

LEXICAL STRESS VARIATION IN ARABIC: AN ACOUSTIC SPECTROGRAPHIC ANALYSIS*

Salman H. Al-Ani

Indiana University, Bloomington

1. Introduction

The speech sounds of a human language are composed primarily of two types: segmentals and suprasegmentals. The segmentals consist of vowel and consonant sound units from which syllables and lexical items are formed. These sound units are usually arranged in a linear manner. Stress, pitch and length are superimposed on the syllables and are considered suprasegmentals. The phonetic and phonological bases of the sound system of any specific language are therefore established on the interaction of the segmentals and suprasegmentals. Even though length is considered a suprasegmental feature it can affect the vowels, consonants and syllables (Ladefoged 1982:14-5).

In Arabic long vowels and geminated consonants play a direct role in the formation and identification of syllables and syllable types. The placement of stress, pitch and length may interact intermittently. Sometimes one or more of these features is associated with the stressed syllable. Although the segmental and suprasegmental features operate in harmony and unison it is important to make the distinction between their two levels. Suprasegmentals or prosodic features are recognized, identified and "are established by a comparison of items in sequence ... whose arrangement in contrastive patterns in the time dimension is not restricted to single segments..." On the other hand, "the segmental features can be defined without reference to the sequence of segments in which the segment appears, and their presence can be established either by inspection or paradigmatic comparison..." (Lehiste 1970:2-3). The paradigmatic comparison is normally accomplished by comparing one item with others in the phonological inventory of the sound system of a specific language.

The speech sounds of a language are always in a state of flux and continuum. In the process of speaking, sounds are always interacting and are continually changing. The state of continuum of the speech sounds is directly associated with the movements and activities of the vocal organs that are used to produce speech. The acoustic physical nature of speech demonstrates this in a crystal clear manner. Just a quick glance at a spectrographic display will prove that speech is not distinctly produced as a discrete sequence of sounds. It is possible to identify individual sounds based on their acoustic characteristics. However, on a spectrogram display, it is difficult to pinpoint exactly where a sound ends and another begins in the sequence of an utterance of speech. In the same analogy, it is difficult in an utterance to separate completely the suprasegmental features from the segmentals. The three suprasegmental features — stress, pitch and length — may interact in an utterance in

* The author is conducting a series of experiments using the MacSpeech Lab II on stress and prosody in Arabic. This is the first article in the series. The analysis and results reported here are based on the experiment of *kitab/katib* (book/writer) recorded in isolation.

the phonological system of a specific language. The domain of stress is the syllable. The function of stress is on a contrastive basis and the operative level of stress requires a minimum of two syllables. The other prosodic features, pitch and length, may prevail at the stressed syllable either simultaneously or one at a time. The suprasegmental features are the focus of this investigation. In this production experiment of two lexical items, *kitāb* and *kātib* (book and writer, respectively), are tested in isolation.

2. Arabic language variety used in the study

The variety of Arabic used in this study is standard literary Arabic (*al-'arabiyya al-fuṣḥā*) as used today by educated speakers. This is not a dialectal study of stress accent. However, any possible linguistic overflow influence, whether regional or dialectal, is to be expected. Each participating speaker of the production experiment of *kitāb/kātib* was handed a sheet of paper containing a list of Arabic lexical items with complete vocalization. The lists adhered fully to the morphological structures of literary Arabic. Each lexical item was designed to be read in pause form or with full vowel endings. In pause form here the vowel endings of an utterance or the last word in a phrase are dropped. The endings of a full form will change into a pause form in the following style: *-a*, *-i*, *-u*, *-in*, and *-un* that all become zero and are not pronounced. Example: *kitābun*, in full form, becomes *kitāb*. Also, the ending *-an* changes into *-ā*. The following feminine case endings: *-ata(n)*, *-atī(n)*, and *-atu(n)* will change into *-ah* (Yushmanov 1961:15). In the design of the *kitāb/kātib* utterances special attention was given to capturing the variations in their pronunciation in pause and full form since both styles are practised in literary standard Arabic in the Arab world today.

3. Material

This study is based on the words *kitāb*, book and *kātib*, writer. Suffixes and prefixes were attached to produce eight lexical items from the stems. In the process of adding the affixes the syllabic configurations and types were also changed as indicated on the following *kitāb kātib* in isolation chart.

<i>kitāb</i> (book)	<i>kātib</i> (writer)
1. <i>ki tāb</i> cv c̄v̄	1. <i>kā tib</i> c̄v̄ cv̄
2. <i>ki tā bun</i> cv c̄v̄ cv̄	2. <i>kā ti bun</i> c̄v̄ cv̄ cv̄
3. <i>'al ki tāb</i> cvc cv c̄v̄	3. <i>'al kā tib</i> cvc c̄v̄ cv̄
4. <i>'al ki tā bu</i> cvc cv c̄v̄ cv̄	4. <i>'al kā ti bu</i> cvc c̄v̄ cv̄ cv̄

The experiment is controlled by using only *kitāb* and *kātib* as the basic stem components. In testing the suprasegmental features of the eight items the configuration and variation of the syllables created a controlled phonetic atmosphere.

4. Subjects

Six informants were asked to read a list of vocalized Arabic words. The speakers were instructed to read in *al-'arabiyya al-fuṣḥā* (standard literary Arabic), each item five times. The items were recorded on tapes and then transferred to disks. The speakers come from three different Arab countries — Iraq, Saudi Arabia and Morocco. They are all educated males and are identified as: SHA and SM (Iraq); SA and AZS (Saudi Arabia); and LH and MH (Morocco).¹

5. Procedure

Each one of the eight utterances of this production experiment was recorded five times by each speaker. The recordings were made by using a TEAC A-3440 tape recorder and a REALISTIC unidirectional dynamic microphone. The material was recorded on reel-to-reel tapes and then saved on three and one half micro floppy disks. From these disks the material was then filtered and digitized on the MacSpeech Lab II computer system that is designed by GW Instruments Software. Spectrograms (both broad and narrow bands) were generated from the waveforms of the different utterances. Also, spectrums of the segment energy envelope displays were produced from the waveforms. The segment energy displays of the tested utterances as well as the spectrograms of the sample waveforms were generated by using the MacSpeech Lab II. Measurements were made from broad band spectrograms, in milliseconds, for the duration of the entire utterance of each syllable and the duration of each vowel. The frequency of the successive syllables in each utterance was measured in Hertz (Hz) at the steady state of the vowel which is the nucleus of the syllable. This mode is obtained through the fundamental frequency (Fo) in the MacSpeech Lab II system where a horizontal wavy line appears. This Fo plot display of the spectrogram menu mode is the basis for the measurement of the frequency in the MacSpeech Lab II system. The Fo corresponds to the rate of the speaker's vocal cords vibration and the plot display of the Fo shows the variation of Fo with time. The MacSpeech Lab II system also allows measurement of the intensity unit. The measurement range scale runs from zero to 100.² The intensity basically corresponds to the amount of energy or effort expended by the speaker. Through the *spectrum* command of the MacSpeech Lab II system the user can obtain the *energy segment* plot display from the waveform of an utterance, going from the left to right on the screen, in the shape of varying peaks and valleys. The size of the plot display of the peaks and their heights depends on the amount of energy exerted in the execution of that particular syllable relative to the other syllables surrounding it in an utterance. Also, the size of the plot display depends on the number of syllables in the utterance being tested. A smaller number of syllables in an utterance shows more details than a larger number.

An information sheet was developed to record the measurements of:

¹ I would like to thank the speakers who participated and gave generously of their time to this study: Lahcen Haddad, Mohammad H irchi, Abdulrahman Al-Shamrani, Abdulaziz Alsebaill and Shaker Mustafa.

² The values of the intensity unit measurements (from zero to 100) is the system used in this study. However, it can be converted into volts by squaring the intensity units. For example: 100 intensity units equals 10 volts. Also, the intensity unit measurements can be converted into amplitude display values.

1) duration, in milliseconds, that includes: a) the entire utterance, b) every syllable of the utterance and c) every vowel of the utterance; 2) frequency in Hertz of the steady state of the vowels of each syllable and 3) the intensity in volts of a) intensity unit of the steady state of every vowel of every syllable and b) the highest point of the intensity unit of every syllable with regard to the utterance.

The sheets were completed for every utterance of the production experiment for each of the six speakers engaged in this study. It is important to mention that measurements were made and recorded of only one token of the five utterances. Hand written notes on the sheets were made at the time of measurement of any unusual observations with regard to other tokens of the same utterance. Measurements were not made of the first and fifth token recordings. It was noted that these two tokens contain more discrepancies than the other three and therefore were not considered as the best items to represent the utterance under consideration. Through the experience of examining the material it was observed that the first utterance was pronounced with more emphasis relative to the rest of the items. Likewise the fifth, being the last, was spoken with less emphasis, particularly the last syllable or two of the utterance. This observation is more noticeable with the intensity measurements. Consequently, all of the recorded measurements of the selected tokens of each utterance were made of the second, third or fourth item. It should be emphasized here that the figures used throughout the research for the intensity measurements are only relative. These figures are valid when they indicate differences of the intensity unit among the different syllables of an utterance. They change with the volume of the recording not only with different speakers but even with the same speakers recorded at different times. Therefore, the value is not in the absolute figures reported here but rather in the relative differences that exist among the measurements of the intensity units of the syllables of an utterance.

Every informant pronounced each of the eight lexical items five times. This totalled 40 items chosen and tested for each speaker. Thus the total number of items tested for the production experiment is 240.

6. Structure of the Syllable

The normal domain of the placement of stress is the syllable. In this regard it becomes imperative to know the number and types of syllables that a lexical item contains. There are six possible syllable types in Arabic. The onset of each of these types always consists of a single consonant and the coda consists of zero, one or two consonants. The two consonant clusters or geminates occur only with lexical items in pause forms. The peak or nucleus of every syllable, that is the most prominent element of the syllable, is always represented by a long or short vowel. The six possible syllable types found in Arabic are:

- | | | |
|--------------------|--------|------------|
| (1) CV | /bi/ | "in, at" |
| (2) C \bar{V} | /mā/ | "what" |
| (3) CVC | /min/ | "from" |
| (4) C \bar{V} CC | /dars/ | "lesson" |
| (5) C \bar{V} C | /bāb/ | "door" |
| (6) C \bar{V} CC | /mārr/ | "passerby" |

The distribution and frequency of occurrence of the above syllable types vary considerably. The first three types occur more often than the last three. The most frequent type used is the CV and the least used is the CVCC (Al-Ani and May 1973:37-49; repr. 1978:113-125). The lexical items of the acoustic experiments of this study utilized only the first four syllable types mentioned above. The minimum number of syllables used is two and the maximum four. It is interesting to note that the syllable types CVCC and CVCC are not utilized in the configuration of the syllable arrangement of the items tested for this experiment. This is not unexpected as these two types occur only in pause form of lexical items or phrases that contain these types of syllables. Thus, they are limited in their distribution and do not occur in the phonological sequence of utterances initially or medially.

7. Acoustic Analysis

The acoustic analysis begins by examining the suprasegmental features. This is achieved by retrieving and saving on diskettes the Arabic utterances that were recorded by the different speakers. The utterances were filtered, digitized and then saved in the form of soundwaves on the MacSpeech Lab II computer either on the hard disk or on micro floppy disks. These soundwaves can be retrieved and analyzed for the various acoustic features of speech, whether segmentals or suprasegmentals. The soundwaves are "caused by actions of the speaker's vocal organs that are (for the most part) superimposed on the outgoing flow of lung air" (Ladefoged 1982:166).

The first step was to establish the point in time of the steady state for every syllable of each utterance from the broad band spectrograms. Normally the steady state for each syllable is represented in the lower formants of the vowel of each syllable. Data information tables were developed to record the measurement results. The aim was to measure the frequency, intensity, and duration of the individual syllables of every lexical item and phrase tested. A duration measurement was made for the entire utterance, for every syllable of the utterance, as well as the duration of measurement for each vowel. The measurement for the frequency was made at the steady state and also at the highest point at the peak of each syllable.

For every utterance five tokens were recorded for each speaker and examined individually. However, only one of the tokens was chosen for the measurements. A free hand sketch of the intensity unit plot display of the chosen utterance was drawn on the data information table.

The acoustic analysis deals with the main components of the investigation of stress accent and other related features. These include (1) stress/intensity, (2) pitch/frequency and (3) length/duration.

7.1 Stress and intensity

One of the powerful features of the MacSpeech Lab II's analysis capabilities is to retrieve the time waveforms that are already digitized and stored in the MacIntosh memory system, whether the utterances are saved on a hard disk or on micro floppy disk. This system allows the researcher to view, in high resolution displays, the energy envelope plots of the utterance under investigation. The measurements of the different syllable peaks are made in intensity units. The scale of the measurements ranges from zero to 100. The successive syllables on the energy envelope plots, starting from the left to right on the screen of the MacSpeech Lab II from which printouts can be

made, appear in the form of peaks and valleys with the peaks reflecting the highest intensity. It is to be understood that the highest peaks, reflecting the most energy, are associated with the stress syllables. Also, on the display plots, the peaks not only indicate the vertical volume of the exerted energy in each syllable but the relative horizontal dimension duration, of that particular syllable, in milliseconds as well.

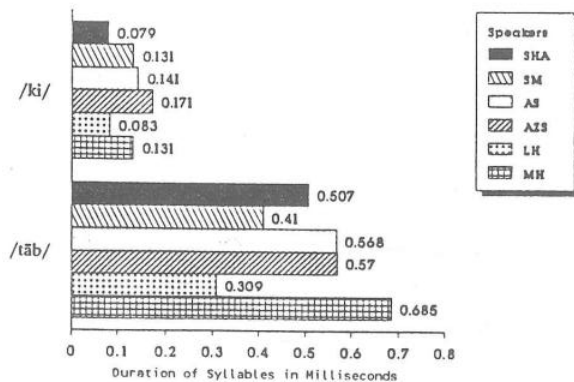
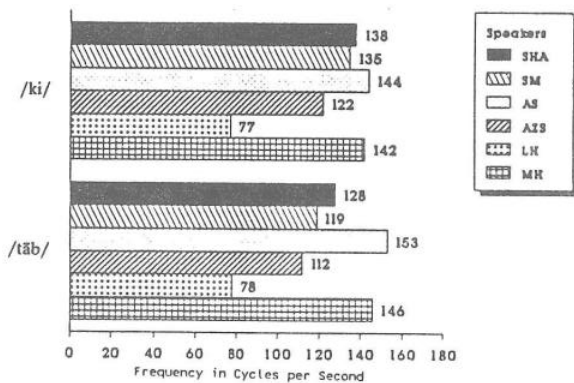
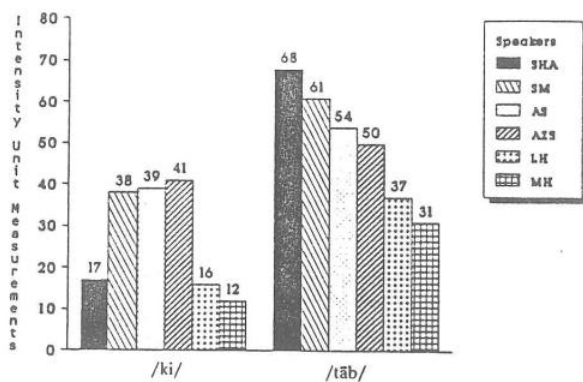
Two measurements were made of the intensity units of each syllable of the utterances in this study. The reason for making and recording two measurements, even though only one is reported here, was to be sure to capture the proper higher points of the syllable peaks. The results, reported below, of the measurements in intensity units are made at the steady state points obtained from the vowels of each syllable through the broad band spectrograms. Reference is made to the measurement of the highest point of the intensity of each syllable whenever there is a considerable difference between the two intensity unit measurements.

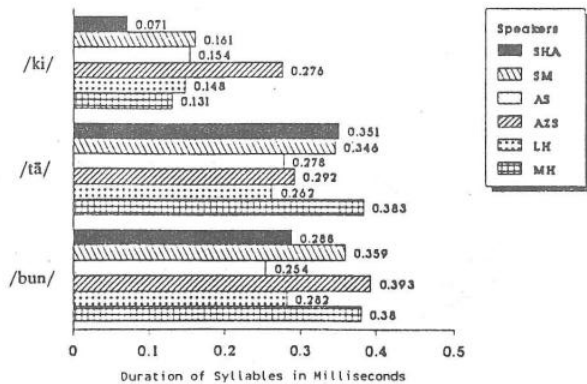
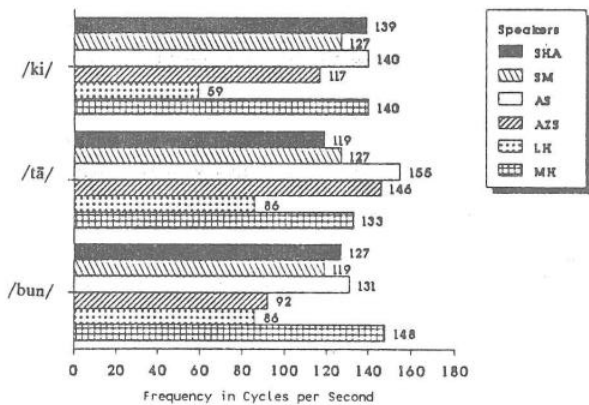
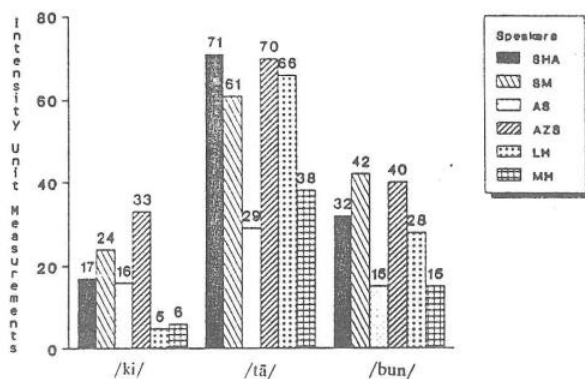
7.2 Pitch and frequency

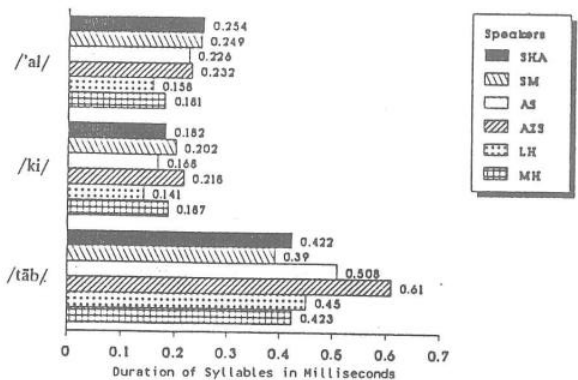
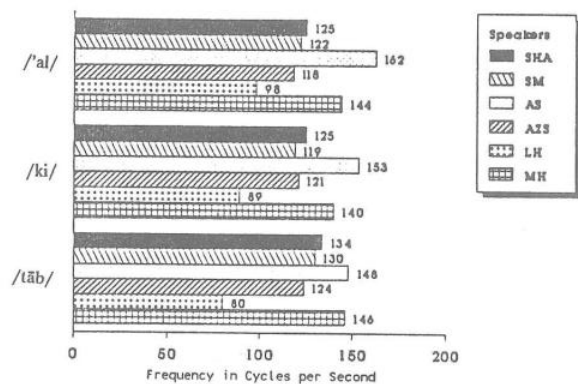
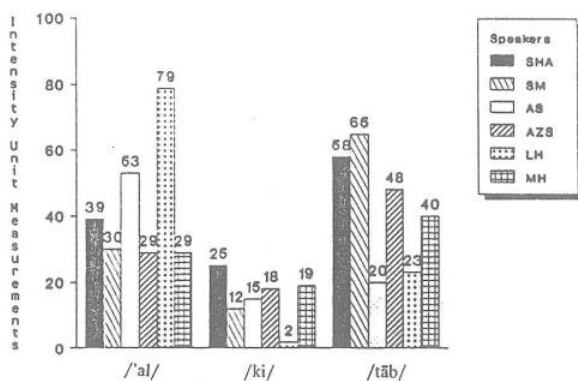
Frequency measurements were made for every syllable in Hertz at the steady state of the vowels of the utterances. Fundamental frequency (F_0) plots obtained through the spectrogram layouts of the MacSpeech Lab II system were generated for each utterance. The figures of the frequency measurements are not listed here. However, the measurements' results will be utilized in the analysis of the suprasegmental features of the production experiment utterances. This will be accomplished by comparing the F_0 of every syllable with the intensity measurements to ascertain whether the syllable that received the highest intensity unit also received the highest frequency measurement. Discrepancies, when a certain syllable in an utterance receives the highest intensity but does not receive the highest F_0 , will be discussed later.

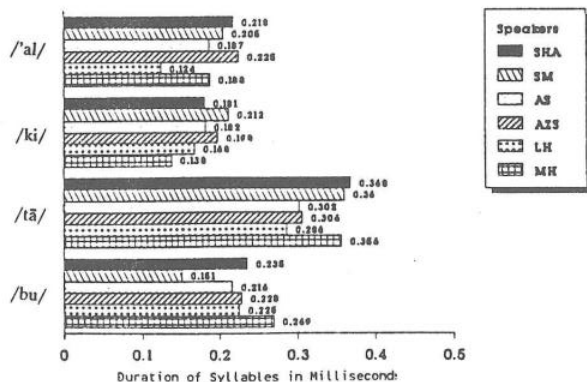
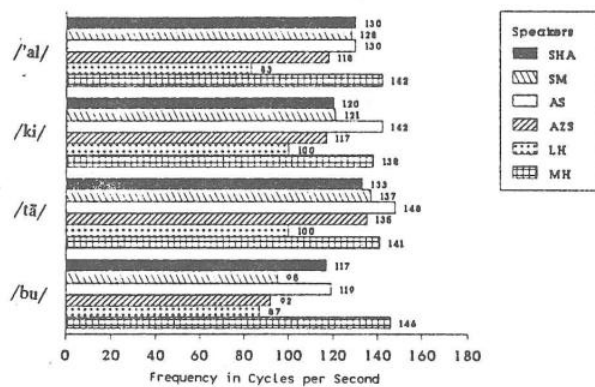
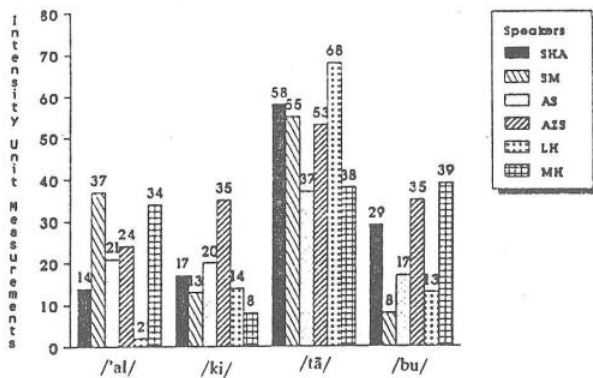
7.3 Length and duration

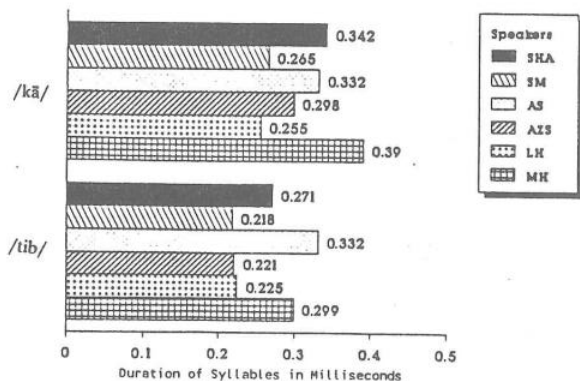
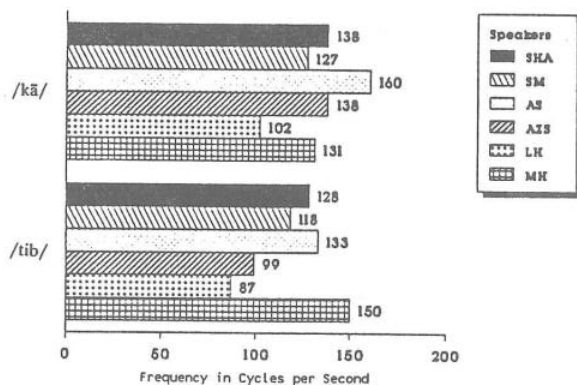
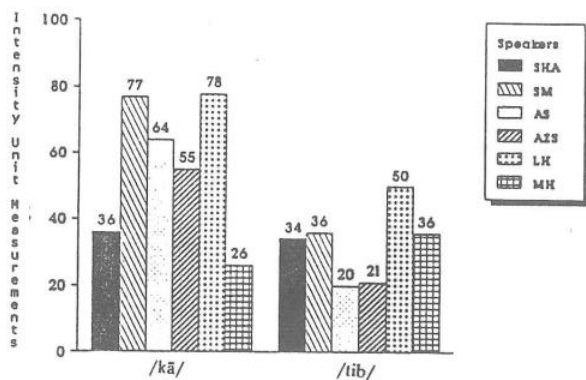
Measurements of the time duration were made, in milliseconds, for every utterance, for every syllable of each utterance and for every vowel of each syllable. Measurements of the duration were obtained from the broad band spectrogram plots. The figures for these measurements will not be listed here but will be used in the analysis.

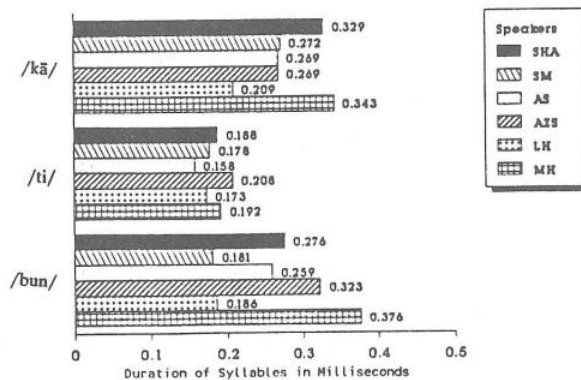
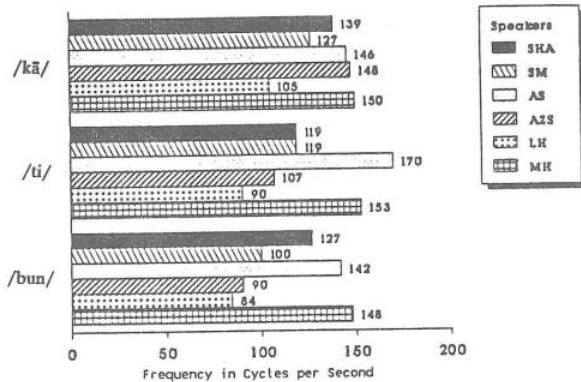
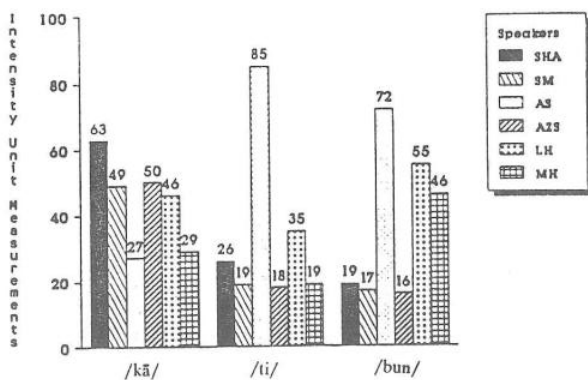


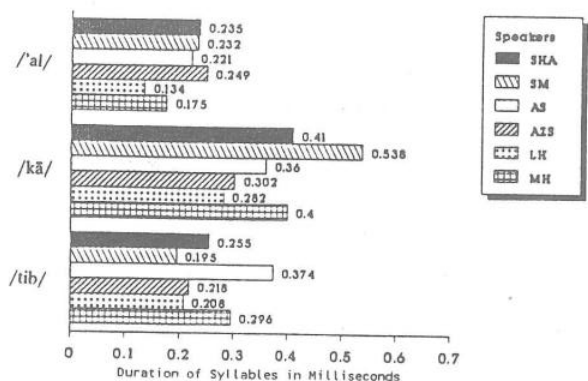
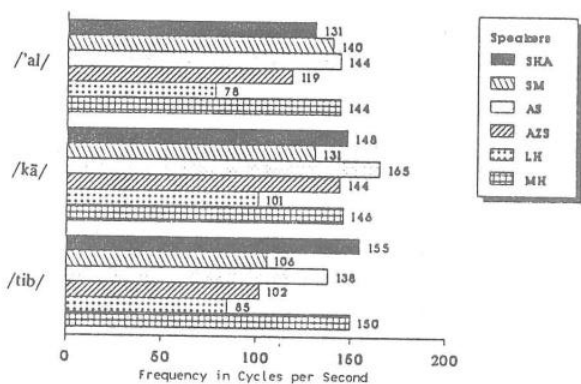
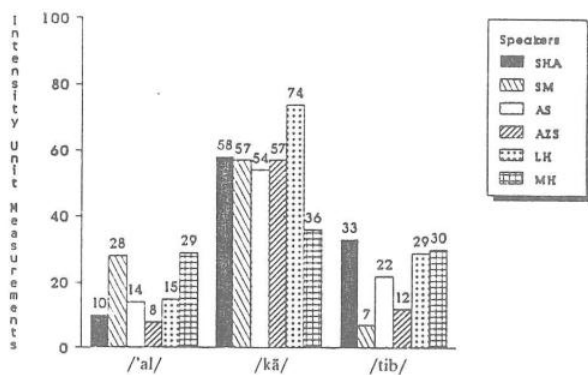


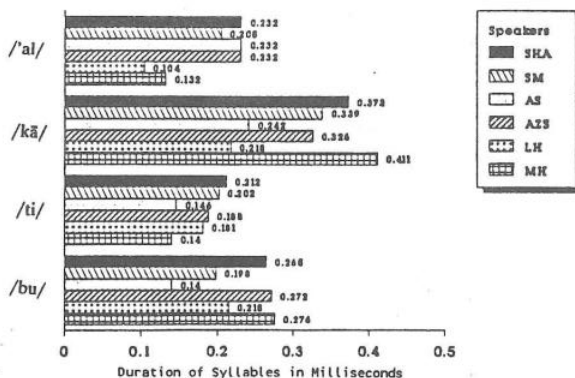
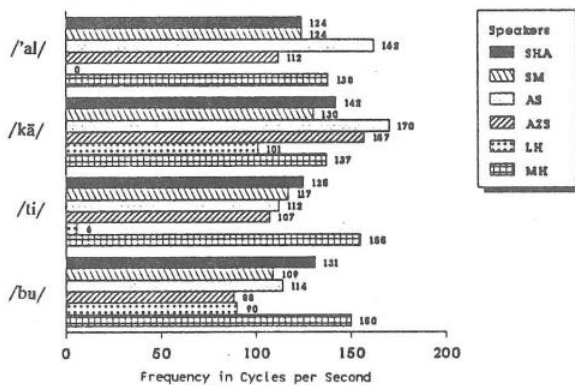
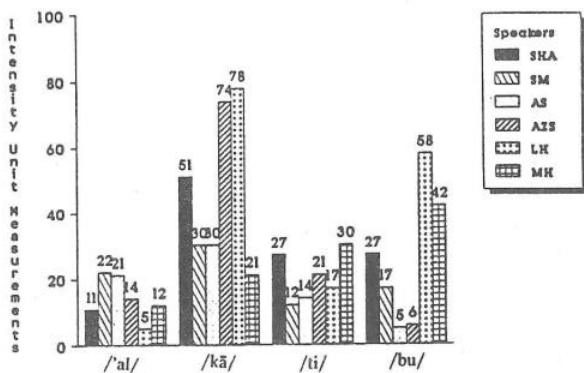












8. Interpretation of Graphs

The measurements of intensity, frequency and duration of the *kitāb/kātib* production experiment for the six speakers are presented graphically from pages fifteen through twenty-two. Every word of the eight different items of the experiment is shown with different columns identifying each speaker. For instance, the solid black column throughout the graphs represents speaker SHA. The figures at the top of the columns of the graphs reflect the actual measurements obtained from the waveforms produced by the MacSpeech Lab II system. It is recommended to identify the columns and match them with the keys provided next to them. The intensity columns are always shown vertically whereas the frequency and duration columns appear horizontally. The reader is advised to interpret the measurement information of the columns on each page both collectively and simultaneously. For instance, the graphs for the word *kitāb* on page 15 demonstrate graphically, in the form of columns, the intensity measurements for the six speakers. The columns reflect the personal characteristics and style of speaking, the mode of recording, the voice aspects, and possible other factors for each speaker. Thus, the variations reflected on the graphs are due to these factors.

The individual columns of the syllable *ki* in the word *kitāb*, for all six speakers, are lower than the columns of the syllable *tāb* despite the relative differences and variations that exist in each set of the columns of these two syllables. Also, the same analogy applies for the duration measurements for the word *kitāb*. The measurements, in milliseconds, of the syllable *ki* for the six speakers are relatively shorter than the duration of the syllable *tāb*. However, the measurements of the frequency exhibit more variation for the syllable *ki* for the six speakers, as apparent on the graph.

The graphs represent the core of the production experiment and should be the focus of interpretation of the data. The figures of the measurements for intensity, frequency and duration should not be taken as absolute but rather in a relative and comparative sense.

9. Summary of Results of the Production Experiment

9.1 Stress variation of *kitāb/kātib*

It is assumed that the syllable that receives the highest intensity and duration measurements in a lexical item is the stressed or accented syllable. However, frequency measurements do not seem to play as strong a role as the intensity and duration measurements. The syllable type and its distribution play a role in determining the primary stress placement. The experiment of this limited corpus reveals that the long syllables CV and CVC, in general, receive the highest intensity measurements. The one apparent exception and variation is speaker MH of Morocco who tends, in the majority of cases, to place the highest intensity on the last syllable.

9.1.1 Intensity

As it appears clearly on the graphs on pages 15-22, the long syllables CV̄ and CVC for the eight utterances of *kitāb/kātib* receive the highest intensity measurements for all six speakers except that in the lexical item *'alkitāb*, speakers AS and LH placed the highest intensity measurements on the first syllable, *'al*. Also, on the word

'*alkitābu*, speaker MH has about equal intensity measurements on the last two syllables. (See graph on page 18).

9.1.2 Duration

It is interesting and more revealing that the long syllables $\bar{C}\bar{V}$ and $\bar{C}\bar{V}\bar{C}$ also receive the highest duration measurements in milliseconds for the eight words of *kitāb/kātib*. The diversity among the speakers is almost minimal. The variations occur in *kitābun/kātibun* with one speaker, AZS of Saudi Arabia, who pronounced the last syllable, *bun*, longer in both words. Also, in *kātibun*, MH of Morocco made the last syllable, *bun*, longer.

9.1.3 Frequency

There is more variation among the six speakers with regard to the frequency measurements in the eight lexical items tested of *kitāb/kātib* than with the intensity and duration measurements. However, the long syllables of the types CV and CVC are still favoured by receiving the highest frequencies with most of the speakers. The one striking exception is speaker MH of Morocco who places the highest frequency on the last syllable of every word but one, *kātibun*, where he placed it on the second syllable *tī*. The two speakers SHA and SM of Iraq placed the highest frequency on the long syllables in the majority of cases. There is more variation in the assignment of stress between AZS and AS of Saudi Arabia.

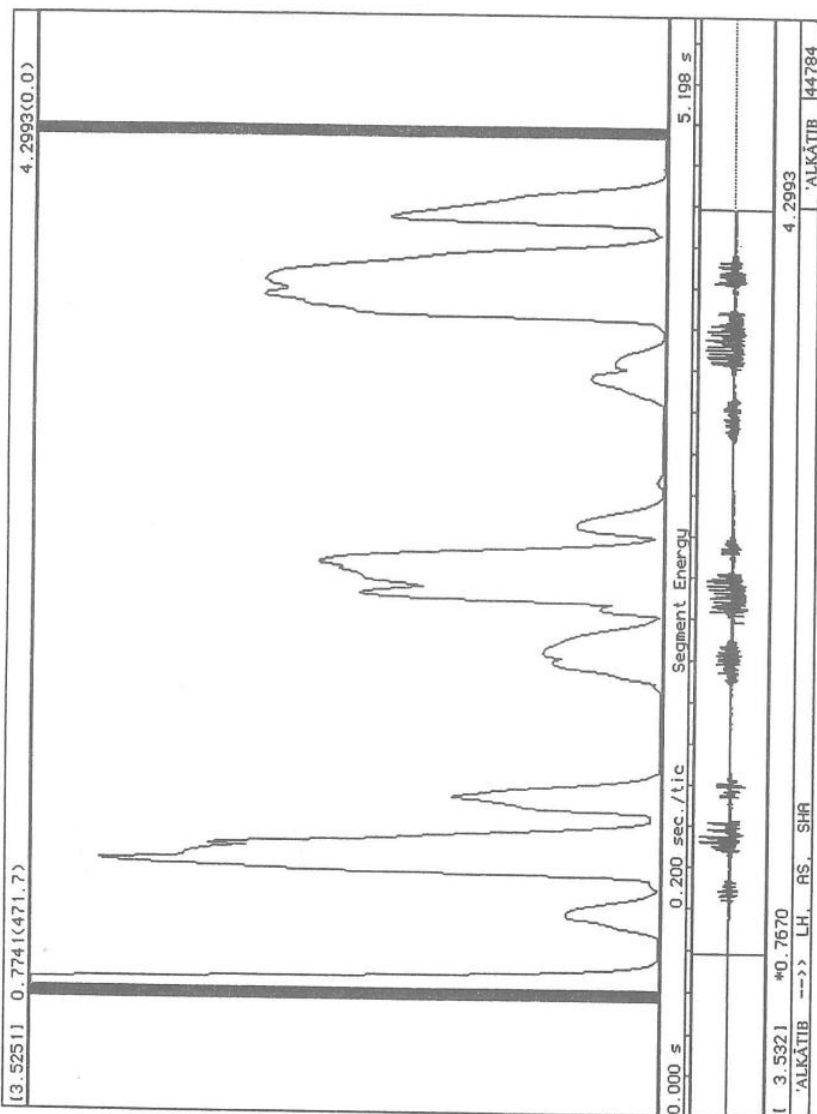
10. Previous Studies on Arabic Stress

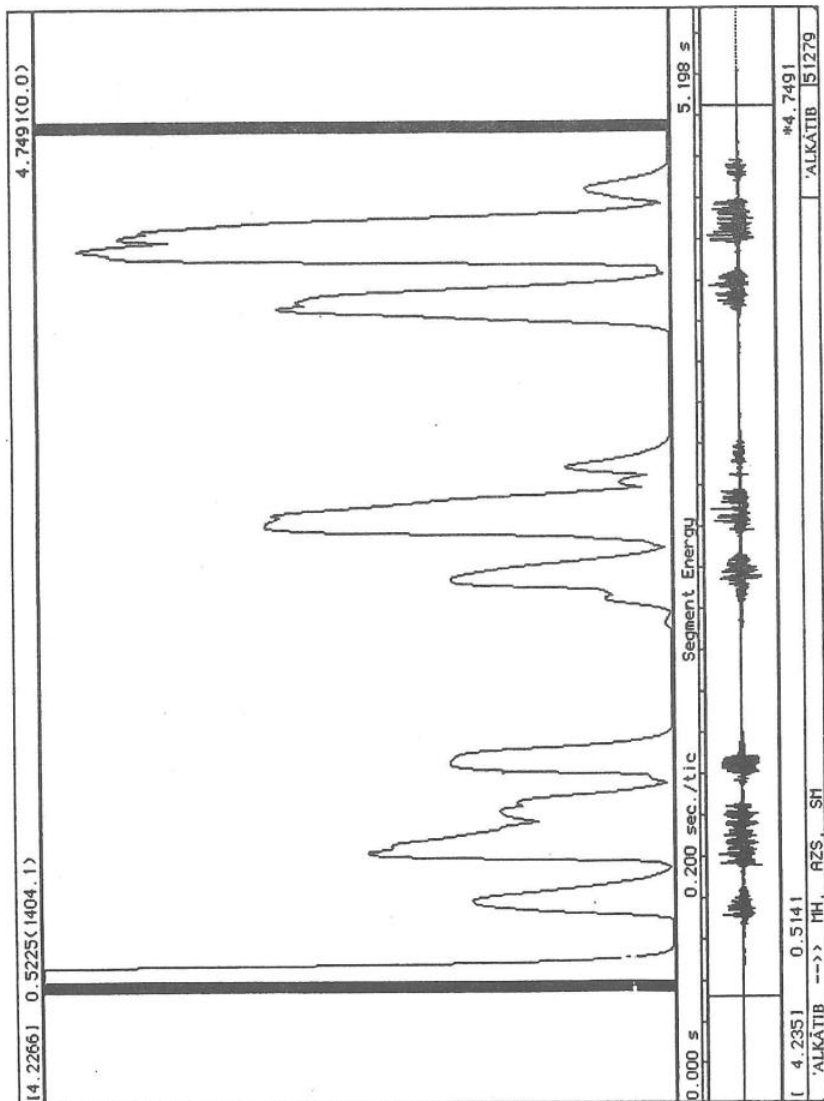
There are five studies that deal with the phenomenon of stress in Arabic. These are: *On Stress and Arabic Phonology* by Daud Abdo; *Arabic Phonology: An Acoustical and Physiological Investigation* by Salman H. Al-Ani; "Computer-Aided Comparative Study of Stress in Modern Standard Arabic" by Fatima Al-Khalifa and Yousif El-Imam; *al-'Aṣwāt al-luġawiyya* by 'Ibrāhīm 'Anīs and *al-Luġa al-'arabiyya: ma'nāhā wa-mabnāhā* by Tammām Hassān. The stress rules formulated for these studies were summarized in the article by Al-Khalifa and El-Imam. In their article, after summarizing the stress rules for the above authors, they presented reformulated stress rules for Arabic that are based on a computer analysis of voice intensity of one Kuwaiti informant.

The results of this production experiment, tentative as they are, confirm to a greater extent the main stress rules of Al-Khalifa and El-Imam. Future work, that this investigator is pursuing, will build on the initial and tentative results presented in this study.

11. Sample displays of the word '*alkātib* for the six speakers

These displays shows the intensity measurements for the six speakers involved in this production experiment. (See pages 25 and 26). The first display illustrates the word '*alkātib* for LH, AS and SHA. The second display also illustrates the word '*alkātib* for MH, AZS and SM. The middle peaks, that represent the syllable *kā*, are the highest for all of the speakers.





REFERENCES

- Abdo, Daud. 1969. *On Stress and Arabic Phonology*. Beirut: Khayats.
- Al-Ani, Salman H. 1970. *Arabic Phonology: An Acoustical and Physiological Investigation*. The Hague, Netherlands: Mouton & Co., "Janua Linguarum series practica 61. Translated into Arabic, 1983.
- Al-Ani, Salman H. and May, Darlene R. 1973. "The phonological Structure of the Syllable in Arabic". *American Journal of Arabic Studies* Leiden: Brill, 36-49; repr. in *Readings In Arabic Linguistics*, 1978. Salman H. Al-Ani, ed. Indiana University Linguistic Club, 113-125.
- Al-Khalifa, Fatima and El-Imam, Yousif. 1989. "Computer-Aided Comparative Study of Stress in Modern Standard Arabic." *Arab Journal for the Humanities*. Kuwait: The Academic Publication Council - Kuwait University, 34:9.313-30.
- 'Anīs, 'Ibrāhīm. 1961. *al-'Aṣwāt al-luġawiyya*. Cairo: Dār an-Nahḍa al-'arabiyya. 3rd ed.
- Ḥassān, Tammām. 1973. *al-Luġa al-'arabiyya: ma'nāhā wa-mabnāhā*. Cairo: al-Hay'a al-Miṣriyya al-'Āmma li-l-Kitāb.
- Ladefoged, Peter. 1982. *A Course in Phonetics*. New York: Harcourt Brace Jovanovich, Inc. 2nd ed.
- Lehiste, Ilse. 1970. *Suprasegmentals*. Massachusetts: The Massachusetts Institute of Technology.
- . 1961. "Some Acoustic Correlates of Accent in Serbo-Croatian." *Phonetica* 7.114-47.
- Yushmanov, N.V. 1961. *The Structure of the Arabic Language*, (Transl. from Russian by Moshe Perlmann). Washington, D.C.: Center for Applied Linguistics of the Modern Language Association of America.

Acknowledgments: The production experiment was conducted at the Phonetic Laboratory, Linguistics Department, Indiana University. I would like to thank Professor Robert Port for his valuable consultation and the time he spent with me. I would also like to thank Kathy Rogers and Nina Fales who introduced me to the Mac-Speech Lab and Salim Al-Ani who spent many hours helping me not only in the laboratory but in designing the beautiful graphs that are included in this paper. Dr. Stuart Davis, also of the Linguistics Department, deserves special mention for all of the time he spent in discussing the various aspects of this work with me.