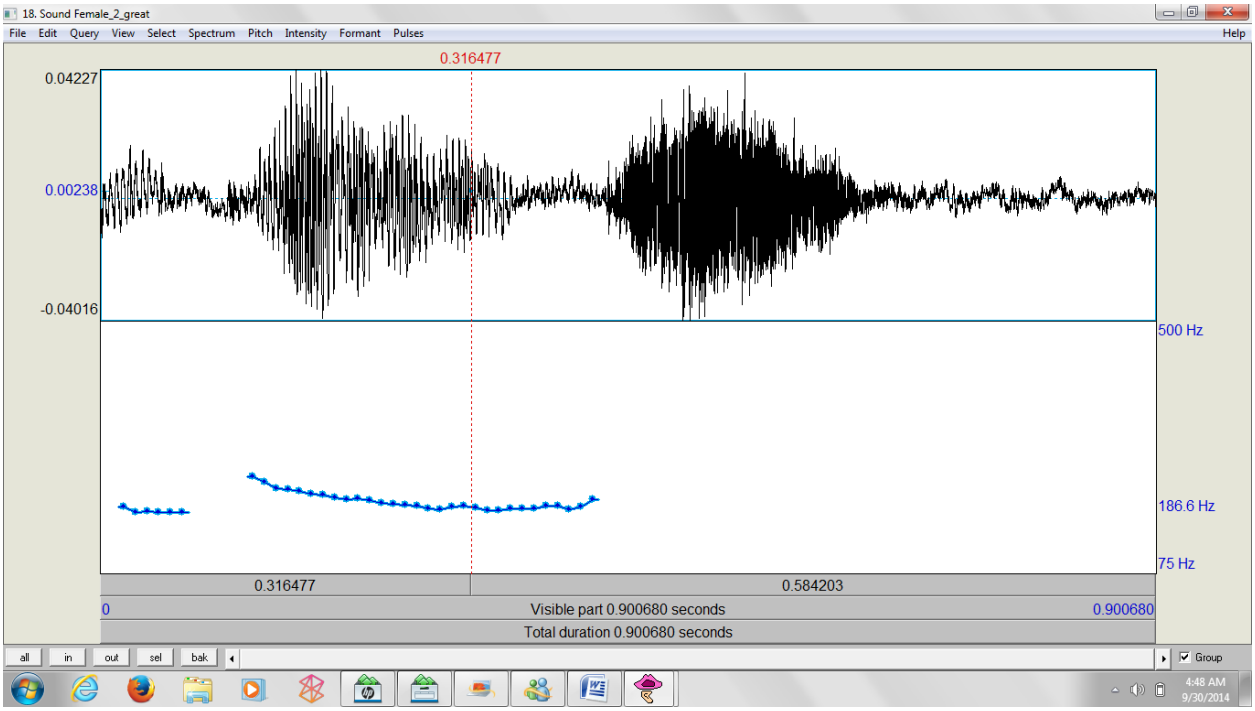


Figure 11(iv) Sound wave and *Fo* and time duration of Female 1 utterance of *I 'think it's ˈgreat.* [aɪ 'θɪŋk ɪts ˈgreɪt]

Female 1 utterance was made at the periodic wave form cycles of 0.102590 secs and 0.559066 secs, both giving the total time duration of 0.661655 secs. The pitch value is 190 Hz. The pitch contour, however, presents a sharp steep fall and a short rise.



(v) Sound wave and *Fo* and time duration of Female 2 utterance of *I 'think it's ˈgreat.* [aɪ 'θɪŋk ɪts ˈgreɪt]

Female 2 rendered the nuclear tone at the periodic wave form cycles of 0.316477 secs and 0.584203 secs. The total time duration sums up to 0.900680 secs, while the highest pitch value 186.6 Hz. The pitch contour presents a picture that resembles a fall-rise, perhaps indicating the diphthong in the nucleus.

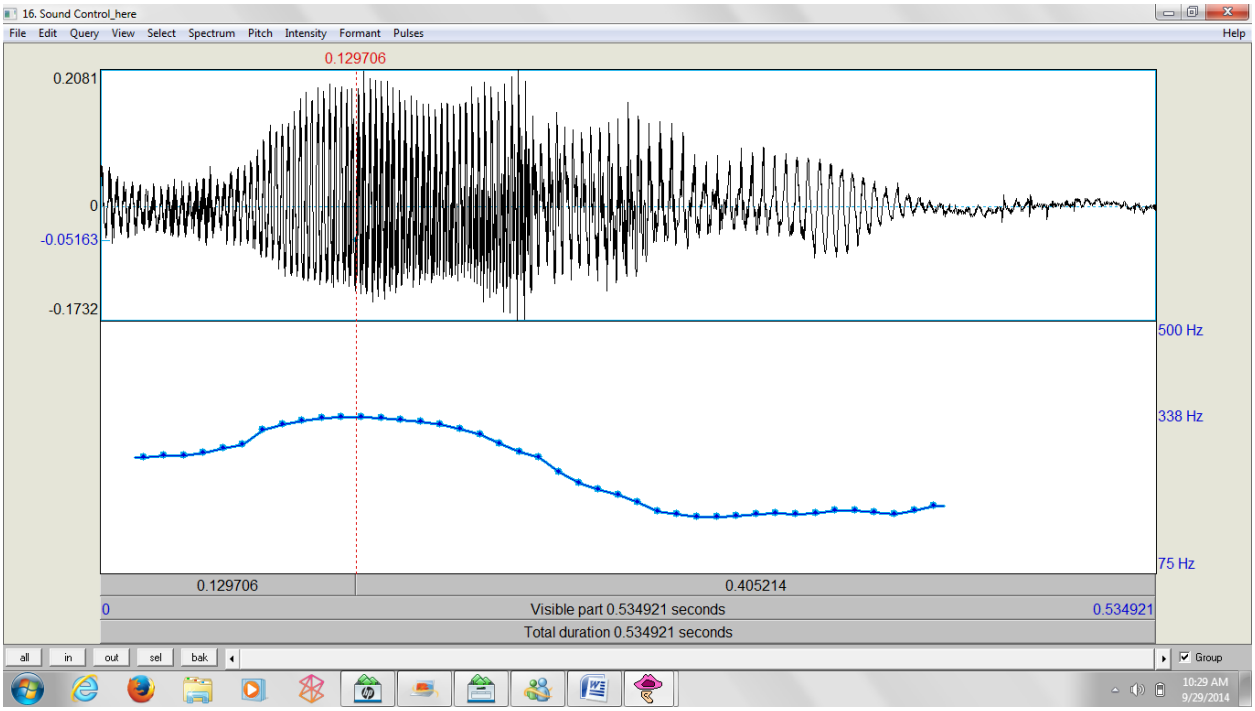


Figure 12(i) Sound wave and *Fo* and time duration of Control utterance of *We're 'all ˌhere.* [wɪəʳ ˈɔːl ˌhɪə]

The Control rendition of the nuclear tone of the statement is made at the periodic wave form cycles of 0.129706 secs and 0.405214 secs. The total time duration is 0.534921 secs while the pitch value is 338 Hz. The pitch contour here clearly indicates a step up to high followed by a short fall.

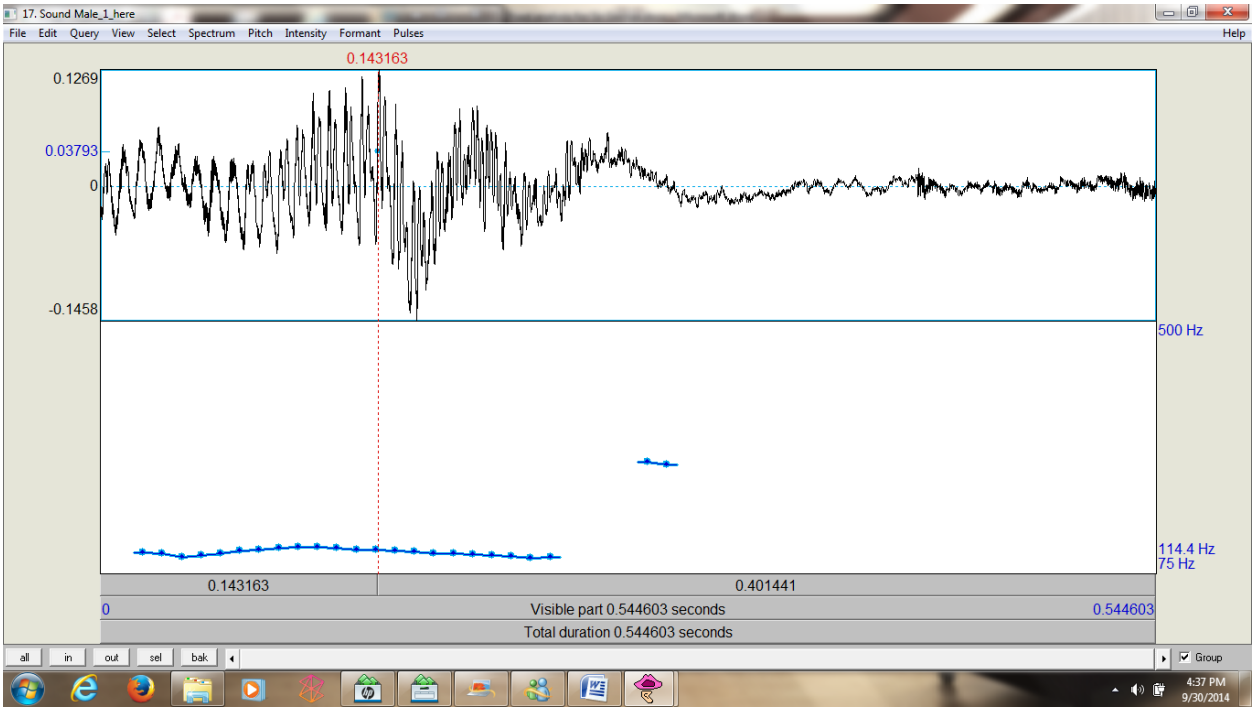


Figure 12(ii) Sound wave and *Fo* and time duration of Male 1 utterance of *We're 'all* ˌ*here.* [wɪəʳ ˈɔːl ˌhɪə]

Male 1 rendered the nuclear tone at the periodic wave forms of 0.143163 secs and 0.401441 secs, giving the total time duration as 0.544603 secs. The pitch value is 114.4 Hz. The pitch contour presents a step up flow which moved down to a fall. However, this was made on a low pitch.

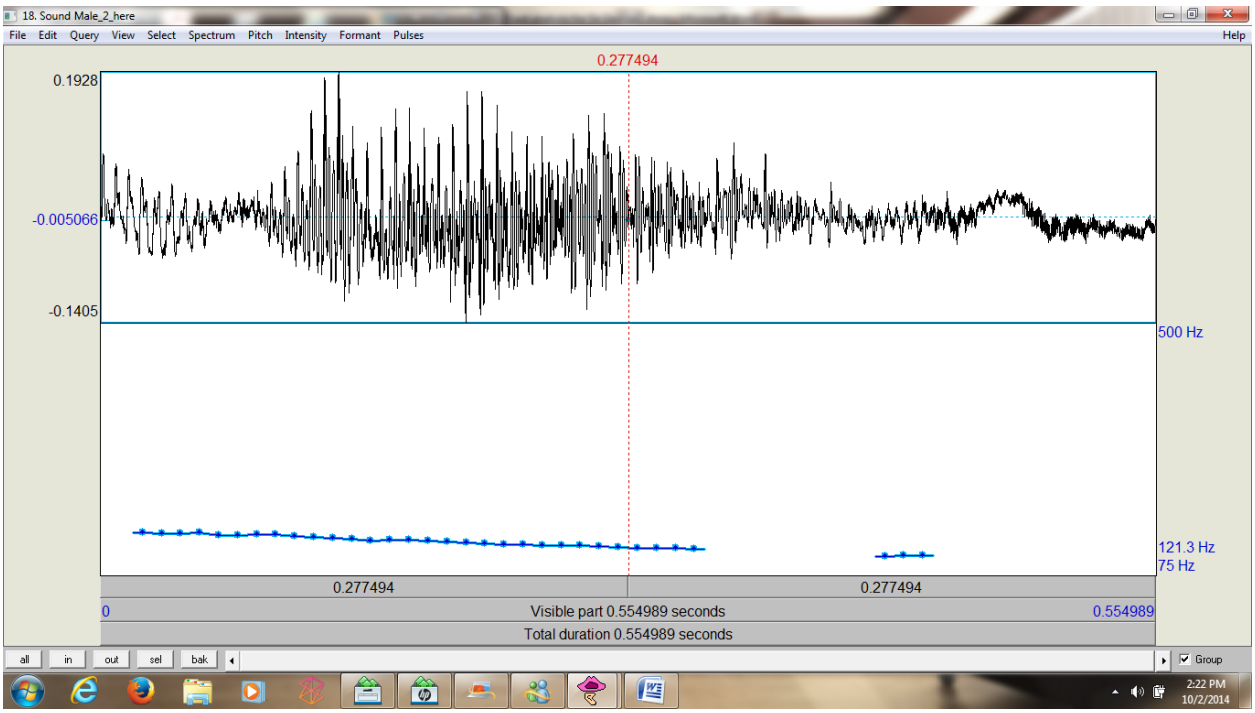


Figure 12(iii) Sound wave and *Fo* and time duration of Male 2 utterance of *We're 'all* ˌ*here.* [wɪəʳ ˈɔːl ˌhɪə]

Male 2 rendition appears to be made on a pitch contour from up straight down. The periodic wave form cycles used in the production are 0.277494 secs and 0.277494 secs, thus giving the total time duration as 0.554989 secs. The pitch value for the utterance is 121.3 Hz.

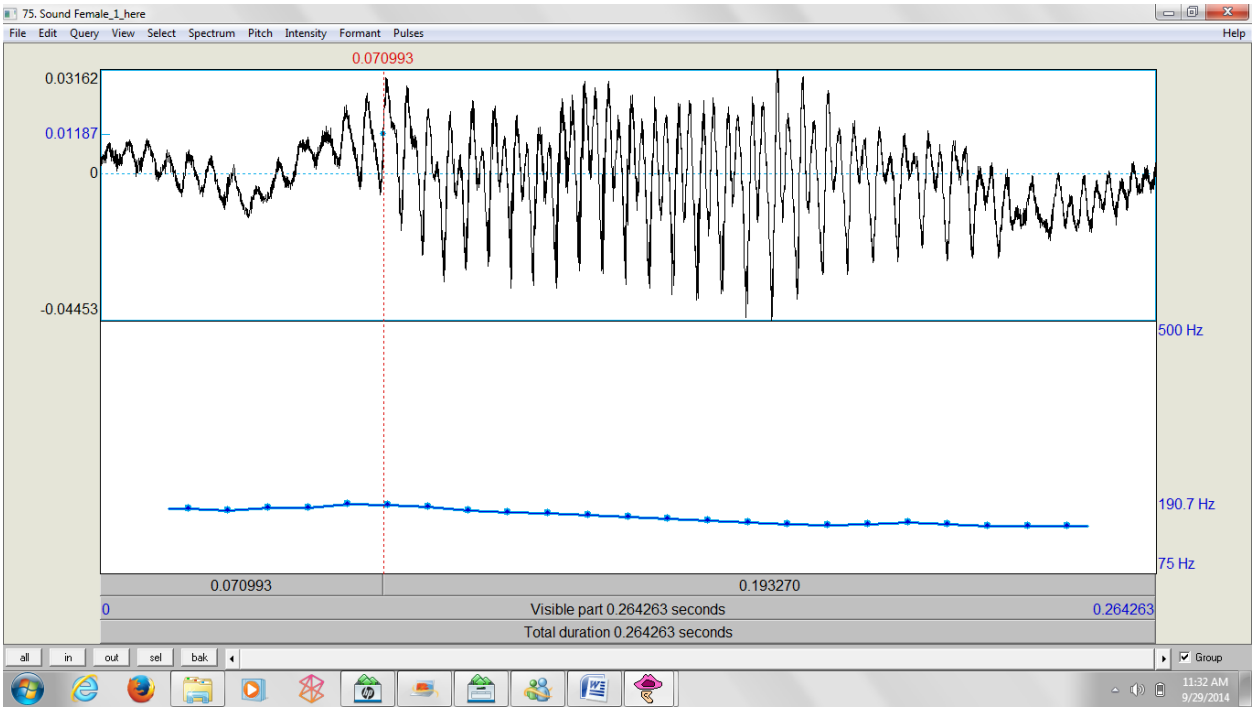


Figure 12(iv) Sound wave and *Fo* and time duration of Female 1 utterance of *We're 'all ˈhere.* [wɪəʳ ˈɔːl ˌhɪə]

Female 1 produced the nuclear tone on a long stretch pitch contour from somewhere low up slightly down. The periodic wave form cycles for the production are 0.070993 secs and 0.193270 secs, giving the total time duration as 0.26463 secs. The highest pitch value is 190.7 Hz.

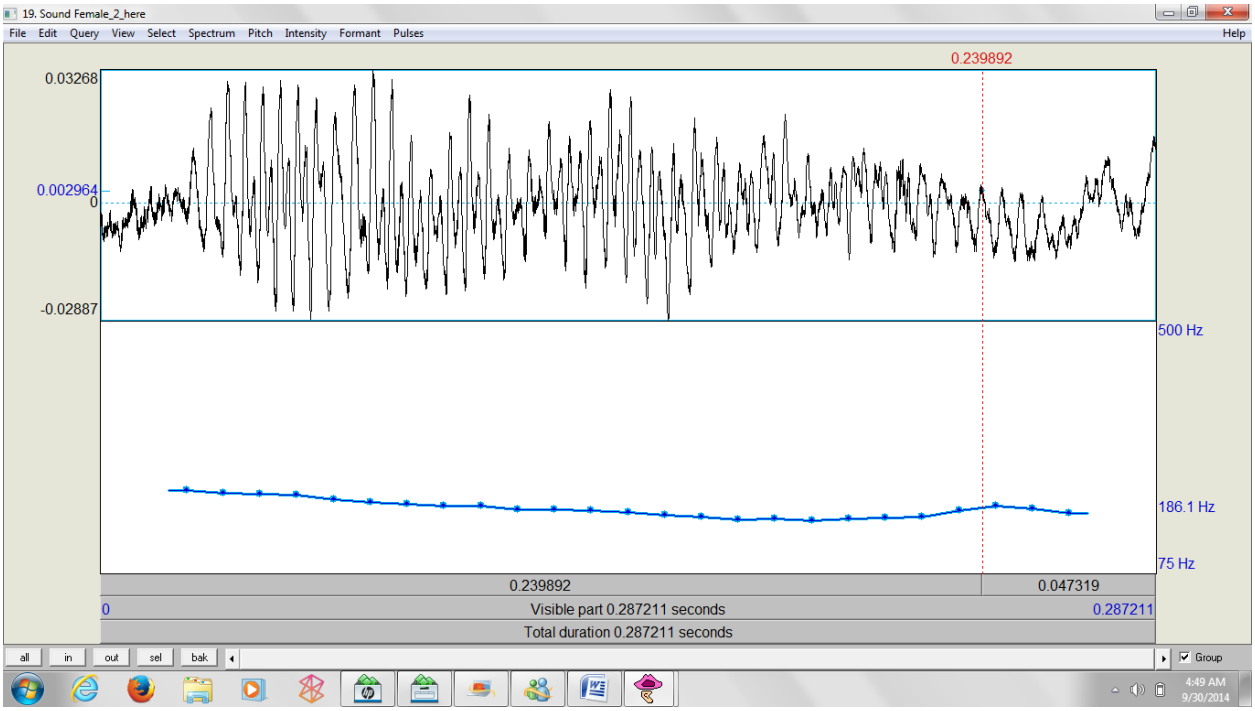


Figure 12(v) Sound wave and *Fo* and time duration of Female 2 utterance of *We're 'all ˌhere.* [wɪəʳ ˈɔːl ˌhɪə]

This a another long stretch pitch contour like the one in Fig 12(iv). It moves from up and goes down to what appears to be a level fall. The consultant produced the tone at the wave forms of 0.239892 secs and 0.047319 secs. The total time duration is 0.28711 secs while the pitch value is 186.1 Hz.

Table 2 Spreadsheet of the acoustic correlates of statements with definitive fall

	CONTROL		CONSULTANTS							
			MALE 1		MALE 2		FEMALE 1		FEMALE 2	
	Pitch value (Hz)	Duration (ms)	Pitch value (Hz)	Duration (ms)	Pitch value (Hz)	Duration	Pitch value (Hz)	Duration (ms)	Pitch value (Hz)	Duration (ms)
(i) I 'think it's ˏgreat.	392.6	0.532472	134.2	0.702517	132.6	0.978367	190	0.661655	186.6	0.900680
(ii) We're 'all ˏhere.	338	0.534921	114.4	0.544989	121.3	0.554603	190.7	0.264263	186.1	0.287211

4.1.2(b) Summary of acoustic analysis of pitch pattern in sentences with definitive fall

It will be recalled that the nuclear tone is the most important accent in the Intonation Phrase (IP). It is also very important in the interpretation of an utterance. Furthermore, it may need to be restated that the default tone for statements is mostly fall, tagged "definitive fall". In the data on statements with definitive fall, the English and native Igbo speakers' utterances all indicate movement from somewhere up to fall. However, the pitch levels of the Igbo speakers' utterances are not as high as that of the English. Furthermore, all but Male 1 observed the diphthong in *great*, though with variations. The pitch value of Male 1 and Male 2 of 134.2 Hz and 132.6 Hz; and those for Female 1 and Female 2 of 190 Hz and 186.6 Hz respectively are each pair very closely related, but greatly at variance with the Control, which measures 392.6 Hz. Also, Male 1 time duration of 0.702517 secs, Male 2 duration of 0.978367secs, and Female 2 duration of 0.900680 secs for the production of the nucleus *great* are widely at variance with that of the Control, while only Female 1 time duration of 0.661655 secs is very close to the Control time duration of 0.532472 secs.

Similarly, the nuclear tone for *We're all* ˏ*here* for all the utterances indicate a fall although the pitch levels for all the Consultants are lower than that of the Control. Consequently, the pitch value

for Control is 338 Hz and those for Male 1, Male 2, Female 1 and Female 2 are 114.4 Hz, 121.2 Hz, 190.7 Hz and 186.6 Hz respectively. The total time duration for Control, Male 1 and Male 2 utterances of 0.534921 secs, 0.544989 secs and 0.554603 secs respectively are more closely related than 0.264263 secs and 0.287211 secs for Female 1 and Female 2.

4.1.3(a) Sound waves of tentative statements

A tentative statement is that which implies some kind of uncertainty or non commitment of the speaker to the demand of the message. It is signalled by the fall-rise tone. In a fall-rise tone as indicative of tentative statements, the pitch of the voice starts relatively high, moves first downwards and then upwards again. The starting point may be somewhere from mid to high, the midpoint low and the end point usually mid.

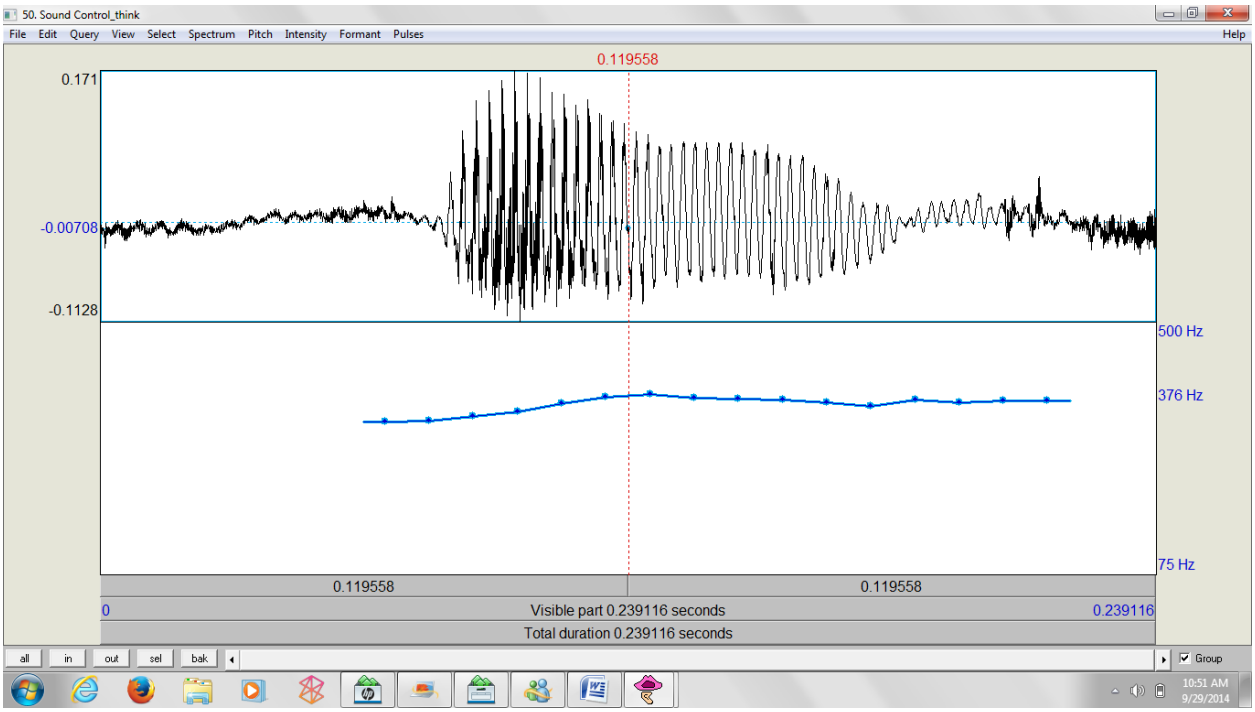


Figure 13(i) Sound wave and *Fo* and time duration of Control utterance of *I ˌthink so.* [aɪ ˌθɪŋk səʊ]

The Control utterance of the nuclear tone of the tentative statement is made at the wave form cycles of 0.119558 secs and 0.119558 secs, giving the total duration as 0.239116 secs. The pitch value is 376 Hz. The pitch contour presents a picture of a relative high pitch at the onset, which moves slightly downwards and then rises slightly again.

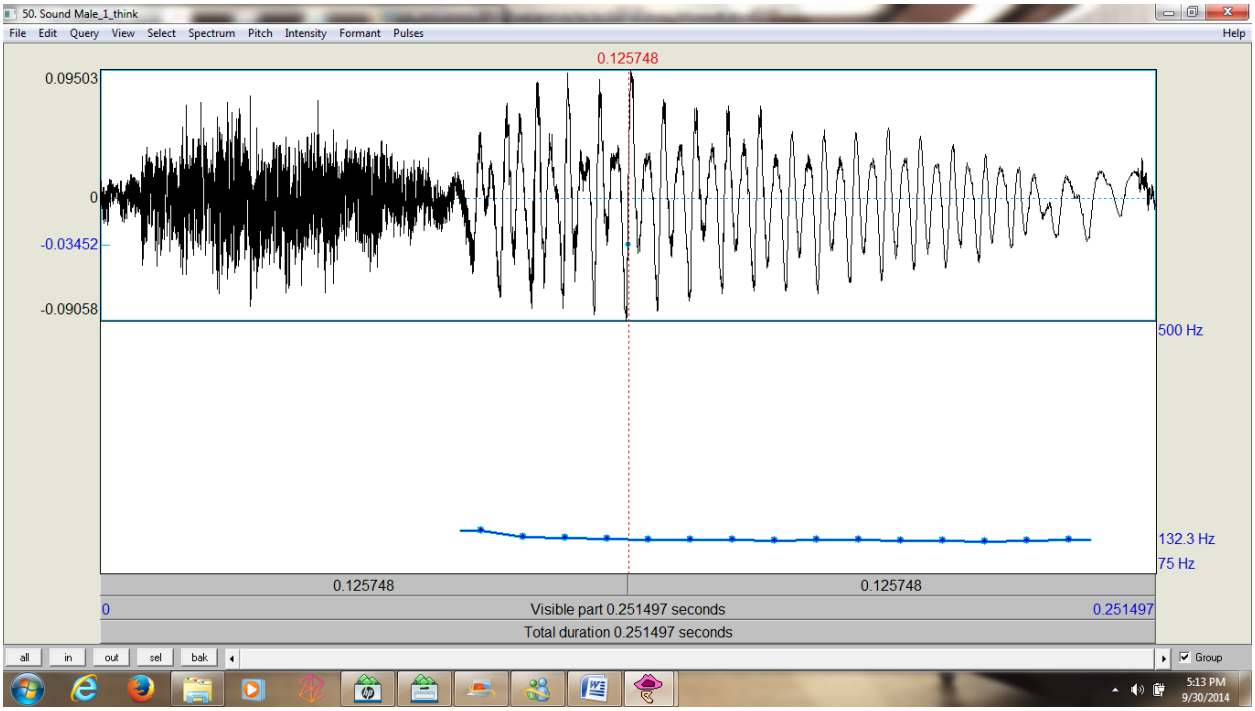


Figure 13(ii) Sound wave and Fo and time duration of Male 1 utterance of *I ˌthink so* [aɪ ˌθɪŋk səʊ]

Male 1 native Igbo speaker produced the nuclear tone of this tentative statement at the periodic wave forms of 0.125748 secs and 0.125748 secs respectively, giving the total duration as 0.251497 secs. The pitch value of the utterance is 132.3 Hz. From the pitch contour, we can see a short step up of pitch at the onset followed by an apparent down and up movement to a long stretch of what appears to be a mid tone.

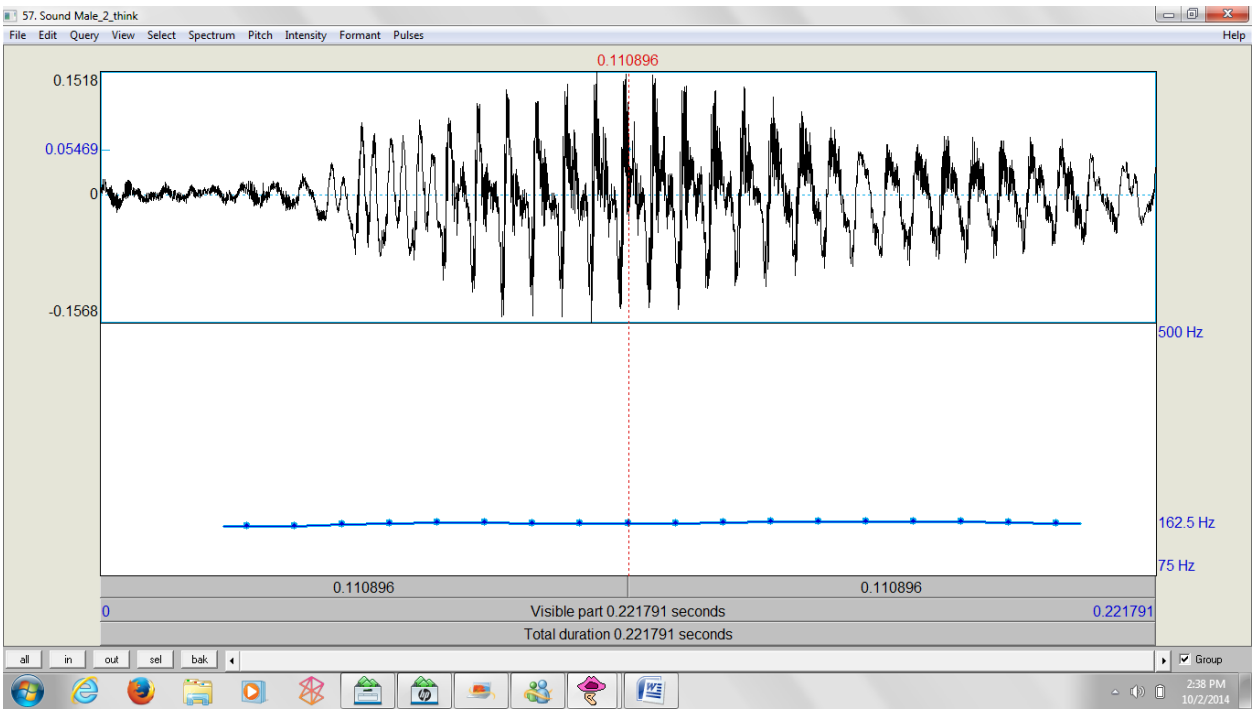


Figure 13(iii) Sound wave and *Fo* and time duration of Male 2 utterance of *I ˌthink so*

[aɪ ˬθɪŋk səʊ]

Male 2 produced the nuclear tone at the equal periodic cycles of 0.110896 secs and 0.110896 secs. The total duration is 0.221791 secs and the highest pitch value 162.5 Hz. The pitch contour does not indicate any clear downward curve reminiscent of fall-rise, rather, we observe what appears like a level stretch of pitch. However, there is evidence of pitch movement from low to high.

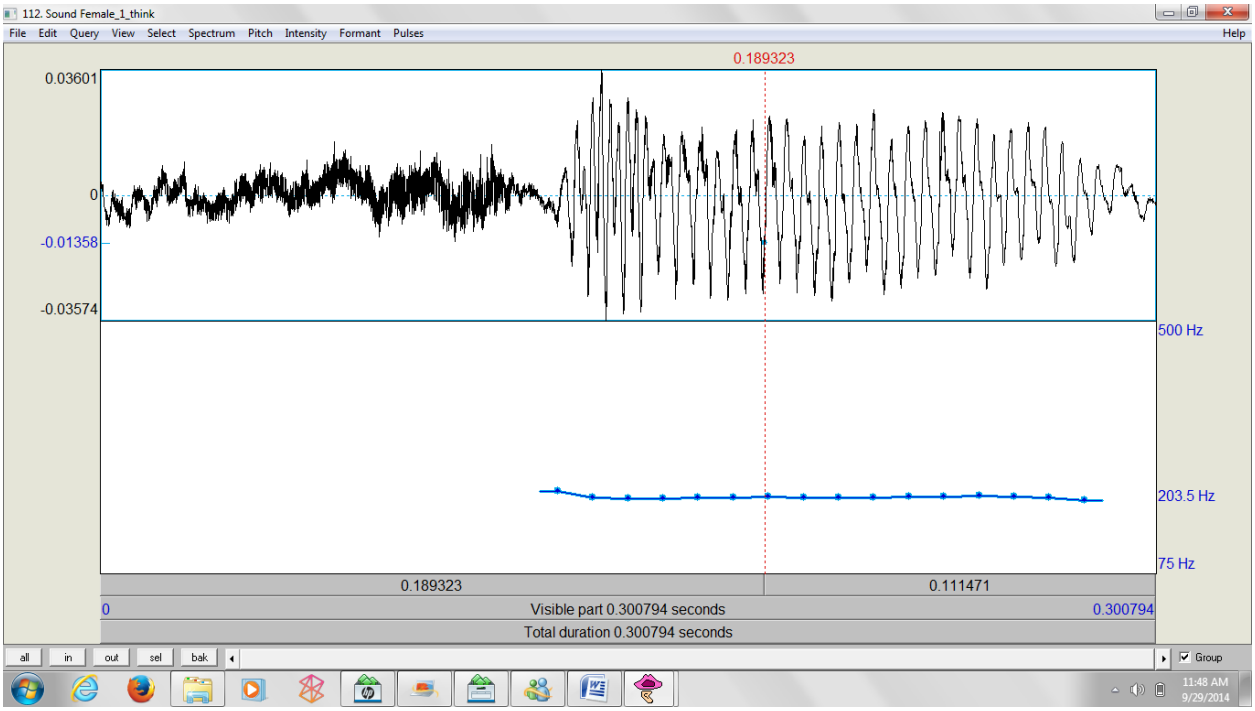


Figure 13**(iv) Sound wave and *Fo* and time duration of Female 1 utterance of *I ˬthink so***
[aɪ ˬθɪŋk səʊ]

Female 1 produced the nuclear tone of the tentative statement at the periodic cycles of 0.189323 secs and 0.111471 secs. The total duration is 0.300794 secs and the highest pitch value 203.5 Hz. The pitch contour reveals an initial step up to high and a slight downward movement, followed by upward movement to a mid.

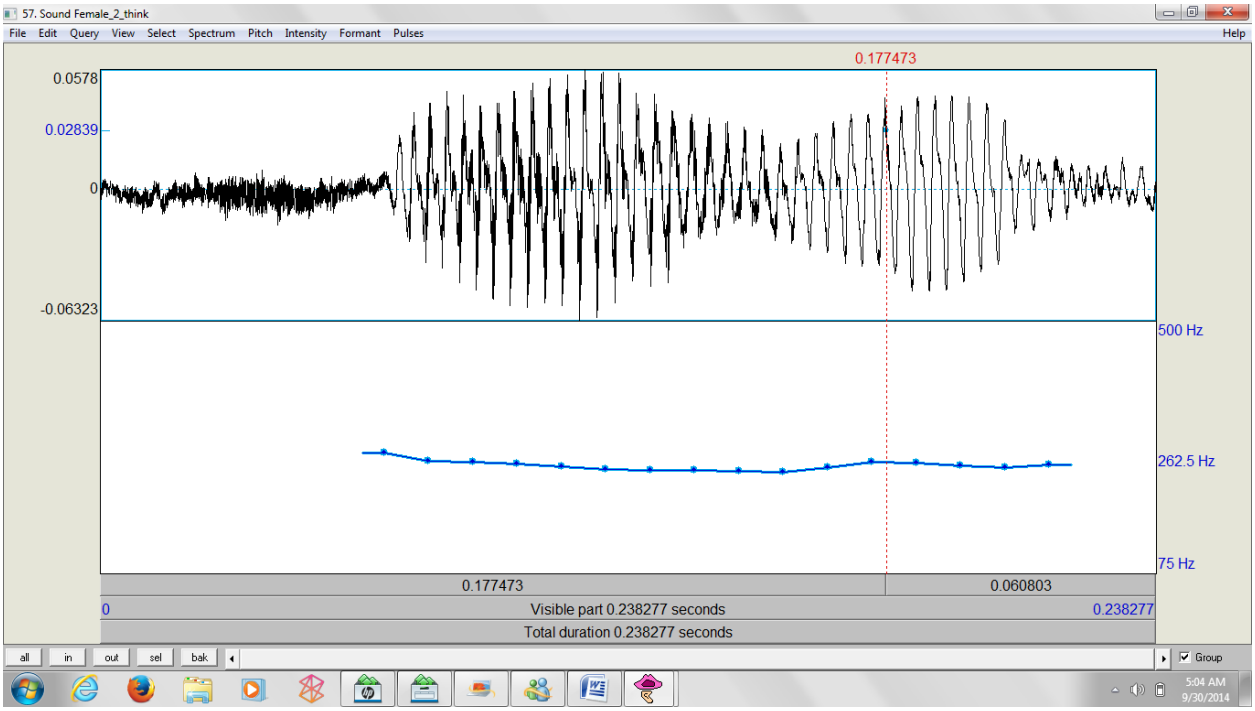


Figure 13(v) Sound wave and *Fo* and time duration of Female 2 utterance of *I ˌthink so* [aɪ ˌθɪŋk səʊ]

Female 2 rendition of the nuclear tone of the tentative statement is presented in periodic cycles of 0.177473 secs and 0.060803 secs, giving the total duration as 0.23877 secs. The highest pitch value is 262.5 Hz. The pitch contour indicates a movement from high to a long-drawn downward movement to another high to mid.

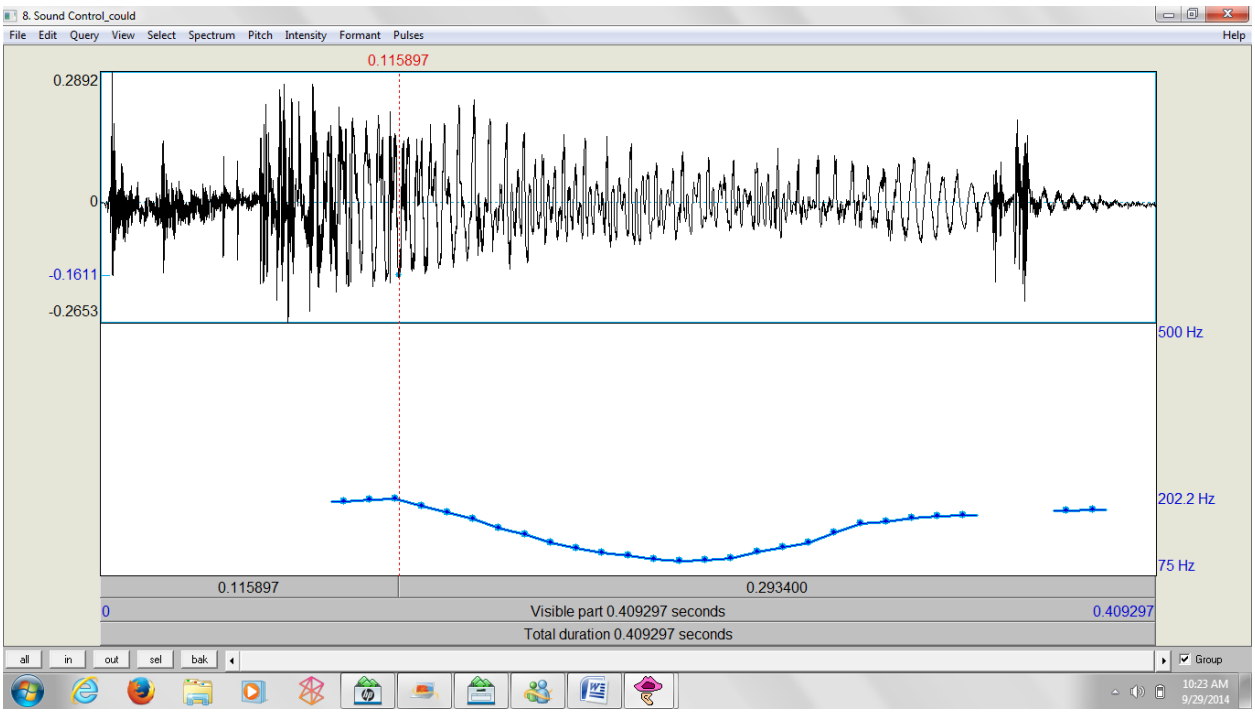


Figure 14(i) Sound wave and *Fo* and time duration of Control utterance of *They ˌcould*

[ðeɪ ˬkʊd]

The Control utterance of the nuclear tone of the tentative statement 'They ˬcould' is rendered at the periodic wave form cycles of 0.11587 secs and 0.293400 secs, giving the total duration of 0.409297 secs. The highest pitch value is 202.2 Hz. The pitch contour of the nuclear tone starts from a high pitch and moves downwards and up to a mid. There is, however, an evidence of silence before the final stop, /d/.

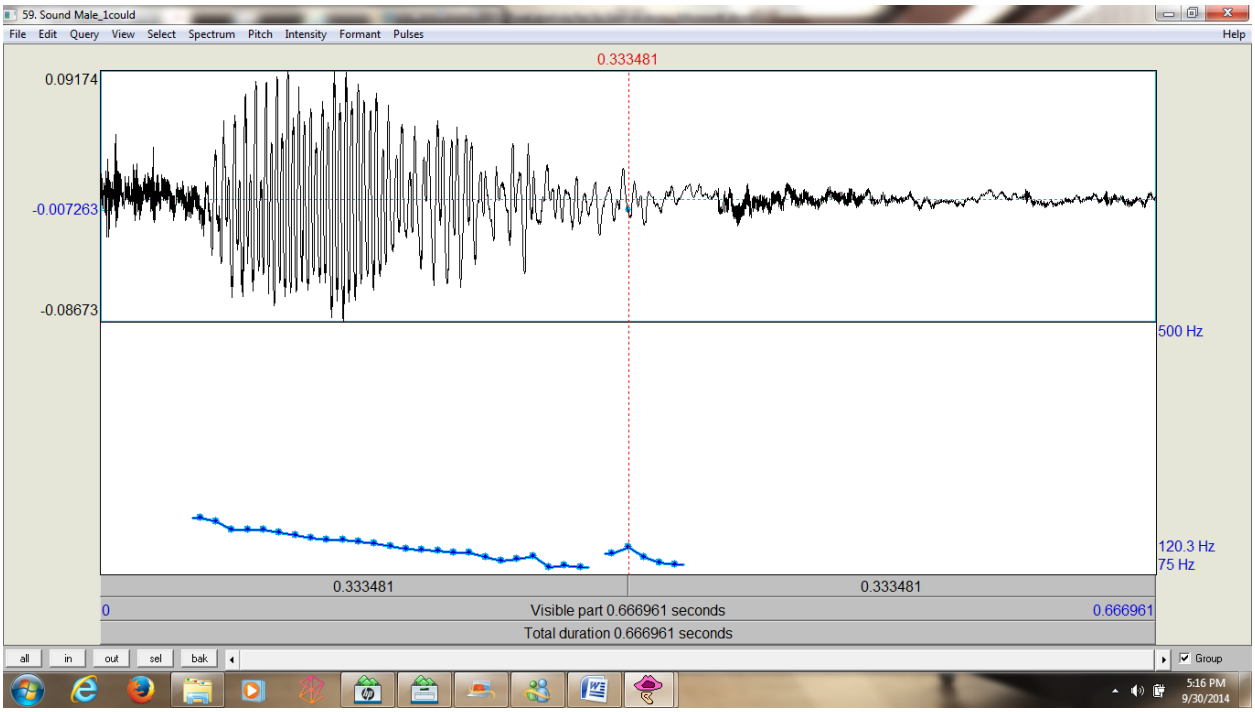


Figure 14(ii) Sound wave and *Fo* and time duration of Male 1 utterance of *They ˬcould* [ðeɪ ˬkʊd]

The Male 1 rendition of the nuclear tone starts from high and moves down and up and down again to a slight pause before another upward movement before the final fall. The periodic wave forms for the utterance are 0.333481 secs and 0.333481 secs, giving the total duration as 0.666961 secs. The highest pitch value measures 120.3 Hz.

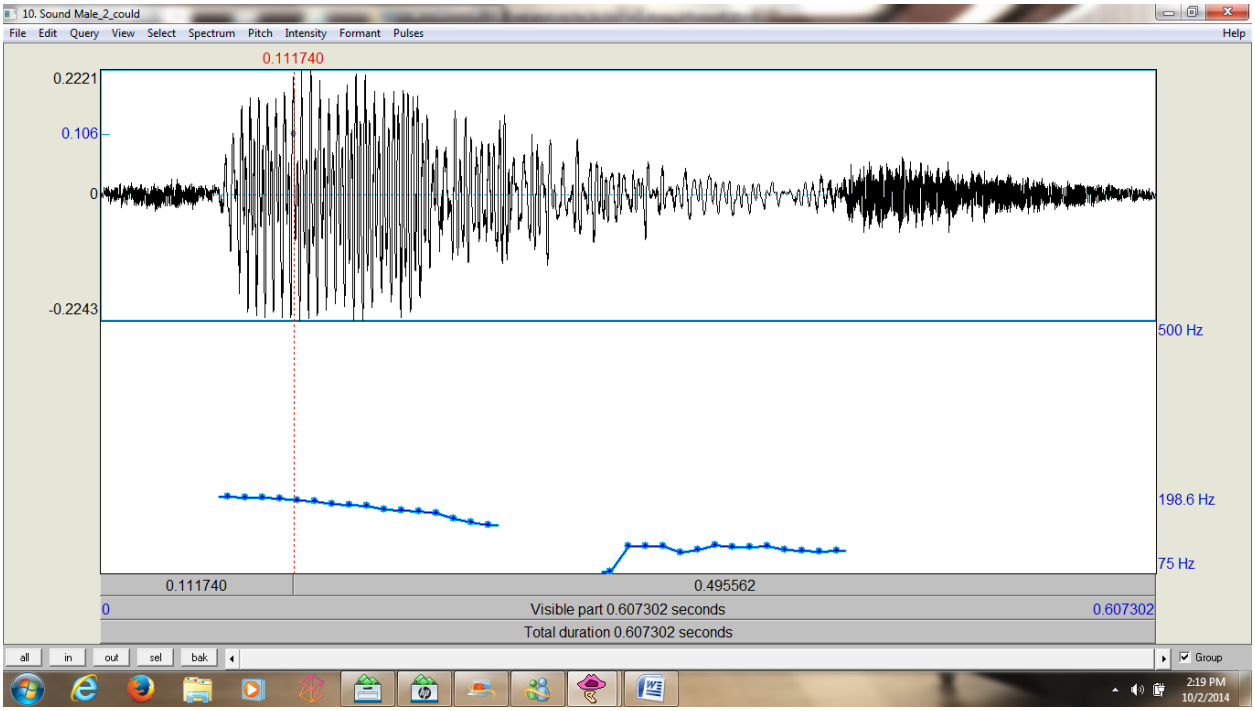


Figure 14(iii) Sound wave and *Fo* and time duration of Male 2 utterance of *They ˌ<u>could</u>* [ðeɪ ˌ<u>kʊd</u>]

Male 2 rendition of the nuclear tone is presented in periodic cycles of 0.111740 secs and 0.495552 secs, giving the total time duration as 0.607302 secs. The pitch value is 198.6 Hz. Here, the pitch contour starts from high to down, then a pause before a sharp upward movement to a shallow down and up movement to a mid.

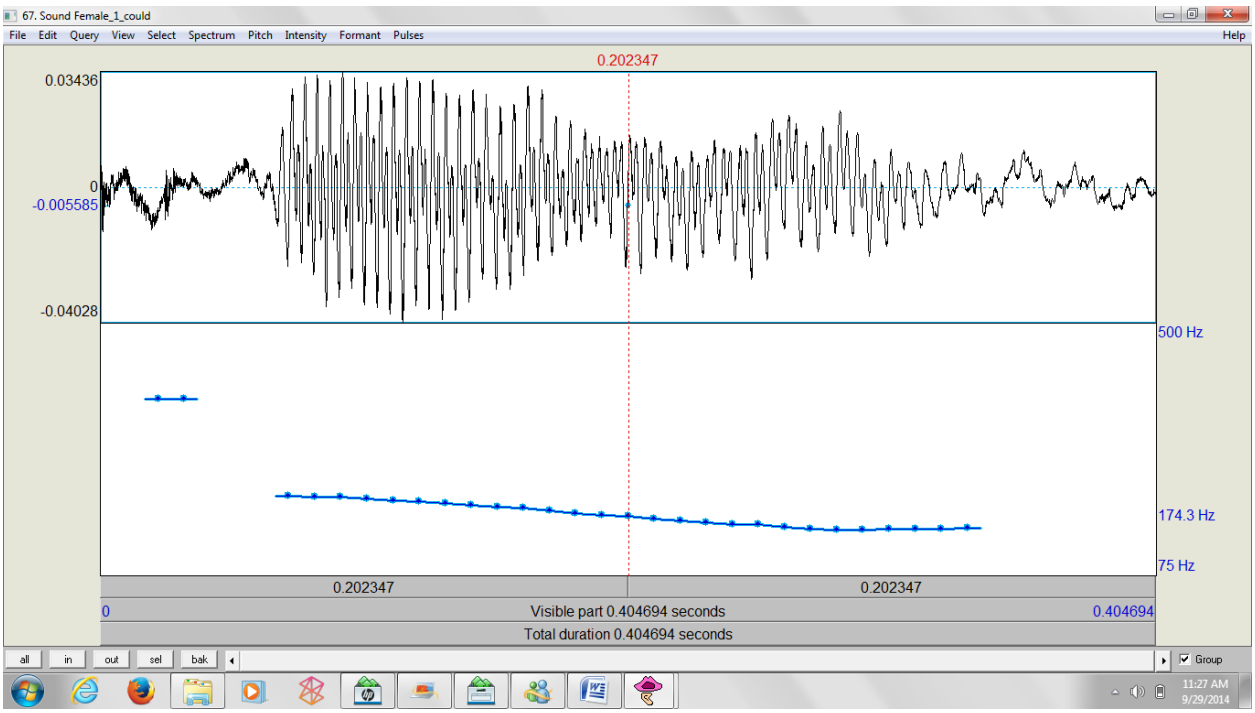


Figure 14(iv) Sound wave and *Fo* and time duration of Female 1 utterance of *They* ˌ*could* [ðeɪ ˌkʊd]

Female 1 produced the nuclear tone at the periodic cycles of 0.202347 secs and 0.202347 secs thus, giving the total duration of 0.404694 secs. The pitch value is 174.3 Hz. The pitch contour starts from high to a downward and upward movement to a mid.

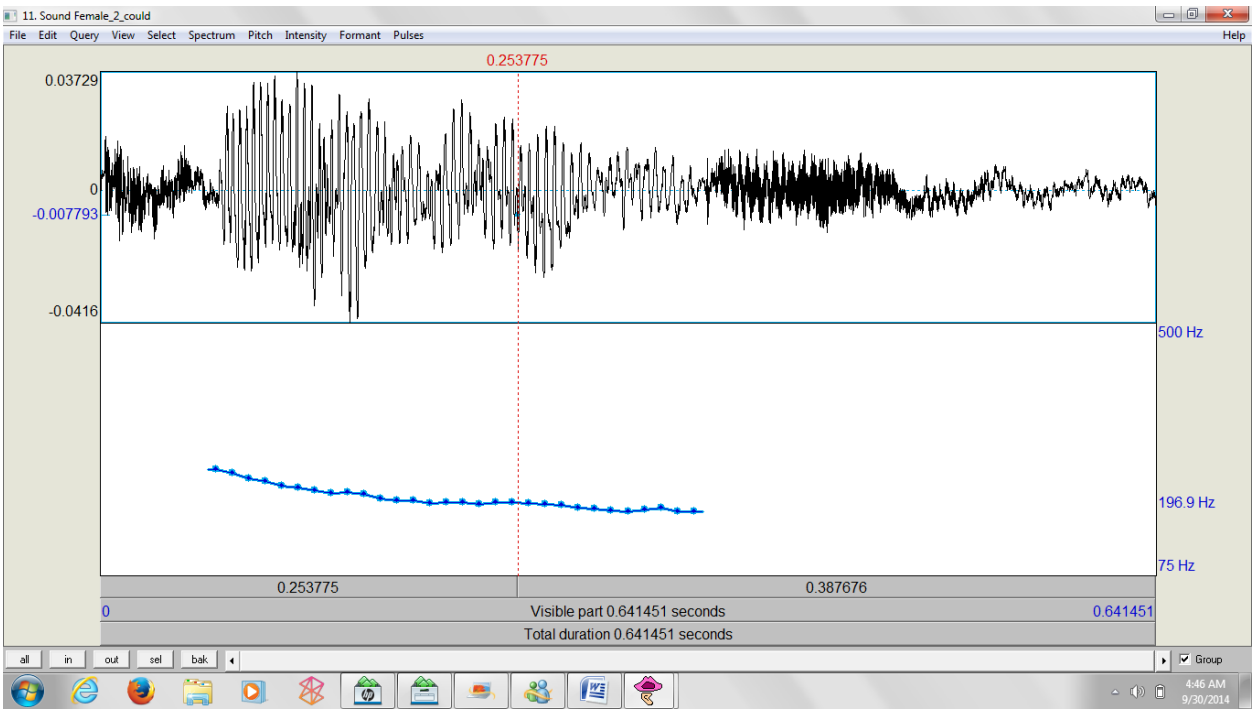


Figure 14(v) Sound wave and *Fo* and time duration of Female 2 utterance of *They* ˌ*could* [ðeɪ ˌkʊd]

Female 2 presents the nuclear tone of the tentative statement at the periodic wave form cycles of 0.387676 secs, giving the total time duration as 0.641451 secs. The pitch value is 196.9 Hz. The pitch contour of the utterance indicates a movement from high to a slight downward and upward movement to mid.

Table 3 Spreadsheet of the acoustic correlates of Tentative Statements

	CONTROL		CONSULTANTS							
			MALE 1		MALE 2		FEMALE 1		FEMALE 2	
	Pitch value (Hz)	Duration (ms)	Pitch value (Hz)	Duration (ms)	Pitch value (Hz)	Duration	Pitch value (Hz)	Duration (ms)	Pitch value (Hz)	Duration (ms)
(i) I ˌthink so.	376	0.239116	132.3	0.251497	162.5	0.221791	203.5	0.300794	262.5	0.238277

(ii) They ˬcould	276.4	0.559615	106.6	0.709728	152.5	0.640091	310.5	3.349660	243.4	0.805147

4.1.3(b) Summary of acoustic analysis of pitch pattern in tentative statements

Fall-rise is noted for tentative statements. In the utterances of the tentative statements: *I ˬthink so* and *They ˬcould*, we observe some similarities among the utterances. The pitch contour of the Control utterance of the nucleus, *ˬthink* have very close relationship with those of the Consultants. Male 1, Female 1 and Female 2 utterances all present slight downward and upward movement of tone. Only Male 2 has the tone with a clear shape of the pitch contour. The pitch levels of Male 1 and Male 2's utterances are lower than those of the English and Female 1 and Female 2. Hence, the pitch value of the Control is 376 Hz, and that for Female 2, is 262.5 Hz. The pitch values for Male 1, Male 2 and Female1 are 132.3 Hz, 162.5 Hz and 203.5 Hz respectively. For the total time duration of the utterances, all but that for Female 1 fall within the same range of between 0.22179 secs and 0.251497 secs, even though the time duration of 0.300794 secs for Female 2 varies only very slightly. This general resemblance speaks positive about the correlation between tone and intonation in tentative statements.

For the nuclear tone in *They ˬcould*, all the pitch contour for the Consultants start from somewhere up and end on a downward note. Also, only the pitch contour of the utterances of Male 1 and Male 2 clearly indicate downward and upward movement reminiscent of the tone of tentative statements. The two utterances equally show evidence of delayed release at the end of the nucleus as is the case with the Control utterance. But Female 1 and Female 2's utterances apparently appear with clumsy curves but still move from somewhere up to somewhere down. The pitch values of 310. 5 Hz for Female 1 and 243.4 Hz for Female 2 are more closely related to the 276.4 Hz for the English, while the pitch value for Male 1 and Male 2 measure 106.6 Hz and 152.5 Hz respectively. Similarly, the total time duration for the Control and those of Male 1, Male 2 and Female 2 relate more, while that for Female 1 is clearly higher than the rest.

4.1.4(a) Sound waves of Yes-no questions

According to Wells (2006), a yes-no question has a default tone, which is a rise. On the other hand, he says that it is equally possible for a yes-no question to be said with a fall. This tone, he calls 'insistent fall'.

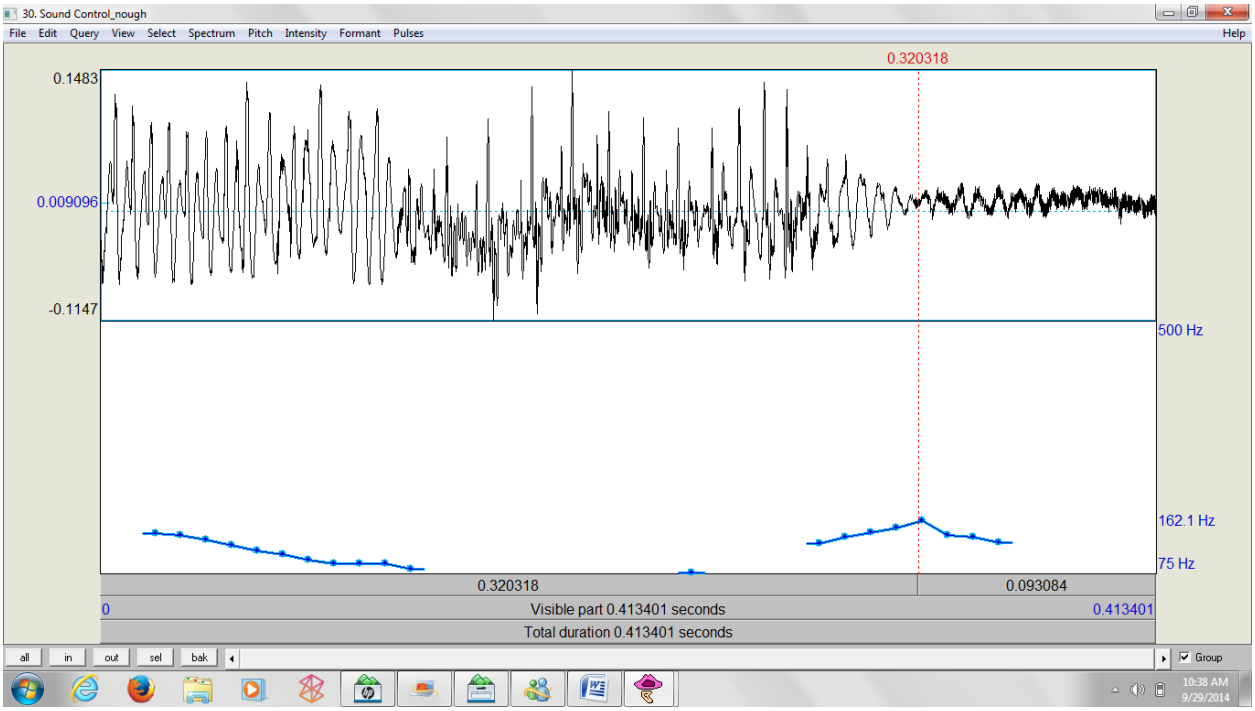


Figure 15(i) Sound wave and *Fo* and time duration of Control utterance of *'Haven't they done e·nough?* ['hæv.ᵊnt ðeɪ dʌn ɪ'nʌf]

The Control production of the nuclear tone of the yes-no question is rendered in periodic wave form cycles of 0.320318 secs and 0.093084 secs, thus giving the total duration of 0.413401 secs. The pitch value is 162.1 Hz. The pitch contour of the utterance actually indicates a rise that ends on a slight drop to mid.

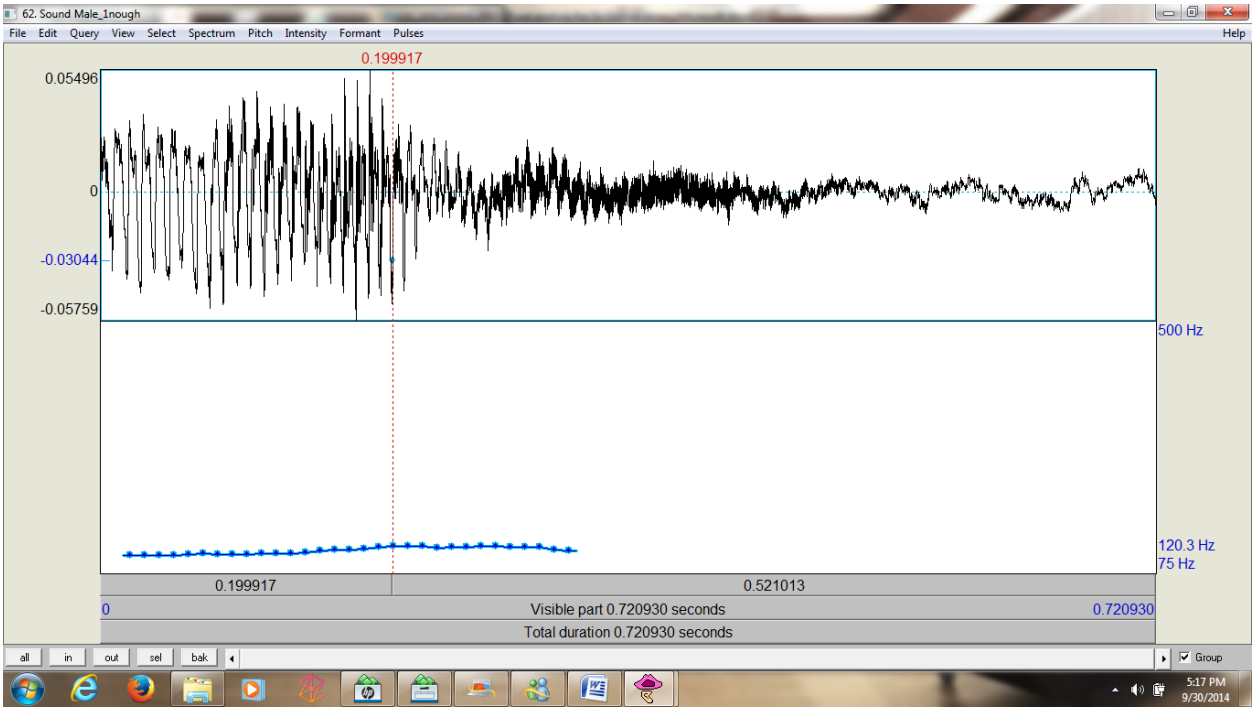
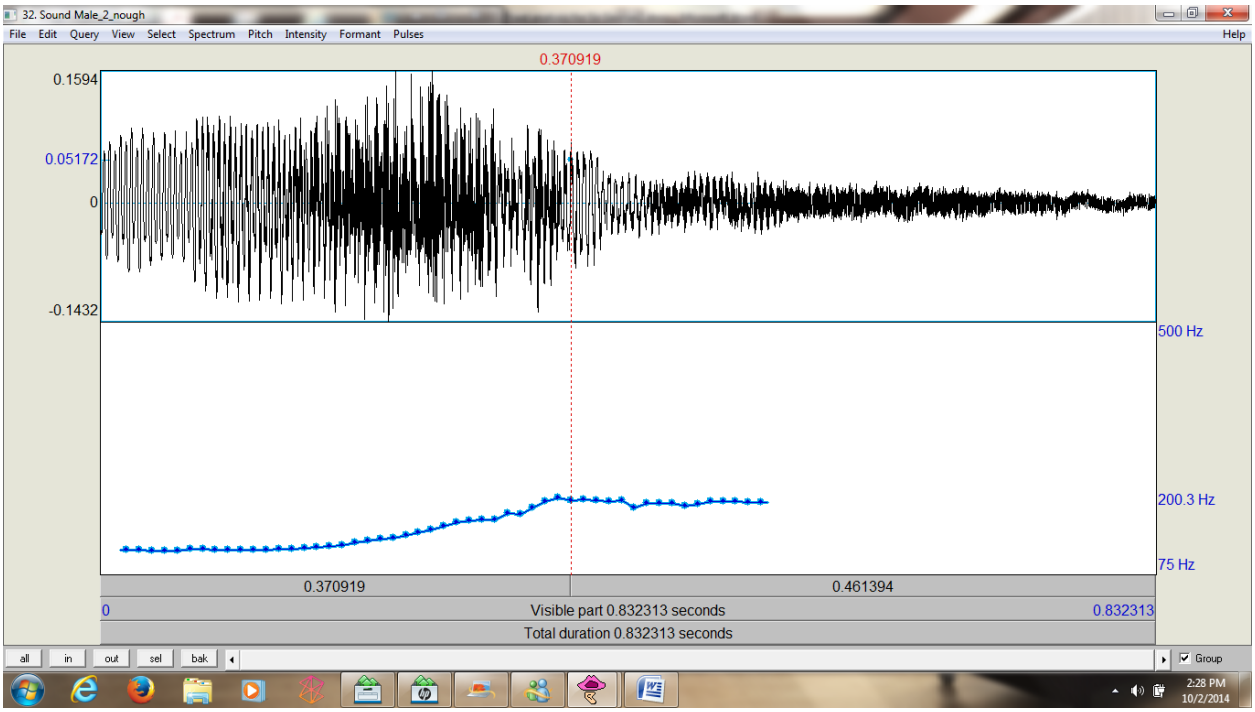


Figure 15(ii) **Sound wave and *Fo* and time duration of Male 1 utterance of *'Haven't they done e-nough?* ['hæv.ᵊnt ðeɪ dʌn ɪ'nʌf]**

Male 1 native Igbo speaker produced the nuclear tone of the yes-no question at the periodic wave forms of 0.199917 secs and 0.521013 secs, giving a total time duration of 0.720930 secs. The highest pitch value is 120.3 Hz. The pitch contour indicates a rise although it is on a low pitch.



(iii) **Sound wave and *Fo* and time duration of Male 2 utterance of *'Haven't they done e-nough?* ['hæv.ᵊnt ðeɪ dʌn ɪ'nʌf]**

The Male 2 utterance of the nuclear tone of yes-no question was produced at the periodic wave forms of 0.370919 secs and 0.461394 secs. The total time duration is 0.832313 secs; and the highest pitch value 200.3 Hz. The pitch contour clearly indicates the default tone of yes-no question, which is rise.

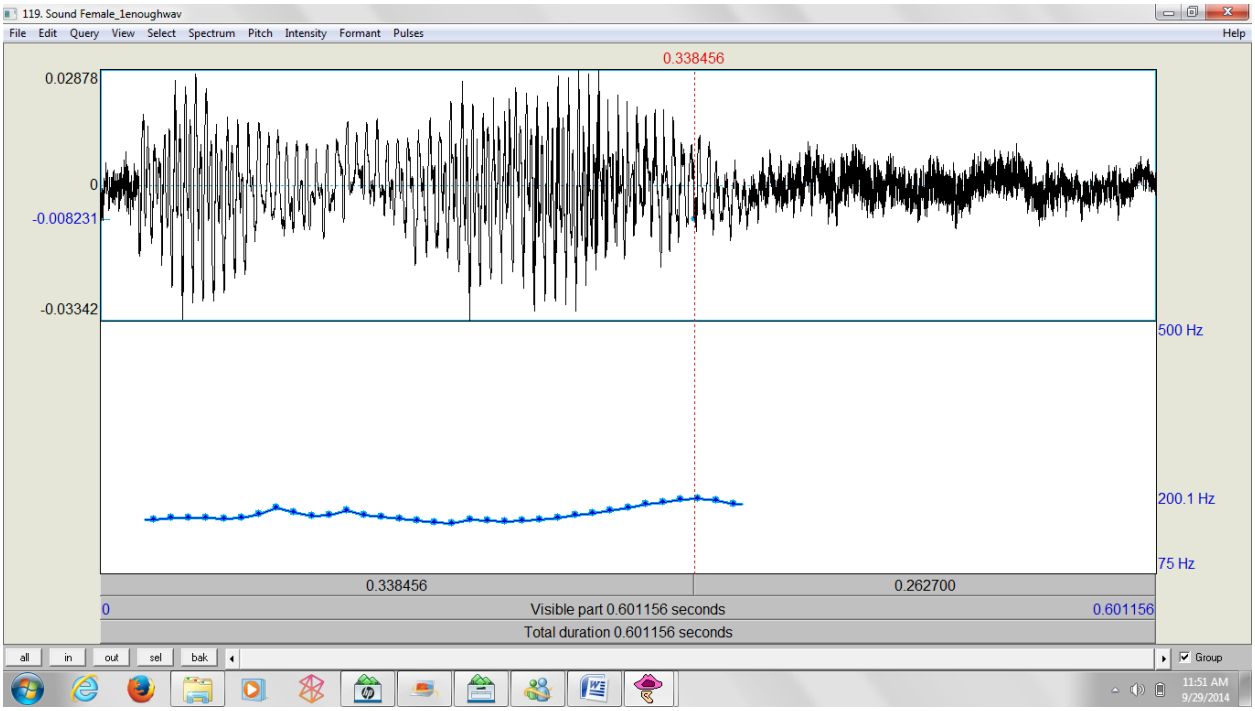


Figure 15(iv) Sound wave and *Fo* and time duration of Female 1 utterance of *'Haven't they done e<u>nough</u>?* ['hæv.ᵊnt ðeɪ dʌn ɪ'nʌf]

Female 1 rendered the nuclear tone at the periodic wave forms of 0.338456 secs and 0.262700 secs. The total duration is 0.601156 secs, while the pitch value is 200.1 Hz. Again, the pitch contour shows a rise typical of a yes-no question.

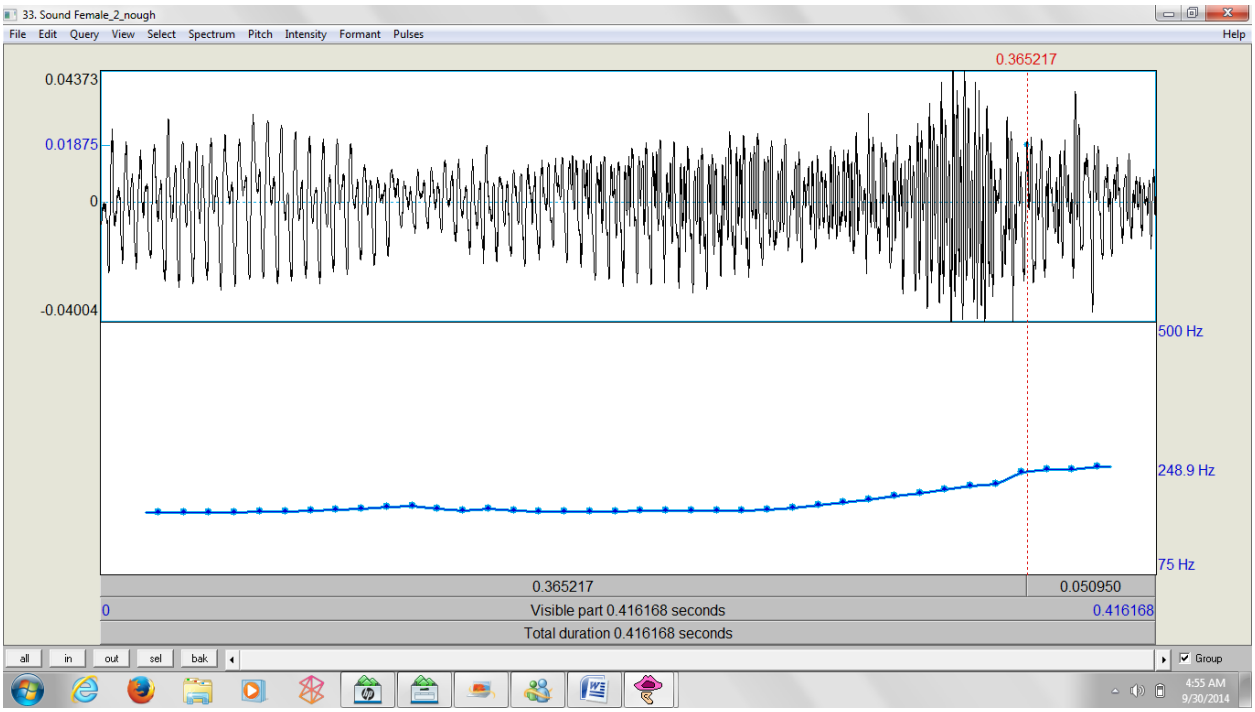


Figure 15(v) **Sound wave and *Fo* and time duration of Female 2 utterance of *'Haven't they done e˗nough?* [ˈhæv.ᵊnt ðeɪ dʌn ɪˈnʌf]**

The Female 2 native Igbo speaker produced the nuclear tone at the periodic wave form cycles of 0.365217 secs and 0.050950 secs. The total time duration is 0.416168 secs; and the highest pitch value is 248.9 Hz. The pitch contour of this utterance presents a long-drawn stretch, which indicates an upward movement to a rise that is reminiscent of yes-no question.

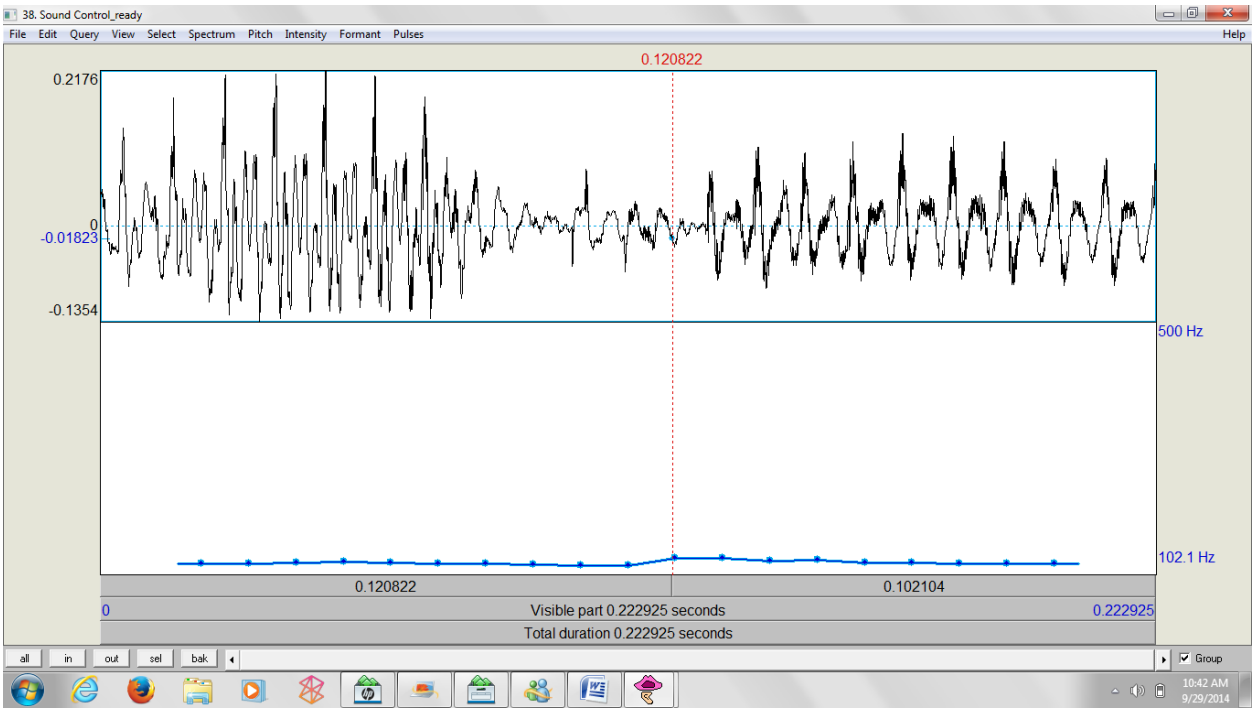


Figure 16(i) **Sound wave and *Fo* and time duration of Control utterance of *'Isn't she ˊready?* [ˈɪz.ᵊnt ʃi: ˊred.i]**

The Control rendition of this nuclear tone of yes-no question is made at the periodic wave form of 0.120822 secs and0.102104 secs. The total time duration is 0.222926 secs, and the pitch value 102.1 Hz. The pitch contour indicates a rise but on a very low pitch.

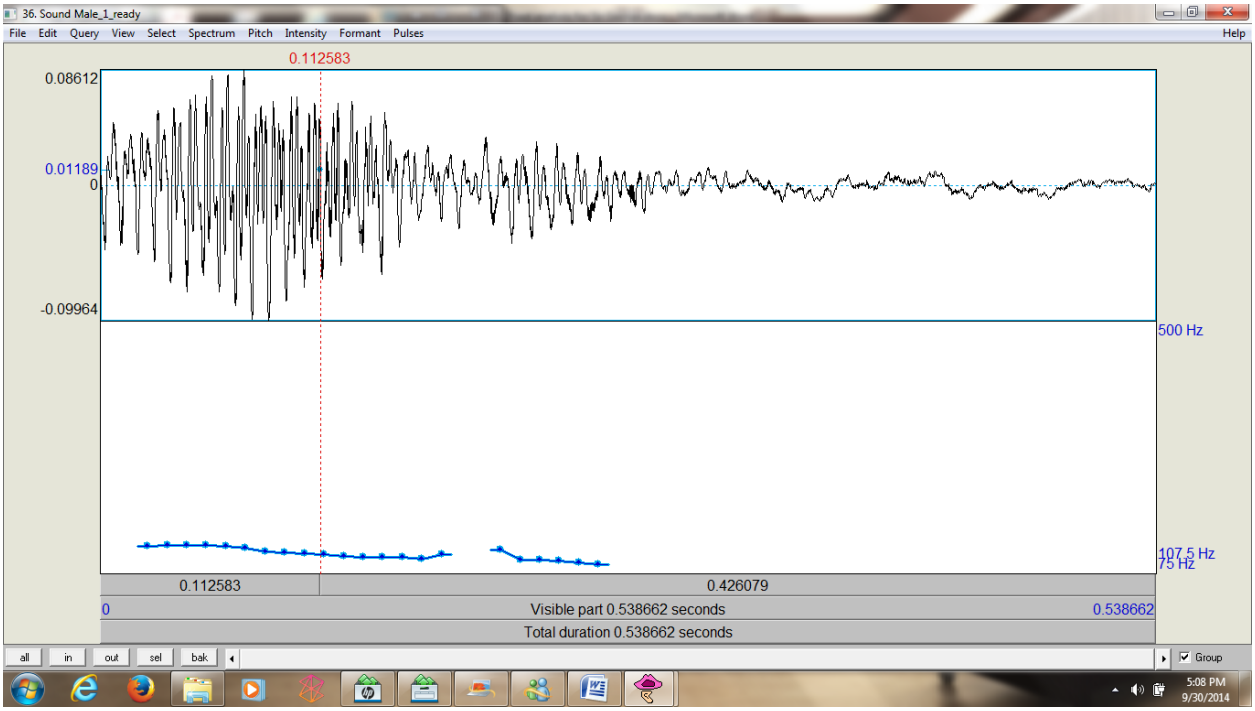


Figure 16(ii) Sound wave and *Fo* and time duration of Male 1 utterance of *'Isn't she ˊready?* ['ɪz.ᵊnt ʃi: ˊred.i]

The Male 1 native Igbo speaker presented the nuclear tone at the periodic wave forms of 0.112583 secs and 0.426079 secs, yielding a total time duration of 0.538662 secs. The pitch value is 107.5 Hz. The pitch contour indicates a pause before the alveolar stop /d/. Secondly, the pitch movement goes down from onset with a slight rise before the pause and then followed by downward movement that rises to mid.

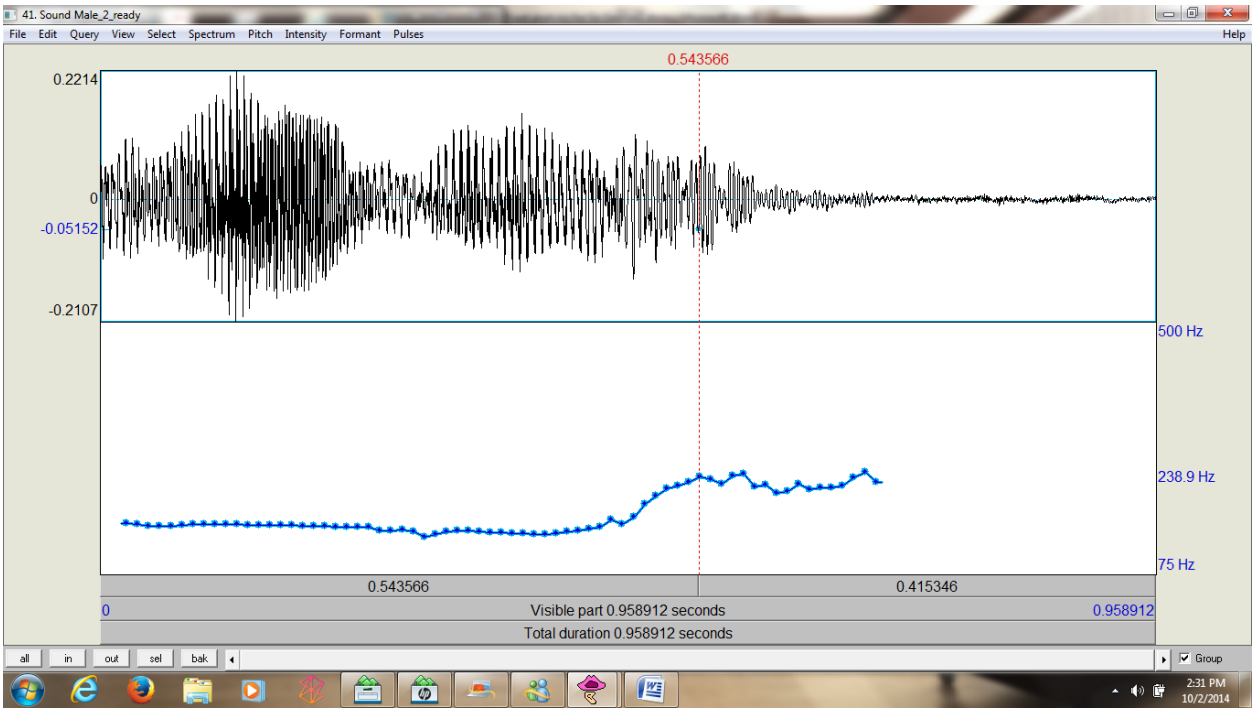


Figure 16(iii) Sound wave and *Fo* and time duration of Male 2 utterance of *'Isn't she ˌready?* ['ɪz.ᵊnt ʃi: ˌred.i]

Male 2's utterance clearly presents an upward movement of the pitch to a rise with a seemengly wobbling end. The wave form cycles used to produce the tone are 0.543566 secs and 0.415346 secs, giving the total time duration as 0.958912 secs. The highest pitch value is 238.9 Hz.

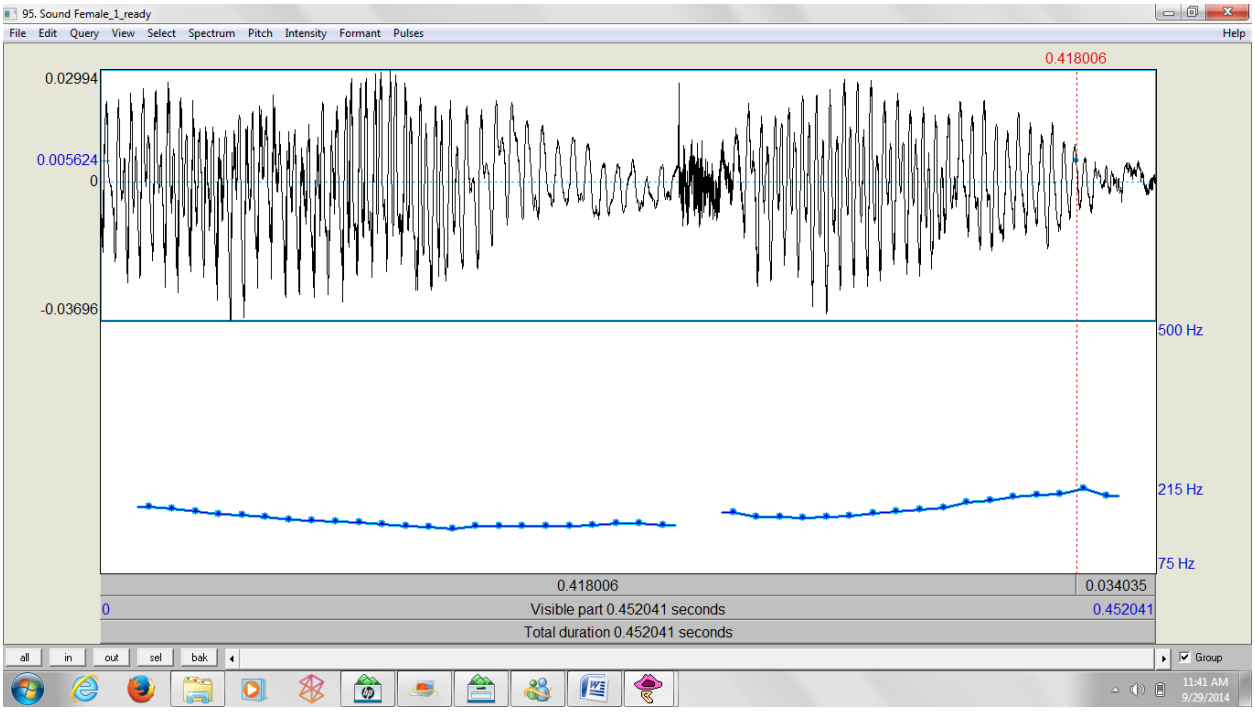


Figure 16(iv) Sound wave and *Fo* and time duration of Female 1 utterance of *'Isn't she ˌready?* ['ɪz.ᵊnt ʃi: ˌred.i]

Female 1 produced the nuclear tone at the periodic wave form cycles of 0.418006 secs and 0.034035 sec. The total time duration is 0.452041 secs, and the pitch value 215 Hz. The pitch contour indicates a pause possibly before the alveolar stop, followed by an upward movement to a rise.

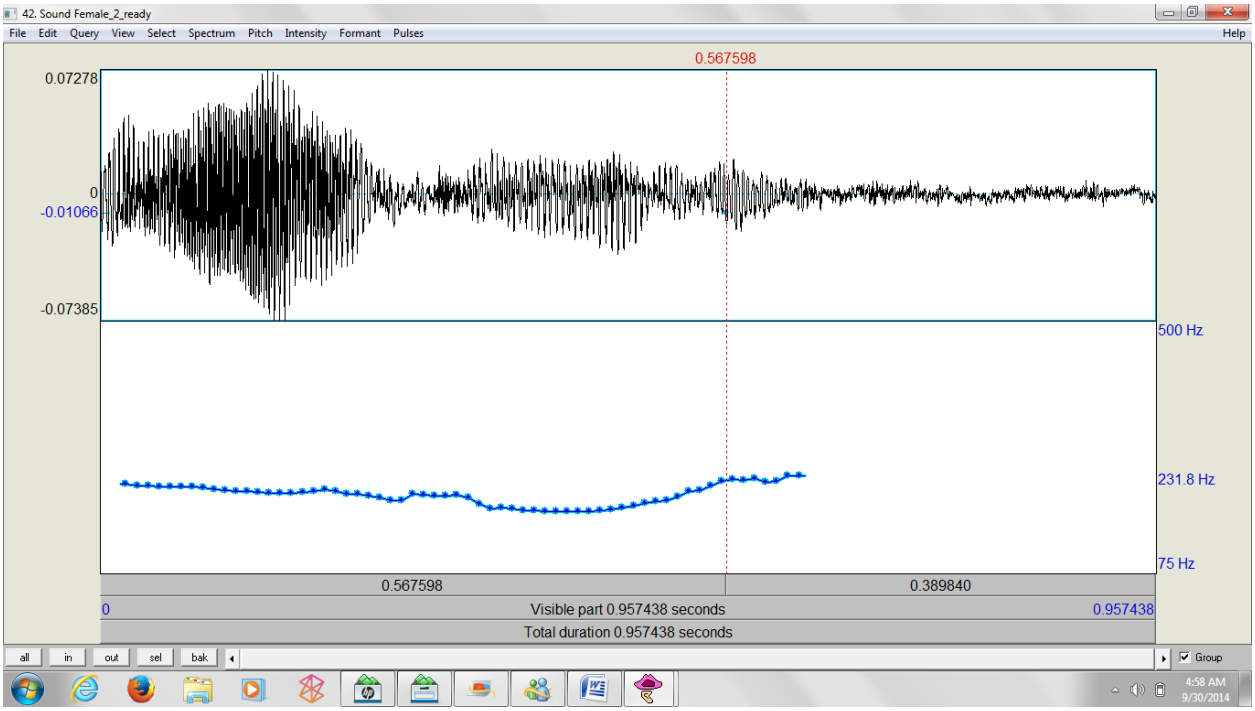


Figure 16(v) Sound wave and *Fo* and time duration of Female 2 utterance of *'Isn't she ˌready?* ['ɪz.ᵊnt ʃi: ˈred.i]

Female 2 presented the nuclear tone of the yes-no question at the periodic wave form cycles of 0.567598 secs and 0.389840 secs, giving the total time duration of 0.957438 secs. The pitch value is 231.8 Hz. The pitch contour of the utterance shows an upward movement to a rise on a relatively high pitch.

Table 4 Spreadsheet of the acoustic correlates of Yes-no questions

	CONTROL		CONSULTANTS							
			MALE 1		MALE 2		FEMALE 1		FEMALE 2	
	Pitch value (Hz)	Duration (ms)	Pitch value (Hz)	Duration (ms)	Pitch value (Hz)	Duration	Pitch value (Hz)	Duration (ms)	Pitch value (Hz)	Duration (ms)
(i) 'Haven't they done eˌnough?	162.1	0.413401	120.3	0.720930	200.3	0.832313	200.1	0.601156	248.9	0.416168
(ii) 'Isn't she ˌready?	102.1	0.222925	107.5	0.538662	238.9	0.958912	215	0.452041	231.8	0.957438

4.1.4(b) Summary of acoustic analysis of Yes-no questions

The default tone for Yes-no question has been identified as rise. In these Yes-no questions under study, there are similarities in the utterances of the Control and the Consultants with the default tone. However, in the nuclear tone of *'Haven't they done e·nough?* there is evidence of pause (delayed) release in the Control utterance before the slight drop at the end. The pause is, however, not observed in the Consultants' utterances. Also, both the English and Male 1 utterances are on very low pitch level when compared with those of Male 2, Female 1 and Female 2. Thus, the pitch value for the Control utterance measures 162.1 Hz and 120.3 Hz for Male 1. Male 2 and Female 1 pitch values measure 200.3 Hz and 200.1 Hz respectively; and 248.9 Hz for Female 2 . The total time duration for the Control and Female 2 are 0.413401 secs and 0.416168 secs respectively, while for Male 1, Male 2 and Female 1, we have 0.720930 secs, 0.832313 secs and 0.601156 secs respectively.

Again, the nuclear tone for *'Isn't she ·ready?* is on low pitch for the English and Male 1 utterances more than is the case in the other utterances by the other Consultants. All the other utterances, however, show great similarities in tone with the default tone. But, Male 1 and Female 1 show evidence of silence before the release of the alveolar stop. The pitch values for the Control and Male 1 measure 102.1 Hz and 107.5 Hz. But, Male 2 and Female 2 pitch values are 238.9 Hz and 231.8 Hz respectively while Female 1 measures 215 Hz. Thus, the Control and the Male 1 utterances are more closely related as Male 2 and Female 2 both relate more closely in F*o* measurement. The total time duration for the utterances varies slightly, which could be an evidence of attitudinal differences of individuals.

4.2 Similarities and Differences between tone and intonation

As the major concern of this study, identifying the point of similarities between tone and intonation will be carried out using the various observations made from the analysis of the data used;

in other words, the findings of the study. Hence, the stress patterns of different forms of words, the pitch contours of different sentence types, the values of fundamental frequencies and the total durations for the production of these individual words and structures (sentences) are all considered basically to identify their meeting points in tone and intonation languages.

Under word stress, all the utterances correlate in their fall-rise stress pattern of the two-syllable word *record (v)*. Similarly, the rise-fall stress pattern in *complete (adj)* and *below (adj)* also correlate but with slight differences in the Male 1 and Male 2 utterances. The pitch values for both Female 1 and Female 2 are higher than those for the Control, Male 1 and Male 2. Nonetheless, the pitch contours indicate that the primary stress correlate with the high tone as it indicates the highest pitch value and attracts the greatest prominence in each occurrence. It, therefore, portends the syllable of focus. Also, in the three syllable words, the utterances relate in many ways, though with slight variations. In these utterances, the stressed syllables produce the highest pitch levels of the utterances with the other forms of stress representing the other levels of pitch. However, Male 2 and Female 1 end their utterances of *interrupt* in an upstep tone. Generally, the results confirm that stressed syllables attract longer durations than the unstressed ones. Moreover, front (open) and tense vowels also attract longer time duration.

In the compound words, both the English speaker and most Consultants share a lot in common. Male 1 and Female 2 are mostly related to the Control in the pitch contour for *second-class*, though with slight differences. Male 1 and Female 1 are mostly different from the Control. The Male 1 duration depicts the Igbo adaptation of loanwords. Male 2 and Female 2, and the Control utterances share similarity in syllable division and Male 1 and Male 2 in pitch values. However, the pitch contours of Male 1 and Female 1 differ clearly with the Control and other Consultants; and with Male 1 registering the lowest pitch level while Female 2 records the highest. The pitch contours of Male 2, Female 1 and Female 2 for *tea-cup* appear different from those of the Control and Male 1. Also, only Female 1 presents a pitch contour that ends on a sharp fall. Again, Female 1 and Female 2

register very high pitch values. Then, except Male 1 and Female 2 in the utterances for *shoelace*, the pitch contours for the Control, Male 1 and Female 2 which end on fall-rise, all the others resemble. Furthermore, the wavelengths of the Consultants' speeches appear shorter than that of the Control.

Therefore, we conclude that the tone language (native Igbo) speakers recognise stress and apply stress patterns correctly in English words the same way tone is assigned in Igbo words where the high tone is assigned to the heavy syllable. However, the pitch heights and duration are largely influenced by the vowels.

Under the sentences, the utterances of the statements with defintive fall all resemble that of the English, except that of Female 2, which tends to have a short rise in the nuclear tone of *I 'think it's ˌgreat.* The same is the case with the utterances of the nuclear tone of *We're 'all ˌhere* where all the Consutants' utterances present similar pitch contours. However, the pich height of the Consultants' utterances are not as high as that of the English, with those of Male 1 and Male 2 being the lowest. This difference in pitch height may be attributed to the diphthong /ɪa/ in which the vowels /ɪ/ and /a/ would be regarded as [-ATR] vowels in Igbo. Secondly, Igbo does not have diphthongs in its phoneme inventory. Consequently, the differences can be attributed to native language interference. Also, the durations of the utterances are closely similar. The relative shortness of the English durations can be attributed to the relative shortness of the wavelength in stress-timed languages.

In the tentative statements, the nuclear tone is usually fall-rise. Here, all the Consultants present similar tone in *I ˌthink so* except Male 2 rendition, which appears to be a level tone. Again, Male 1 and Male 2 produced theirs on a pitch much lower than the English 376 Hz. Female 1 and Female 2 are equally low, but not as low as those of Male 1 and Male 2. The reason for the low pitch of Male 1 and Male 2 can be attributed to low pitch range for male adult speakers.

Similarly, the default tone for the tentative statement is also clearly indicated in the utterances of the Control, Male 1 and Male 2 in *They ˌcould*, and slightly by Female 1 and Female 2. Evidence of delayed release is noticed in the utterances of the English, Male 1 and Male 2, but totally absent in

those of Female 1 and Female 2. Male 1 records the lowest pitch level, followed by Female 1; while Male 2 and Female 2 register pitch levels very close to that of the English. The total time durations for the two tentative statements are more closely related to the Control in *I think so*, while only that for Female 1 appears widely at variance with the Control in *They could*. Generally speaking, therefore, there is very great affinity between the English and native Igbo speakers' utterances of the fall-rise nuclear tone of the tentative statements.

The default tone for Yes-no question which is rise seems to synchronise among the Control and the Consultants. For *'Haven't they done enough?* The native English speaker produced it on a very low pitch and with a pause before the fricative /f/ which none of the Consultants observed. Again, this may be attributed to the native language (Igbo) interference where no word ends with a consonant. Male 2 and Female 1 utterances are produced on a relatively high pitch, while Female 2 is the highest.

In the nuclear tone of *'Isn't she ready*? the Control and Male 1 appear on lower pitch level, while the rest of the Consultants render theirs on varying pitch levels. Male 1 and Female 1 indicate silence, perhaps before the pronunciation of the voiced alveolar stop /d/. Again, this can be a case of individual differences in pronunciation.

With the above findings of the study, therefore, we can make categorical statements and generalisations on tone intonation interface.

CHAPTER FIVE

SUMMARY OF FINDINGS AND CONCLUSION

This work has delved into a relatively novel area in the study and analysis of speech, which is acoustic study. This chapter therefore summarises the procedure for the research and juxtaposes the findings with the research questions to ascertain how far they answer the questions and justify the purpose of the study.

5.1 Summary of Findings

The study was designed to carry out an acoustic (experimental) study of the speech in English of native English speakers vis-a-vis those of the native Igbo speakers. Recognising the fact that English is an intonation language and Igbo a tone language, the study specifically set out to investigate the interface of tone and intonation through the study of the suprasegmental features, which they both share. It also set out to determine their points of correlation and to authenticate the relevant information about their relatedness. The study equally set out to ascertain whether the native Igbo speakers encounter problems speaking English (intonation) language, as well as identify the nature of such problems. Furthermore, the study made relevant statements on the attitude of the speakers towards generating (intended) meanings. At the end of the investigation (study), revealing and interesting findings were made, which go a long way in achieving the objectives as well as provide answers to the research questions.

The data comprised words and sentences. The words (two and three syllable words) were used to investigate stressing; and the nucleus of sentences used to investigate tone. The acoustic correlates of fundamental frequency and duration form the rubrics for our analysis in the work.

As the primary focus of our study is to investigate the point of relatedness of intonation and tone, our findings show that both tone and intonation are aspects of suprasegmental phonology, which necessarily have to coexist. Intonation involves raising and lowering the pitch of the voice, that is change in the pitch of the voice as we speak which could range from light to heavy (or low or

high), and tone is itself the height of pitch and change of pitch (associated with pronunciation of syllables or words and which affects the meaning of the word). It was again confirmed from the results that no two speakers speak on the same pitch level; one may speak on a lower pitch than the other. On the other hand, every individual has control over his or her pitch. Hence, an individual may choose to speak with a pitch higher than normal to achieve intended meaning, or as a result of individual difference in pronunciation. The above observations therefore, answer our first and second research questions on different modes of realisation of intonation and regular suprasegmental features of tone.

Our third research question seeks to establish the tones and stress that correlate. Strictly speaking, our study observes that stress is not used in Igbo as it is used in English. All the same, as stress is the pronunciation of a syllable or word with more respiratory energy or muscular force than other words or syllables in an utterance, stressed syllables have greater overall acoustic intensity than the more weakly stressed or unstressed ones as indicated by pitch values and durations of the utterances investigated and it has greater length and is more prominent and audible. Thus, we identify various types of stress as primary stress, secondary stress and tertiary stress with the primary stress as the most prominent. Stress therefore results from a combination of loudness, pitch and duration and is applied to syllables and not individual segments.

Stress is used lexically in English to distinguish different words such as word class and establish the meaning of the word. Similarly, in tone language, tone distiguishes words with the same segments and as well establishes their individual meanings. Notably, tone can be high or low. High tone results when vocal cords vibrate rapidly making the affected syllable to be more prominent and to sound higher, thus, becoming the focus in that word or structure as is the case with stressed syllable.

The above observations therefore relatively equate stressed syllables with high tone in this study in line with their acoustic properties. They also justify the objective of the research, which aims to determine the features of tone that English and Igbo (intonation and tone) languages share.

On tone, it needs to be restated that English does not have lexical tone, that is, fixed pitch as is the case with tone language like Igbo. In English tone, pitch is used over syntactic units to distinguish meaning. All the same, the tone could be high (rise) or low (fall), but they are increased or reduced at the instance of the speaker to achieve the meaning he wants to convey, thus resulting in different intonation patterns. However, whereas certain pitch pattern (tone) is applied to every syllable of an utterance in tone languages, only one syllable in an intonation phrase bears the default tone, which is used mainly to express attitude. Hence, the tonal specification of an English utterance is said to be closely related to the notion of focus.

In this work, the fall and rise pitches respectively correlate with the low and high tone of the tone language. These tones have been ably reproduced in the nuclear tones of the statements (with definitive fall) and yes-no questions of the native Igbo speakers. The fall-rise default tone of the tentative statement is greatly manifested in the utterances of both the Control and the Consultants. The variations noticed here are quite insignificant and can be attributed to individual differences usually noticed during speech. These tones satisfy both the attitudinal and grammatical functions of intonation. Therefore, this aspect of the analysis answers the fourth research question on the areas to expect correlation between tone and intonation, and justifies the objective of the study that was set to indentify the nature and point of relatedness of the suprasegmental features.

Expectedly, the results confirm that some challenges do manifest as the native Igbo speakers (tone language speakers) speak English (intonation language), although the challenges are few. Some variations are noticed in the utterances of the native Igbo speakers of English in the areas of pitch height which are lower than those of the Control. In the same vein, some of the utterances of the Consultants last longer than those of the English. The reason for these variations can be attributed to

the vowel types and vowel quality of the English vowels, some of which are noticeably different from the Igbo vowels. Secondly, English, being stress-timed as against Igbo that is syllable-timed contributed to the variations noticed in the rhythm of the utterances. With these explanations, the fifth research question is answered, and the fourth objective justified. Both were set to ascertain if and where native Igbo speakers encounter problems in articulating non-segmental features of English.

Finally, our findings have all helped to answer the research questions and justify the objectives of this study. We have as well succeeded in presenting the necessary information about the interface of intonation and tone using the speeches of native Igbo speakers of English. It can therefore be asserted that this study confirmed that virtually all languages of the world are intonational.

5.2 Conclusion

This study adopted an instrumental (acoustic) approach to the analysis of speech. Therefore, live utterances of native English speakers and those of native Igbo speakers were recorded and studied. The test frame is English (intonation) language and the objective was to establish areas of relatedness in the prosodic features of pitch, stress, tone and intonation. So, the utterances of the native English speaker(s), here tagged the Control, served as the benchmark of the suprasegmental phonology of English upon which the utterances of the native Igbo speakers (two males and two females) were juxtaposed and areas and points of similarities identified.

One fact that the study establishes is that intonational features are also present in tone languages, such as use of pitch to distinguish meaning over syntactic units. Thus, intonation is described as a suprasegmental feature superimposed on tones in tone languages. However, the mode of occurrence of these suprasegmental features may vary across languages as is the case in the tone (Igbo) and intonation (English) languages of study. All the same, the degree of variation in the languages of study is very minimal in most cases.

The analysis of word stress has revealed some relatedness between stress and tone. Although they are different features that manifest different characteristics, one sure area of correlation is that the syllable with high tone in tone language records high pitch and great prominence as does the stressed syllable in the intonation language of study. This is noticed in the pitch contours of the stressed syllables where the stressed syllables indicate the highest pitch values of the utterances. In the case of nuclear tones of the sentence types used in the study, it has become obvious from the utterances of the native Igbo speakers that tone is equally used in distinguishing meaning over syntactic units, hence, the similarity in the fall and rise nuclear tones of the statements with definitve fall and yes-no quesitions respectively; and the apparent similarity in the fall-rise tone of the tentative statements.

This study has, therefore, contributed to knowledge in linguistic studies. It will also benefit phonologists, semanticians and scholars of acoustic phonetics, including language teachers, broadcasters, legislators, actors and actresses and indeed some literate Nigerian public particularly those of the Igbo extraction interested in acoustic phonetics, as the English language remains the official language of Nigeria.

It must be noted, however, that the findings of this research are not conclusive about acoustic analysis of speech or interface of tone and intonation. The study has only added to the few available information on acoustic phonetic studies and remain open for improvement following further studies especially in this area of interface of tone and intonation.

REFERENCES

Abercrombie, D. (1967). *Elements of General Phonetics.* Edinburgh : Edinburgh University Press.

Abiodun, M. and Ajiboye, O. (2006). "A Phono-Syntactic Account of the High Low Tones Simplification in Moba Dialect". Paper presented at the 20th Annual Conference of the Linguistic Association of Nigeria held at Nigerian Educational Research & Development Council, Abuja, Nigeria; November, 13th to 17th, 2006.

Abramson, A. S. (1975). 'Tones of Central Thai: some perceptual experiments', In J. G. Harris and J. R. Chamberlain (Eds.). *Studies in Tai Linguistics in Honour of William J. Gedney* pp. 1-16. Bangkok: Central Institute of EnglishLanguage, Office of State Universities.

Agbedo, C. U. (2000). *General Linguistics: An Introductory Reader*. Nsukka: ACE Resource Konsult.

Anagbogu, P. N., Mbah, B . M. & Eme, C.A. (2010). *Introduction to Linguistics* (Second Edition). Awka: Amaka Dreams.

Anidobe, C. J. (2007). "Contrastive Analysis of Igbo and Hausa Phonemes and Phonotatics". M. A. Thesis of UNN.

Anyakoha, M. N. (2010). *New School Physics for Secondary Schools*. Onitsha: Africana First Publishers.

Ashby, M. & Maidment, J. (2005). *Introducing Phonetic Science*. Cambridge: Cambridge University Press.

Bakare, C. A. (1995). "Discrimination and Identification of Yoruba Tones : Perception Experiments and Acoustic Analysis", In K. Owolabi (Ed.). *Language in Nigeria: Essays in Honour of Ayo Bamgbose*. Ibadan: Group Publishers.

Bamgbose, A. (1969). 'Yoruba', In E. Dunstein (Ed.), *Twelve Nigerian Languages*, pp. 163-172. London: Longmans.

Blake, B. T. (2008). *All About Language Today*. London: Hutchinson University Library.

Bolinger, D. (1968). *Aspects of Language*. New York: Harcourt Brace Jovanovich.

Catford, J. C. (1994). *A Practical Introduction to Phonetics.* Oxford: Claredon Press.

Celik, M. (2001). "Teaching English to EFL/ESL Students", *The Internet TEST Journal, VOL. VII, No. 12, December 2001.* http://iteslj.org/Techniques/Celik-Intonation.html; Accessed 16th June, 2005.

Chang, N. T. (1972). "Tones and Intonation in Chengtu Dialect Szechuan China", In D. Bolinger (ed.), *Intonation*, pp.391-413. Harmondsworth: Pengiun Books.

Clark, J.,Yallop, C. & Fletcher, J. (2007). *An Introduction to Phonetics and Phonology*. Malden: Blackwell Publishing.

Clements, G. N. & Goldsmith, J. (1980). "What is Downstep? A Reply to Clark", In *Studies in African Linguistics,* Vol. 11, No. 2.

Clements G. N. & Rialland, A. (2008). "Africa as a phonological area", In B. Heine and D. Nurse (Eds).*A Linguistic Geography of Africa*. Cambridge: Cambridge University Press.

Collins, B. & Mees, I. M. (2003). *Practical Phonetics and Phonology: A Resource Book for Students*. London: Routledge.

Cruttenden, A. (1997). *Intonation*. Cambridge: Cambridge University Press.

Crystal, D. (1997). *The Cambridge Encyclopedia of Language*. Cambridge: Cambridge University Press.

Crystal, D. (2003). *A Dictionary of Linguistics*, Fifth Edition. Malden: Blackwell Publishing.

Dauer, R. M. (1983). "Stress-timing and syllable-timing reanalysed", *Journal of Phonetics 11:51-62.*

Dimmendaal, G. (1975). "Metatomy in Benue-Congo: Some FurtherEvidence for an Original Argument", In E. Nolue Emenanjo and Ozo-Mekuri Ndimele (Eds.), *Issues in African Languages and Linguistics (IALL): Essays in Honour of Kay Williamson.* Aba: National Institute for Nigerian Languages.

Du, T. C. (1988). "Tone and Stress in Taiwanese", Ph.D Dissertation, University of Illinois at Urbana Champaign.

Ejele, P. E. (2003). "Intonation in Tone Languages: Insights from Esan", In Ozo-Mekuri Ndimele (Ed.) *The Linguistic Paradise: A Fetschrift for E. Nolue Emenanjo.* Aba: National Institute for Nigerian Languages.

Emenanjo, E. N. (1978). *Elements of Modern Igbo Grammar: A Descriptive Approach*. Ibadan: University Press.

Gandour, J. T. (1978). 'The perception of tone', In V. A. Fromkin (Ed.), *Tone: a Linguistic Survey*, New York: Academic, pp. 94 – 105.

Gleason, A. A. (1979). *An Introduction to Descriptive Linguistics.* New Delhi: Oxford and DBH Publishing Coy.

Halle, M. & Vergnaud, J. –R. (1975). *An Essay on Stress.* Cambridge MA: MIT Press.

Harley, T. A. (1995). *The Psychology of Language: From Data to Theory*. Hove: Psychology Press.

Hombert, J. M. (1976). 'Development of tones from vowel height', In J. M. Hombert (Ed.), *Studies on Production and Perception of Tones*. University of California at Los Angeles; *Working Papers in Phonetics 33*: 55 – 56.

Hornby, A. S. (1975). *Oxford Progressive English for Adult Learners.* Oxford: Oxford University Press.

Hyams, L. M. (1975). *Phonology: Theory and Analysis*. New York: Holt, Rinehart and Winston.

Ikekeonwu, C. I. (1993). "Intonation and Focus: A Reanalysis of Downdrift and Downstep in Igbo", In *Working Papers.* Department of Linguistics, Lund University, 40: 95 – 113.

……..... (2003). "Semantic Categorization in Tonal Assignment: An Investigation of Hausa, Igbo and Yoruba", In Ozo-Mekuri Ndimele (Ed.), *Four Decades in the Study of Languages and Linguistics in Nigeria: A Festschrift for Kay Williams*. Aba: National Institute for Nigerian Languages.

Ikekeonwu, C., Ezikeojiaku, P. A., Anosike, U. & Ugoji, J. (1999), *Fonoloji Na Grama Igbo.* Ibadan: University Press.

Inkelas, S. (1999). "Exceptional Stress-attracting suffixes in Turkish: Representations versus the grammar", In R. Kager, H. Van der Hulst and W. Zoneveld (Eds.), *The Prosody-Morphology Interface,* 134 – 187. Cambridge University Press.

Jones, D. (2006). *Cambridge English Pronouncing Dictionary*. Cambridge: Cambridge University Press.

Kelly, J. (1969). 'Urhobo', In E. Dunstein (Ed.), *Twelve Nigerian Languages.* London: Longmans Green, pp. 153 – 62.

Kentowicz, M. (1994). "Sonority-driven Stress", In *Rutgers Optimality Archive.* ROA -172.

Kratochvil, P. (1968). *The Chinese Language Today.* London: Hutchinson University Library.

Ladefoged, P. (1962). *Elements of Acoustic Phonetics.* Chicago: University of Chicago Press.

…….. (1971). *Preliminaries to Linguistic Phonetics*. Chicago: The University of Chicago Press.

…….. (1982). *A Course in Phonetics.* Boston: Thomas Wadsworth.

…….. (2003). *Phonetic Data Analysis: An Introduction to Fieldwork and Instrumental Techniques.* Malden: Blackwell.

…….. (2006). *A Course in Phonetics.* Wadsworth: Cengage Learning.

Laver, J. (1994). *Principles of Phonetics.* Cambridge: Cambridge University Press.

Leech, G. N. (1969). *A Linguistic Guide to English Poetry*. London: Longman.

Liebeman, P. & Blumstein, S.E. (1988). *Speech Physiology, Speech Perception and Acoustic Phonetics*. Cambridge University Press.

Lyons, J. (1995). *Linguistic Semantics: An Introduction.* Cambridge: Cambridge University Press.

Maidment, J. A. (1990). "Focus and Tone in English Intonation", In S. Ramsaran (Ed.), *Studies in the Production of English*. London: Routledge, 19 - 26.

Mbah, B. M. (1996). "Tonemic Commutability and Fission: The Case of Register Tone Languages; Ph.D. Seminar Paper of the Department of Linguistics, University of Nigeria, Nsukka.

…… (2005). "Tone and Syntactic Integrity in Igbo", In B. M. Mbah (Ed.), *Nsukka Journal of African Languages and Linguistics (NJALL)*, Vol. 2, No.2.

Mbah, B. M. & Mbah, E. E. (1999). *Topics in Phonetics and Phonology: Contribution from Igbo.* Nsukka: AP Express Publishers.

…….. (2010). *Topics in Phonetics and Phonology: Contributions from Igbo.* (2nd Edition). Nsukka: AP Express Publishers.

Mutaka, N. N. & Tamanji, P. N. (2000). *An Introduction to African Linguistics.* Muenchen: E. C. Lincom Europa.

O'Connor, J. D. & Arnold, G. F. (1973). *Introduction of Coloquial English*. London: Longman Group.

Okorji, R. I. (1998). "Aspects of Tonology in Umuchu Dialect: A Syntactic Appraisal", *Nsukka Journal of the Humanities*, No. 9.

……(2002). "The Inland West Igbo Dialects: A Phono-Prosodic Appraisal", Ph.D. Thesis, Department of Linguistics and Nigerian Languages, UNN.

Oluikpe, B. O. (1979). *Igbo Transformatonal Syntax: The Ngwa Dialect Example.* Onitsha: Africana Publishers.

Oyebade, F. O. (2006). "Prosodic Structure Preservation and English Loanwords in Yoruba: A Constraint-Based Account". A Paper presented at the 20th Annual Conference of the Linguistic Association of Nigeria held at the Conference Centre of the NERDC, Abuja, Nigeria, (13th – 17th Nov. 2006).

Onwudiwe, G. E. (2014). "Juncture in speech: Illustrations from the Nnewi Dialect of Igbo"; Ph.D. Seminar of Department of Igbo, African and Asian Studies, Nnamdi Azikiwe University, Awka.

Partee, B. H., Meulen, A.T. & Wall, R.E. (1990). *Mathematical Methods in Linguistics*. Dordrecht: Kluwa Academic Publishers.

Peterson, G. E. & Barney, H . L. (1952). " Control methods used in a study of vowels", *JASA 24*: 175 – 84.

Pickett, J. M. (1980). *The sounds of speech communication: A Primer of Acoustic Phonetics and Speech Perception.* Baltimore: University Park Press.

Pike, K. L. (1948). *Tone Languages*. An Arbor: The University of Michigan Press.

Richards, J. C. & Schimidt, R. (2002). *Longman Dictionary of Language Teaching and Applied Linguistics.* London: Longman.

Roach, P. J. (1982). "On the Distinction between Stress-timed and Syllable-timed Languages", In D. Crystal (Ed.). *Linguistic Controversies.* London: Edward Arnold.

…… (1983). *English Phonetics and phonology: a practical course*. Cambridge: Cambridge University Press.

…… (2000). *English Phonetics and Phonology: A self-contained, comprehensive pronunciation course,* (Third Edition). Cambridge: Cambridge University Press.

Selkirk, E. O. (1980). "The role of prosodic categories in English word stress", *Linguistic Inquiry 11*: 563 – 605.

Sezer, E. (1981). "On Non-final stress in Turkish", In *Journal of Turkish Studies 5*: 61 – 69.

Tench, P. (1996)). *The intonation systems of English.* London: Cassell.

Trask, R. L. (1996). *Historical Linguistics.* London: Hodder Arnold Publication.

Udoh, I. I. L. (2003). *An Introduction to Phonemic Analysis*. Lagos: Concept Publications.

Uguru, J. O. (2000). "The Interaction of Tone and Intonation in Ika". M.A. Dissertation of the Department of Linguistics and Nigerian Languages, University of Nigeria, Nsukka.

…… (2006). "Stress and Tone in a Tone Language: Focus on Ika". A Paper presented at the 2nd Annual Conference of the Igbo Studies Assocaition held at the University of Nigeria, Nsukka.

…… (2006). "Acoustic Analysis of English and Ika Intonation Patterns: A Comparative Study". Ph.D. Dissertation of Department of Linguistics and Nigerian Languages, University of Nigeria, Nsukka.

Uzochukwu, S. (2001). *Traditional Funeral Poetry of the Igbo*. Akoka-Lagos: University of Lagos Press.

Wells, J. C. (2006). *English Intonation: An Introduction*. Cambridge: Cambridge University Press.

Welmers, W. E. (1973). *African Language Structures*. Berkeley: University of California Press.

Westerman, D. & Ward, I. C. (1990). *Practical Phonetics for Students of African Languages*. London: Kegan Paul International.

Williamson, K. (1993). "The Noun Prefixes of New Benue-Congo", *Journal of African Languages and Linguistics, 14.*

http://www.uni-potsdam.de/u/anglistik/carrol/Texte/SS/SS lectures 2-5.htm "English Segments & Syllables", Accessed 25th June, 2005.

APPENDIX

A: PRELIMINARY TO THE STRUCTURED INTERVIEW FOR COLLECTION OF DATA

Good day Sir/Madam!

This interview was designed to assist the researcher in carrying out a study on interface of tone and intonation in speech. Thank you for your anticipated co-operation.

(i) Respondent's Biodata

Please, supply the required information below:

Name…… Sex…… Nationality…….. Tribe……

Highest educational qualification……. Occupation……

Do you speak and write your native language well?.....

Can you speak English language well?.....

Now, you are going to read out the under listed words and sentences loud. Please, as much as possible, feel free and relaxed, and produce the words and sentences in the most natural way you speak (not minding that they are English words and sentences).

Thank you.

TEST MATERIALS

1(i) Word stress

(a) Two-syllable words	**(b) Three-syllable words**
sharpen (v)	imagine (v)
record (v)	interrupt (v)
autumn (n)	abnormal (adj)
complete (adj)	disaster (n)
below (adj)	transistor (n)

(ii) Compound words

(a) *First element adjectival, stress on second element*

loudspeaker

second-class

(b) *First element nominal, stress on first element*

typewriter

tea-cup

(c) *Mixture of (a) and (b)*

shoelace

gear-box

2 Stess placement in sentences

Statements: The *definitive fall*

The fall should be on the nuclear syllable

(i) I 'think it's ˎgreat

(ii) We're 'all ˎhere

(iii) The're 'waiting outˎside

(iv) It's 'half past eˎleven

(v) I'll 'go and get some ˎmilk

Tentative statements

(i) I ˬthink

(ii) I'll ˬtry

(iii) She ˬmight

(iv) They ˬcould

(v) ˬProbably

Yes-no questions

(i) 'Haven't they done e-nough?

(ii) 'Didn't you bring an um-brella?

(iii) ˈCouldn’t you send a -text message?

(iv) ˈIsn’t she -ready?

(v) ˈAren’t you going to intro-duce us?