

PSYCHOMETRICALLY EQUIVALENT ARABIC MONOSYLLABIC WORD
RECOGNITION MATERIALS

by

Maida Christine Robertson

A thesis submitted to the faculty of

Brigham Young University

in partial fulfillment of the requirements for the degree of

Master of Science

Department of Communication Disorders

Brigham Young University

August 2006

BRIGHAM YOUNG UNIVERSITY

GRADUATE COMMITTEE APPROVAL

of a thesis submitted by

Maida Christine Robertson

This thesis has been read by each member of the following graduate committee and by majority vote has been found to be satisfactory.

Date

Richard W. Harris, Ph.D., Chair

Date

Shawn Nissen, Ph.D.

Date

Ron Channell, Ph.D.

BRIGHAM YOUNG UNIVERSITY

As chair of the candidate's graduate committee, I have read the thesis of Maida Christine Robertson in its final form and have found that (1) its format, citations, and bibliographical style are consistent and acceptable and fulfill university and department style requirements; (2) its illustrative materials including figures, tables, and charts are in place; and (3) the final manuscript is satisfactory to the graduate committee and is ready for submission to the university library.

Date

Richard W. Harris, Ph.D.
Chair, Graduate Committee

Accepted for the Department

Ron W. Channell, Ph.D.
Graduate Coordinator

Accepted for the College

K. Richard Young, Ph.D.
Dean, David O. McKay School of Education

ABSTRACT

PSYCHOMETRICALLY EQUIVALENT ARABIC MONOSYLLABIC WORD RECOGNITION MATERIALS

Maida Christine Robertson

Department of Communication Disorders

Master of Science

The purpose of this study was to develop, digitally record, evaluate, and psychometrically equate a set of Arabic monosyllabic word lists to use in the measurement of the word recognition score. Familiar Arabic monosyllabic words were digitally recorded by a native male talker from Jordan who was judged to have a standard Arabic dialect. Twenty native Arabic participants with normal hearing were used as subjects to determine the percentage of correct word recognition for each word at 10 intensity levels ranging from -5 to 40 dB HL in 5 dB increments. The monosyllabic word data were analyzed using logistic regression. The words producing the steepest psychometric function were included in the final word lists. Four lists of 50 words each were created and eight half-lists (25 words each) were created from the four lists. A Chi-square analysis was performed, revealing no statistical differences among the lists and

half-lists. The mean monosyllabic psychometric function slopes at 50% for lists and half-lists were 4.8%/dB.

ACKNOWLEDGMENTS

I would like to thank everyone involved in this project—the willing participants, my committee, everyone—for being so willing to participate and for being so supportive and helpful. I would like to thank Rose, my partner during this whole project for her help and friendship, and yes—endurance! Thanks to my committee Dr. Channell, Dr. Nissen, and especially Dr. Harris for all their help and patience. Thanks to Reem and Mais, our interpreters for their hard work.

Thanks to my family and friends for their constant support, love and encouragement throughout the last year, and of course, thanks to David who has given me a new reason to finish.

TABLE OF CONTENTS

	Page
List of Tables	viii
List of Figures	ix
List of Appendixes	x
Introduction.....	1
Literature Review.....	2
Method	10
Participants.....	10
Materials	12
Calibration.....	14
Procedures.....	14
Results.....	15
Discussion.....	20
References.....	26

List of Tables

Table	Page
1. Age (in years) and Pure-tone Threshold (dB HL) Descriptive Statistics for the Twenty Arabic Participants.....	11
2. Arabic Male Monosyllabic Lists in Rank Order from Most Difficult to Easiest.....	16
3. Arabic Male Monosyllabic Half-lists in Rank Order from Most Difficult to Easiest..	18
4. Mean Performance of Arabic Male Monosyllabic Lists and Half-Lists.....	19

List of Figures

Figure	Page
1. Psychometric Functions for the Four Arabic Monosyllabic Lists for Male Talker Recordings	21
2. Psychometric Functions for the Eight Arabic Monosyllabic Half-lists for Male Talker Recordings.....	22

List of Appendixes

Appendix	Page
A. Informed Consent Research Participation Form.....	32
B. Description of BYU Arabic Speech Audiometry Materials CD.....	33

Introduction

A standard audiologic evaluation typically includes a measure of one's ability to understand and process speech, also known as speech audiometry. An essential component during a speech audiometry evaluation is determining the individual's word recognition score. Word recognition testing determines one's ability to correctly perceive and repeat a list of words presented at a suprathreshold level. Though word recognition tests are widely available in the United States and in a variety of languages, there are unfortunately none currently adequate to evaluate word recognition for patients whose native language is a standard Arabic dialect. Pure-tone testing is widely available in most countries and cultures because there is no language barrier, allowing the test to be performed using the same procedures and criteria across many languages. Though pure-tone testing is beneficial in determining a hearing loss, further evaluation is needed to determine the effect of hearing impairment on speech comprehension (Carhart, 1951; Cramer & Erber, 1974; Davis, 1948; Doyone & Steer, 1951; Erber, 1974; Hirsh, et al., 1952). Word recognition testing is able to determine the affects a hearing impairment has on one's ability to discriminate among different words. The problem lies in the language-based nature of the test. The test needs to be administered in the same language of the listener, if not, the results may be invalid. Word recognition testing is dramatically influenced by the phonetic, melodic, and intonational differences among languages (Carhart, 1951). In order to accurately determine one's word recognition ability, the test must be administered in the native language of the individual.

Lack of appropriate speech audiometry material becomes an obstacle, especially in developing countries. In developing countries, there is a high rate of hearing

impairment but insufficient trained audiologists, deaf educators, or technicians to provide full service intervention. According to the World Health Organization (WHO), the number of deaf and hard of hearing individuals in developing countries is twice as many than in developed countries (Smith, 2001). In the United States, the American Academy of Audiology (2004) has reported up to 3 in every 1,000 healthy newborns have a hearing impairment. For comparison, in Jordan, it is estimated that 6 in every 1,000 healthy newborns have a hearing impairment (Al-Masri, 2003). This high rate of hearing impaired infants is most likely due to the high rate of consanguineous marriages in the Arabic culture. These statistics may actually underestimate the prevalence of hearing impairment due to the lack of epidemiology services in developing countries. According to Smith (2001) the number of infants with a hearing impairment is estimated to be double what has been reported. According to the WHO, the prevalence of ENT's range from 1/30,000 to 1/150,000 in developed countries and 1/2,000,000 in less developed countries.

Though hearing impairment is more prevalent in developing countries, services and technology are still very limited (Jauhiainen, 2001; WHO, 1998). There is a need for adequate audiologic personnel, equipment and materials to correctly identify hearing losses early and thereby implement early intervention. The purpose of this study is to develop word recognition test materials for the assessment of patients who speak the Arabic language, specifically, for individuals from Jordan and Palestine who speak Arabic.

Review of Literature

Speech audiometry has been used to evaluate hearing long before the pure-tone

test was created. The first tests were not created for the purpose of evaluating the hearing impaired population, but for the evaluation of various communication devices. It was not until much later that some of the techniques were also found useful in evaluating the hearing impaired population. Since the first test, speech audiometry has progressed significantly. Adequate quantification of speech audiometry tests is a fairly recent accomplishment.

In the 1920s Campbell (as cited in Berger, 1971) developed speech audiometry techniques to evaluate the efficiency of telephone sound transmitting equipment. Campbell used a list of nonsense syllables transmitted over the phone. The listener would then repeat the nonsense syllables and compare their results with the original vocalization. This test became very popular and was known as the Standard Articulation Test. The main advantage to this test was its unbiased nature. The listener's previous vocabulary knowledge would not affect the outcome of the test. However, while this was a good technique to eliminate confounding variables, it also created one. The human ear is trained to listen for the specific sounds and sound combinations of their native language. It is unnatural and unrealistic to expect an accurate measure of speech comprehension when the listener is evaluated using syllables that have no meaning in their native language (Berger, 1971).

In 1929, Fletcher proposed the use of monosyllabic words for testing deaf populations (Fletcher, 1953). Monosyllabic words have a simple phonetic structure yet still have meaning and were thereby functional to the listener. At the close of World War II, the Harvard Psychoacoustic Laboratories created the PAL-PB 50 Lists to evaluate the

hearing impaired. This test was used in Army and Navy rehabilitation programs (Berger, 1971).

Since that time, numerous speech audiometry tests have been created, evaluated and used daily with hearing impaired populations. Historically, speech audiometry was primarily used to make audiological diagnoses: was the hearing impairment sensorinerual, conductive, or mixed? Although, these diagnoses are now made through much simpler techniques and tests, speech audiometry still plays a vital role in the audiological evaluation. Pure-tone testing is not capable of evaluating the functional hearing abilities of patients because our world does not communicate through simple tones, but through complex speech. Herein lies the value of speech audiometry. Currently, the primary role of speech audiometry is in monitoring auditory rehabilitation, and the assessment of hearing aids and social handicaps (Lyregaard, 1997). The ultimate success of auditory rehabilitation and hearing aids of any kind is based on the individual's ability to perceive and understand speech. The purpose of speech audiometry is to answer three questions: (a) What is the lowest intensity at which the listener can barely identify simple speech material? (b) How well does the listener understand everyday speech under everyday conditions? and (c) What is the highest intensity at which the listener can tolerate speech? (Berger, 1971). These questions and the ability to comprehend spoken language can be evaluated through speech audiometry.

The two basic components of speech audiometry are the Speech Reception Threshold (SRT) and the word recognition score. Each is an important and essential part of the audiological exam. The SRT is the lowest level at which the listener can identify familiar speech materials 50% of the time. Lists of spondaic words are presented to the

listener at various levels to determine the threshold. The SRT is valuable because it lends validity to results obtained from pure-tone testing (American Speech-Language-Hearing Association, 1988). Pure-tone thresholds and SRT typically agree with one another for both normal and hearing impaired populations. The SRT also has a high degree of validity (Carhart & Porter, 1971; Fletcher, 1929; Fletcher, 1950; Hughson & Thompson, 1942; Wilson, Morgan, & Dirks, 1973). The SRT is also used as a reference level for supra-threshold testing (word recognition) and often serves as the starting point in an audiological evaluation.

The word recognition score has had many names, including speech discrimination and suprathreshold speech recognition testing. For the purposes of this paper, the term word recognition will be used. This term better reflects the true nature of the test, since the listener is not asked to discriminate between words but to identify words. In word recognition, the test is conducted at intensity levels above the listener's SRT to determine the percentage of words the listener can correctly perceive and comprehend. This test is generally used to evaluate the listener's everyday hearing function. Unlike the SRT word lists, in the word recognition test, lists of monosyllabic words are typically used in English. Typically in pure-tone audiometry, the stimulus, or tone, is presented several times until the statistical average is found. In word recognition testing, the stimulus cannot be presented multiple times due to the possible influence of learned effects, thereby invalidating the results of the test. For this reason, in word recognition testing there are a list of words chosen and each stimulus is presented once (Lyregaard, 1997). Though word recognition has high face validity due to the functionality of the stimuli used, it has been proposed that perhaps using sentences would lend a higher validity,

since oral communication generally takes the form of sentences rather than single words. However, when sentences are presented as the stimulus, the results are a measurement of both central and peripheral hearing, thus possibly confounding the results. Due to this confounding variable, monosyllabic words are the best choice for word recognition stimuli due to their simple, yet functional nature (Lyregaard, 1997).

Several word recognition tests have been created for the evaluation of the hearing impaired. Common word recognition tests that are clinically used in English are: Central Institute for the Deaf W-22 (CID W-22; Hirsh et al., 1952), Northwestern University Auditory Test No. 6 (NU-6; Tillman & Carhart, 1966), Phonetically Balanced Kindergarten Test (PBK-50; Haskins, 1949), and Northwestern University Children's Perception of Speech (NU-CHIPS; Elliot & Katz, 1980). Similarities between all studies include the measures of correct responses as a function of intensity presentation and the use of familiar and high usage words. A common representation of word recognition test results is through a psychometric function as displayed by an ogive or S-shaped curve. This curve represents how an individual's hearing ability increases with an increased intensity level (Carhart, 1951).

Numerous studies have looked at specific aspects of word recognition that make tests more accurate. Multiple studies have found that speech materials produced by different talkers elicit varied psychometric functions (Beattie, Edgerton, & Svihovec, 1977; Beattie, Svihovec, & Edgerton, 1975; Carhart, 1965; Doyne & Steer, 1951; Hirsh, Reynolds, & Joseph, 1954; Krueger, Bell, & Nixon, 1969; Wilson & Carter, 2001; Wilson & Oyler, 1997). Carhart (1951) stressed the importance of using familiar words of the native language of the listener. In 1948, Egan proposed six criteria for selecting words to

formulate word lists for word recognition testing in English: (a) monosyllabic structured words, (b) equal average of difficulty among lists, (c) equal range of difficulty within lists, (d) equal phonetic composition among lists, (e) composition representative of spoken English, and (f) commonly used words.

From the studies of Carhart (1951) and Egan (1948), one can conclude that word lists for word recognition testing not only need to be familiar, monosyllabic words, but the presentation of words also needs to be unvarying, as well to satisfy Egan's criteria. That is, the presentation of word lists need to be similar across all listeners. This can most easily be done through digital recordings, and by using the same talker to record the test materials.

In the Arabic language, there are primarily two different studies which have addressed word recognition. Studies done by Alusi, Hinchcliffe, Ingham, Knight, and North (1974) and Ashoor and Prochazka (1982) are the only two word recognition tests available for Arabic speaking populations. Both studies used monosyllabic words for the word lists, based on the research done by Fry (1964) and Hirsh et al. (1952). In addition, Knight and Littler (1953) reported that monosyllabic words detect word recognition difficulty associated with hearing impairment better than other word structures. Studies by Alusi et al. (1974) and Ashoor and Prochazka (1982) used the common monosyllabic structures of CV, CVCC, CVVC, CVVCC with all 28 consonants and 6 vowels of the standard Arabic dialect. Standard Arabic was used to guarantee familiarity of words used since standard Arabic is used in schools and universities for teaching and is the language of the mass media. All words were selected from elementary school books, children's stories, daily newspapers, and excluded any absurd or technical words. Both studies

attempted to create phonetically balanced word lists on based on frequency of consonant and vowel occurrence. Alusi et al. (1974) divided 150 words into 6 phonetically balanced lists of 25 words each and Ashoor and Prochazka (1982) divided 120 words into 6 phonetically balanced lists of 20 words each. Ashoor and Prochazka collected familiarity ratings from a large number of participants from Saudi Arabia to ensure the commonality of the words included on the word lists. Ashoor and Prochazka recorded the words by analog in the standard Saudi dialect and Alusi et al. recorded the words in the standard Baghdad dialect. Though both tests were found to be valid for their select populations, they are not widely distributed. This may be due to the fact that there is a great shortage of audiological services in these countries. Another cause may be due to the dialectal differences in the Arabic language. The Arabic language has many different dialects, and test materials for one dialect may not be appropriate for accurate testing of another dialect, as implied by Carhart (1951).

Though studies done by Alusi et al. (1974) and Ashoor & Prochazka (1982) appear to be adequate for their respective dialects, Baghdad and Saudi, based on research discussed previously regarding efficient and accurate word recognition testing, a few challenges still remain. The primary dilemma is that each specific regional dialect of Arabic needs to have a specific test using that specific dialect. Furthermore, not all current tests take into account the complex characteristics of Arabic, standard versus colloquial dialects. There are 21 independent Arabic nations with an approximate population of 183 million Arabic speakers. Within this vast population there are varying dialects of the Arabic language, with a common dialect known as Modern Standard Arabic (American Association of Teachers of Arabic, 2003). The difference between

standard Arabic and colloquial dialects is known as *diglossia*. Diglossia is defined as a stable language situation where in addition to the typical dialect; there is a very divergent variation which is learned mostly through formal education and used only in written form and in formal spoken occasions. However, it is not used for ordinary conversations in the everyday language used in the community (Ferguson, 1959). Standard Arabic is the main language used by the mass media, and is the primary language used to teach in schools and universities whereas colloquial dialects are used in social settings. Therefore the general Arabic speaking population has been exposed to and is familiar with both Standard Arabic and colloquial dialects (Altoma, 1969; Fatihi, 2001; Ferguson, 1959). Standard Arabic is considered to be syntactically more complex and more prestigious than colloquial dialects of Arabic. Colloquial dialects are regionally based. The closer the regions are in physical proximity, the dialects are more compatible. The further apart the regions, the dialects are less compatible (Fatihi, 2001). Similarities between languages include numerous shared words and the basic syllable structure of words (Altoma, 1969). The differences between colloquial dialects and standard Arabic need to be considered before developing speech audiometry materials.

For the current study, standard Arabic was used since it is the language of formal oral communication and is typically understood by all Arabic populations. In addition, word lists were recorded using digital technology. While analog technology is adequate, digital technology is far superior and more appropriate and accurate for modern test situations. Digital recordings have increased channel separation, increased dynamic range, improved signal-to-noise ratio, reduced harmonic distortion, elimination of wow and flutter associated with tape recordings, and have a longer storage life without

distortion. Digital recordings can also be saved and modified efficiently and uniformly. With the use of custom software, digital recordings can be presented in varied intensity levels and can be easily randomized (Kamm, Carterette, Morgan, & Dirks, 1980; Ridgeway, 1986; Sony, 1991). Due to these advanced features, it is clear that digital recordings will provide a more accurate and uniform presentation of stimulus words.

All of the factors discussed above lead to a call for a more standardized, appropriate word recognition test for the standard Arabic language. Thus, this study will attempt to create psychometrically equivalent monosyllabic word lists following these criteria (Carhart, 1951; Egan, 1948). The creation of these materials will take into account the specific characteristics of the Arabic language to produce lists for accurate word recognition testing for patients whose native language is Arabic.

Method

Participants

Twenty native Arabic speakers participated in this study by evaluating the monosyllabic words; 8 females and 12 males. All participants originated from Jordan, Israel, or Palestine, ranged in age from 18 to 39 years ($M = 23.1$ years), and self-reported speaking Arabic regularly. Participants in this study possessed pure-tone air-conduction thresholds ≤ 15 dB HL (ANSI, 2004) at octave and mid-octave frequencies from 125 to 8000 Hz. Static acoustic admittance was between 0.3 and 1.4 mmhos with tympanometric peak pressure between -100 and +50 daPa (American Speech-Hearing-Association, 1990; Roup, Wiley, Safady, & Stoppenbach, 1998). A summary of participant thresholds are listed in Table 1.

Table 1

Age (in years) and Pure Tone Threshold (dB HL) Descriptive Statistics for the 20 Arabic

Participants

	<i>M</i>	<i>Minimum</i>	<i>Maximum</i>	<i>SD</i>
Age	23.1	18	39	4.9
125 Hz	1.8	-10	10	4.9
250 Hz	0.5	-10	10	5.4
500 Hz	4.0	-5	15	5.0
750 Hz	3.5	-5	15	4.3
1000 Hz	2.8	-5	10	4.1
1500 Hz	3.3	0	10	3.7
2000 Hz	1.8	-5	10	4.1
3000 Hz	2.3	-10	15	6.2
4000 Hz	1.8	-10	15	6.1
6000 Hz	1.0	-10	15	7.7
8000 Hz	-0.8	-10	15	7.1
PTA*	2.8	-5	8	3.2

*PTA = average pure-tone thresholds at 500, 1000, & 2000 Hz.

Materials

Words. Monosyllabic words were selected for the purpose of creating lists of 25 words to create word recognition testing material. Words were chosen based on Egan's (1948) six criteria for selecting words for word recognition lists: monosyllabic structured words, equal average of difficulty among lists, equal range of difficulty within lists, equal phonetic composition among lists, composition representative of spoken English (in this case, spoken Arabic), and commonly used words. To fulfill these criterion, Arabic words were selected from three different sources: a word frequency list provided by Landau (1959) which he derived from modern prose and daily newspapers from Lebanon, Egypt, Iraq and Palestine; a list of Arabic key words for learners of everyday Arabic as a second language (Quitregard, 1994); and children's books. Words chosen also represented colloquial and standard Arabic, in that each word had similar forms in colloquial and standard Arabic. Words chosen for this study adhered to the following syllabic structures: CV, CVC, CVCC, CVVC, and CVV (Altoma, 1969). Words chosen followed those phonetic patterns. The initial list of monosyllabic words contained 210 words. Additional words were added to the list from Landua and Quitregard and were rated by a native audiology doctoral student from Palestine. The student ranked each word in order from most familiar to least familiar and eliminated ambiguous words. The highest rated 250 words were selected for recording. Recorded words were found to be familiar and culturally sensitive to the Arabic language.

Talkers. Initial test recordings were made using five male native Arabic-speaking individuals from Jordan. Each participant made 5 minute preliminary recordings of continuous speech for the purpose of further evaluation. Five native Jordanians were

asked to judge the samples of continuous speech using a 10 point scale based on the ease of understanding and the accuracy of the Jordanian dialect. The highest ranked talker became the talker for all subsequent recordings.

Recordings. All recordings were made on the Brigham Young University campus in Provo, Utah, USA in a large anechoic chamber. Ambient noise levels in the anechoic chamber were measured at approximately 0 dB SPL, allowing for a signal-to-noise ratio of approximately 65 dB during recording sessions. A Larson-Davis 2541 microphone was placed approximately 15 cm from the talker at a 0° azimuth and was covered by a 7.62 cm windscreen. The microphone signal was amplified by a Larson-Davis model 900B microphone preamp, coupled to a Larson-Davis model 2200C preamp power supply. The signal was digitized at a 44.1 kHz sampling rate with 24-bit quantization by an Apogee AD-8000 analog-to-digital converter, with preamp gain set to utilize the full range of the 24-bit analog-to-digital converter. The digitized signal was stored on a hard drive for later editing.

The talker was asked to pronounce each monosyllabic word at least four times during all recording sessions. The first and last word of each recording was eliminated to avoid possible list effects. The remaining words were evaluated by a native Arabic-speaking judge (an audiology doctoral student from Palestine) for the best perceived quality of production. The best word was then included on the Arabic speech audiometry test recordings. Any words that were judged to be poorly recorded (peak clipping, extraneous noise, etc.) were rerecorded or eliminated from the study. The final recordings of words were digitally stored as wav files. A total of 250 words were chosen to be included in the word recognition test. Each word was saved as a separate file and digitally

edited to have the same level of root mean square (RMS) as the 1 kHz tone. The 250 words were randomly divided into ten lists of 25 words each.

Calibration

The audiometer was calibrated weekly and prior to each data collection session in accordance to ANSI (2004) standards. No changes in audiometric calibration were necessary during the course of data collection.

Procedures

Custom software was used to control the presentation and randomization of each word file. The software also has the ability to document gathered data in a spreadsheet, including the following: date and time of presentation, participant assigned number, participant gender, test ear, talker gender, intensity level, signal to noise ratio, list name, time of recording per list, wav file, stimulus word, and participant response. The signal was routed from a computer containing the custom software to the external inputs of a Grason Stadler model 1761 audiometer. The stimuli were routed via a single TDH-50P headphone from the audiometer to the subject, who was seated in a double-walled sound suite meeting ANSI S3.1 standards (ANSI, 1999) for maximum permissible ambient noise levels for the ears not covered condition using one-third octave-band measurements. For consistency, each participant's test ear was covered by the right headphone. Before testing each participant, the inputs to the audiometer were calibrated to 0 VU using the 1 kHz testing tone in combination with the custom software.

Participants were not familiarized with the monosyllabic word lists prior to testing. The word lists were presented at ten different levels: -5 to 40 dB HL in 5 dB steps. Across all participants, each list was heard at every presentation level an equal

number of times. Prior to testing, each participant was read the following instructions:

You will hear monosyllabic Arabic words (words with one syllable) at several different loudness levels. At the very soft loudness levels it may be difficult for you to hear the words. Please listen carefully and repeat the words you hear. If you are unsure of the word, you are encouraged to guess. If you have no guess, please be quiet and listen for the next word. Do you have any questions?

All participants spoke English fluently and a native Arabic speaker was available to answer any questions and aid in the interpretation of the words repeated. Words were marked correct if participant's vocal response matched the written form shown on the computer screen through the customized software. Dialectal differences were not considered incorrect as long as the meaning of the word was not affected.

Results

The 250 words were ranked from "difficult" to "easy" to perceive. The most difficult words were eliminated and the easiest 200 words were kept. The 200 words which were easiest to perceive were divided into 4 lists of 50 words each. Each list was counterbalanced using random block assignment. The first 4 words from the rank-ordered list of 200 were randomly assigned to one of 4 lists (list 1, 2, 3, or 4). This process was repeated until each list consisted of 50 words. Table 2 contains the four balanced lists of monosyllabic words ranked from most difficult to easiest.

Once the 4 lists were constructed 8 half-lists of 25 words each were developed. The half-lists were created from each list by assigning the first word in the list as an A or a B, assigning the second word with the opposite letter, and then counterbalancing the designation of the remaining words. Once all the words were assigned a letter the list was

Table 2

Arabic Male Monosyllabic Lists in Rank Order from Most Difficult to Easiest

List 1	List 2	List 3	List 4
عكس	رَفَع	ضَغَط	ضُعْف
شاق	سَطَرَ	حِرْص	سَيَّر
نَعْل	أَصْل	كُون	كَشَف
رَبِح	حَلَف	دور	حَسِب
مُلك	عَبء	عُنف	حُكْم
حقد	حَرَف	عِرْق	ضَبِيق
كَم	رَمَل	بَدْر	أَلْف
حُر	عاج	بَيْع	فُرن
وَضَع	حَطَّ	حِسر	عَد
نَفَس	عَمَق	حَق	نَقَد
جَرَح	رَعَد	نار	حار
صَبَح	موت	حَلَق	جاد
فَرَض	نَقَل	سِيف	خَنَق
بُرْج	تَلَج	غاز	صِين
عَصْر	وَصَف	شَك	خَوْف
عَجَز	خَيْل	رُبْع	خَلَف
جُنْد	خاص	صَلَح	كَنْز
نمر	شَيْبِل	سِجِن	شَاب
طَيْر	رَف	وَعَد	أَرْض
صَوْت	طَعِن	سَطَح	غَيْث
نوع	زَرع	رُوح	رِي
حال	مَنع	شَرَق	سوء
أهل	حَبْر	قُصر	نَهْر
رَبِح	بؤس	بَرَق	وَفَد
زُوج	حَجْم	جَار	فَحَص
			دَخَل
			لَمَس
			حَوَظ
			طُول
			رَمَز
			بُخَل
			أَمْر
			وَقْت
			سَبَع
			سُوق
			حِزْب
			دَم
			مِلْح
			عُود
			بَنَك
			جَهْل
			فَجْر
			شَخَص
			حُزْن
			صَيْف
			عِيد
			صَبْر
			عُرس
			شُكْر
			عِين
			عَلِم
			ضَرَب
			حَزَم
			عَطَف
			بَعَد
			عَزَم
			قَبِر
			حِيس
			رَبَط
			حَلَم
			قُرن
			مَرَج
			صَنِيد
			نُجَل
			ساق
			فَهْر
			هَام
			دَفَع
			فَقَر
			نَخَل
			مَجْد
			شَر
			رِيف
			نَشْر
			شَهْر
			نَحَل
			لَفْظ
			سَهْل
			كُرْه
			حَرْب
			دَمَج
			بَحْر
			حَجَز
			تَبَع
			حَبَس
			دِين
			عَرَض
			فِيل
			عَهْد
			بَحْث
			كَلْب
			قَمَح
			قَلْب
			قَرَض
			كَاس
			نَقَش
			شُوق
			عَرَش
			قُوس
			شِعْر
			شِيبَه
			صَنَعَب
			عَرَف
			عُدْر
			فَهْم
			جَبَل
			حَقَل
			قِسْم
			مِتر
			وَهْم
			طَرَح
			فُصَل
			نَبَض
			سَهْم
			وَجَه
			حَظ
			أَمَس
			فِكْر
			مَدَح
			فَقَه
			صَيْف
			جَهْد
			مِصر
			حَل
			سَنِبَت

divided into 2 half-lists. Half-lists of monosyllabic words are included in Table 3 ranked from most difficult to easiest. All lists are shown in Arabic going from left to right.

Logistic regression was used to calculate regression slopes and intercepts for each of the monosyllabic lists and half-lists. The logistic regression slope and intercept values for each list and half-list are presented in Table 4. These regression slope and intercept values were then inserted into a modified logistic regression equation (Equation 1) that was designed to calculate percentage of correct performance for any given intensity level. Percent correct values were then used to construct psychometric functions.

$$P = \left(1 - \frac{\exp(a + b \times i)}{1 + \exp(a + b \times i)}\right) \times 100 \quad (1)$$

In Equation 1, P is percentage correct recognition, a is the regression intercept, b is the regression slope, and i is the presentation intensity level in dB HL. By inserting the regression slope, regression intercept, and intensity level into Equation 1, it is possible to predict the percentage of correct word recognition for any given intensity level.

Percentage of correct word recognition was predicted for each of the monosyllabic lists and half-lists for a range of -10 to 40 dB HL in 5 dB increments. Psychometric functions were then produced using the predicted percentages. The threshold (presentation intensity required for 50% word recognition performance), the slope at threshold, and the slope from 20 to 80% were calculated for the monosyllabic lists and half-lists by inserting the specified proportions into Equation 2. Data for the threshold, slope at threshold, and slope from 20 to 80% for each list and half-list are also presented in Table 4 along with the regression slope and intercept values.

Table 3

Arabic Male Monosyllabic Half-lists in Rank Order from Most Difficult to Easiest

List 1A	List 1B	List 2A	List 2B	List 3A	List 3B	List 4A	List 4B
حَلْف	رَفَع	أَلْف	حُكْم	عَلِمَ	عَطَفَ	جَبَلَ	فَهِمَ
أَصْل	حَرَفَ	ضَعُفَ	ضَبِيقَ	عَزَمَ	بَعَدَ	عَزَفَ	صَعَبَ
سَطَرَ	عَبَأَ	كَشَفَ	سَيَّرَ	حَزَمَ	ضَرَبَ	عُذِرَ	شَبِهَ
رَمَلَ	عَاجَ	عَدَ	حَسَبَ	حَسَ	قَبِرَ	حَقَلَ	فَصَلَ
خَطَ	نَقَلَ	فُرنَ	جَادَ	صَيَّدَ	رَبَطَ	قَسَمَ	وَهَمَ
تَلَجَ	مَوَتَ	خَنَقَ	حَارَ	فُرنَ	تُبِلَ	طَرَحَ	سَهَمَ
رَعَدَ	عَمَقَ	نَقَدَ	صَيَّنَ	مَرَجَ	حَلَمَ	تَبَيَّنَ	مَتَرَ
خَيَّلَ	وَصَفَ	غَيَّثَ	أَرْضَ	فَقَرَ	سَاقَ	وَجِهَ	مَدَحَ
مَنَعَ	طَعَنَ	رَيَّ	شَابَ	قَهَرَ	نَخَلَ	حَظَ	فَقِهَ
شَيَّلَ	رَفَ	كُنَزَ	خَالَفَ	نَفَعَ	هَامَ	فَكَّرَ	صَيَّفَ
زَرَعَ	خَاصَ	خَوَّفَ	فَحَصَ	مَجَّدَ	خَوَّضَ	مَيَّصَرَ	أَمَسَ
حَجَمَ	حَبَرَ	عَنَفَ	سَوَّءَ	دَخَلَ	رَيَّفَ	سَهَّلَ	سَبَّطَ
رَبَجَ	شَاقَ	وَقَدَ	دَوَّرَ	رَمَزَ	لَمَسَ	حَلَّ	كُرِهَ
مُلِكَ	بَوَّسَ	حَرَصَ	ضَغَطَ	طَوَّلَ	شَرَّ	لَفَظَ	جَهَدَ
نَعَلَ	عَكَسَ	كَوَّنَ	نَهَرَ	شَهَرَ	نَشَرَ	حَجَزَ	نَحَلَ
جُرِحَ	نَفَسَ	حَقَّقَ	حَلَقَ	حَزَبَ	سَوَّقَ	دَمَجَ	حَرَبَ
فَرَضَ	صَبَحَ	بَيَّعَ	جَسَرَ	سَبَّعَ	بُخَلَ	تَبَّعَ	حَبَسَ
حَقَدَ	وَضَعَ	نَارَ	سَيَّفَ	دَمَ	وَقَّتَ	بَحَرَ	عَرَضَ
كَمَ	حَرَّ	بَدَرَ	عَرِقَ	أَمَرَ	عَوَدَ	فَيْلَ	دَيَّنَ
عَصَرَ	عَجَزَ	صَلَحَ	سَطَحَ	جَهَلَ	شَخَّصَ	كَالَبَ	عَهَدَ
صَوَّتَ	طَيَّرَ	وَعَدَ	رُبِعَ	بَنَكَ	صَيَّفَ	قَمَحَ	قَلَبَ
جُنَدَ	بُرَجَ	شَنَّكَ	سَيَّجَنَ	مَلَحَ	فَجَرَ	قَرَضَ	بَحَثَ
حَالَ	نَمَرَ	غَازَ	شَرَّقَ	حَزَنَ	عَيَّدَ	شَوَّقَ	نَفَّشَ
أَهَلَ	نَوَعَ	فَصَّرَ	رَوَّحَ	صَبَّرَ	عَرَّسَ	كَأَسَ	عَرَّشَ
زَوَّجَ	رَبَجَ	بَرَّقَ	جَارَ	شَكَرَ	عَيَّنَ	شَعَرَ	فَوَّسَ

Table 4

Mean Performance of Arabic Male Monosyllabic Lists and Half-lists

List	Intercept ^a	Slope ^b	Slope at 50% ^c	Slope from 20 to 80% ^d	50% Threshold ^e
1	3.3716	-0.1927	4.8	4.2	17.5
2	3.3482	-0.1908	4.8	4.1	17.6
3	3.6010	-0.2058	5.1	4.5	17.5
4	3.2958	-0.1862	4.7	4.0	17.7
<i>M</i>	3.4042	-0.1938	4.8	4.2	17.6
<i>Minimum</i>	3.2958	-0.2058	4.7	4.0	17.5
<i>Maximum</i>	3.6010	-0.1862	5.1	4.5	17.7
<i>Range</i>	0.3053	0.0196	0.5	0.4	0.2
<i>SD</i>	0.1350	0.0084	0.2	0.2	0.1
1A	3.8167	-0.2181	5.5	4.7	17.5
1B	3.0437	-0.1739	4.3	3.8	17.5
2A	3.2279	-0.1855	4.6	4.0	17.4
2B	3.4774	-0.1964	4.9	4.3	17.7
3A	3.5614	-0.2035	5.1	4.4	17.5
3B	3.6417	-0.2081	5.2	4.5	17.5
4A	3.3051	-0.1856	4.6	4.0	17.8
4B	3.2869	-0.1867	4.7	4.0	17.6
<i>M</i>	3.4201	-0.1947	4.9	4.2	17.6
<i>Minimum</i>	3.0437	-0.2181	4.3	3.8	17.4
<i>Maximum</i>	3.8167	-0.1739	5.5	4.7	17.8
<i>Range</i>	0.7730	0.0442	1.1	1.0	0.4
<i>SD</i>	0.2505	0.0145	0.4	0.3	0.1

^aregression intercept. ^bregression slope. ^cPsychometric function slope (%/dB) at 50% was calculated from 49.999 to 50.001%. ^dPsychometric function slope (%/dB) from 20-80%. ^eIntensity required for 50% intelligibility.

$$i = \frac{\log \frac{p}{1-p} - a}{b} \quad (2)$$

In Equation 2, i is the presentation level in dB HL, p is the proportion of correct recognition, a is the regression intercept, and b is the regression slope.

A Chi-Square (χ^2) analysis was done on the completed lists and half-lists. There were no significant differences among the lists; $\chi^2(3, N = 20) = 2.01, p = .57$, nor among the half-lists; $\chi^2(7, N = 20) = 8.56, p = .29$. The psychometric functions for the intensity adjusted lists for the 4 lists and 8 half-lists are shown in Figures 1 and 2 respectively.

Discussion

The purpose of this study was to create a set of homogenous Arabic monosyllabic word lists for use in measuring speech recognition. Figures 1 and 2 show that the lists and half-lists created are homogenous in performance with respect to audibility and psychometric function slopes. A Chi-Square analysis indicated that there were no statistically significant differences among the monosyllabic lists or half-lists for subjects with normal hearing.

As shown in Table 3 and Figures 1 and 2, psychometric function slopes at the 50% location for the monosyllabic lists and half-lists ranged from 4.7 to 5.1 %/dB ($M = 4.8$ %/dB). The psychometric function slopes at the 50% threshold for the monosyllabic lists and half-lists ranged from 4.7 to 5.1 %/dB ($M = 4.8$ %/dB). Slopes of the psychometric functions when measured from 20 to 80% were slightly shallower (4.0 to 4.5 %/dB; $M = 4.2$ %/dB).

When compared to speech recognition materials of other languages, the current Arabic recordings have been found to have similar or slightly lower slopes. Beattie et al.

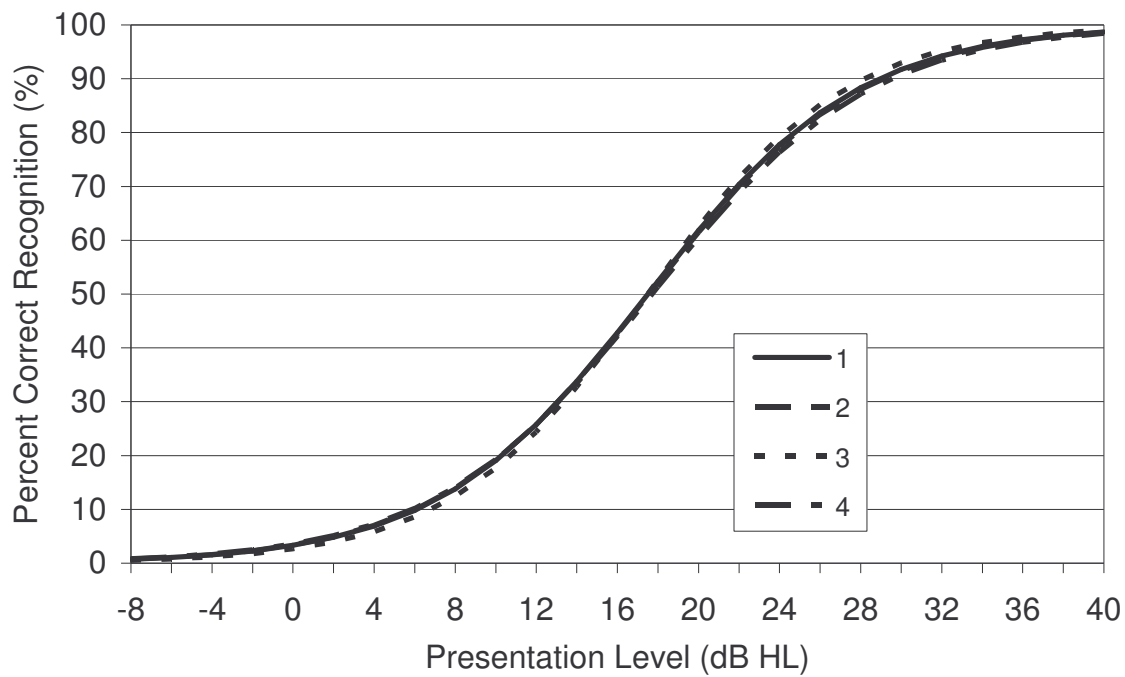


Figure 1. Psychometric functions for the four Arabic monosyllabic lists for male talker recordings.

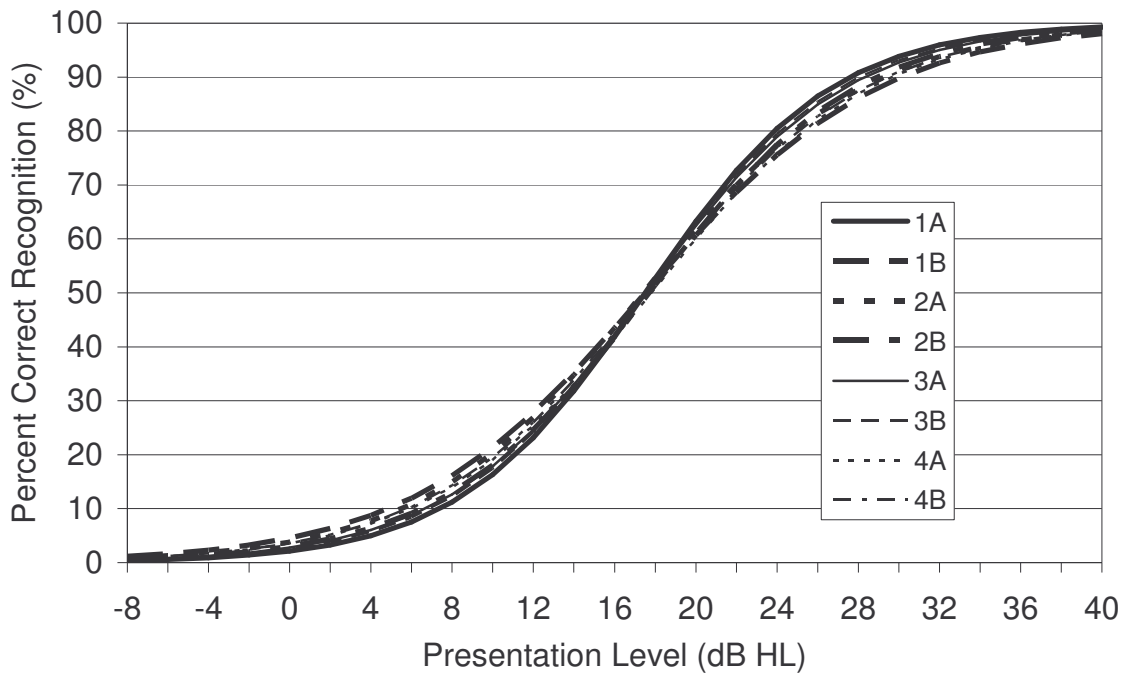


Figure 2. Psychometric functions for the eight Arabic monosyllabic half-lists for male talker recordings.

(1977) reported a mean slope (when measured from the 20 to 80% points) of 4.2 %/dB for the English NU-6 word lists and a mean slope of 4.6 %/dB for the English CID W-22 word lists. In a Mandarin study by Nissen, Harris, Jennings, Eggett, & Buck, (2005), the mean psychometric function slope when measured from the 20 to 80% points was 6.3%/dB for male recordings.

Though when compared to other languages, the recordings of the current Arabic study yielded similar psychometric function slopes, the threshold levels were significantly higher (percent correct at 50 dB HL). As shown in Table 3 the thresholds for monosyllabic lists and half-lists range from 17.5 to 17.7 dB HL ($M = 17.6$ dB HL). In a recent study done evaluating Taiwan Mandarin the psychometric function threshold for all lists was 5.4 dB HL (Dukes, 2006). In speech audiometry materials developed for Japanese, the psychometric function thresholds ranged from 11.5 to 11.7 dB HL ($M = 11.6$ dB HL; Mangum, 2005). As seen in Figures 1 and 2 the psychometric function thresholds yielded from the Arabic monosyllabic word lists are both skewed to the right, indicating higher thresholds. According to the results reported previously, it can be deduced that the current Arabic recordings and possibly the Arabic language in general is difficult to hear due to the phonetic nature of the words used. Though the exact reason for the high thresholds is unclear, it can be concluded that the current Arabic recordings are more difficult to hear than speech audiometry materials created in other languages (Dukes, 2006; Mangum, 2005; Nissen, et al., 2005).

Due to the high thresholds of the monosyllabic Arabic lists and half-lists, the presentation of these materials for the clinical setting must be considered. In the current study, word lists had to be presented at 40 dB HL before all words were heard correctly

by participants. In future clinical use, it is recommended to present lists at 40 dB HL for normal hearing populations.

Though the current study attempted to create psychometrically equivalent monosyllabic words for the Arabic language, there is much more research needed in this area. For instance, though the current lists are appropriate for the Jordanian-Palestinian Arabic dialect, they may not be appropriate for the many other dialects of Arabic. More speech audiometry materials need to be created for the other dialects of Arabic.

In a recent study done by McArdle & Wilson (2006), equivalent speech audiometry lists established by testing young adults with normal hearing were administered to elderly adults with typical presbycusis sensorineural hearing losses. The results showed that though the lists were equivalent for the young adult population, they were drastically different for the elderly group. In order to achieve equality among the lists, 9 of the 18 lists had to be eliminated. From these results, McArdle & Wilson concluded that in order to gain true homogeneity between speech audiometry lists, one needs to create word lists based on the hearing population it will be testing. Due to this report, a future study may include creating Arabic speech audiometry materials for a specific hearing impaired population. The current Arabic recordings are statistically equivalent only for populations with normal hearing. There is a need to evaluate the current materials on hearing-impaired populations.

Lastly, in the current study, it was noted that several times the participants missed the first word on a new list being presented. To eliminate any effects this may have had on the data collected, it is recommended that a non-test word be placed at the beginning of each list, thereby removing the possible first-word-bias.

In summary, the purpose of this study was to develop digitally recorded monosyllabic Arabic speech recognition lists and half-lists of familiar words. These lists and half-lists are homogeneous with respect to audibility and psychometric function slope. These lists can be used to evaluate speech recognition in individuals whose native language is a standard Jordanian-Palestinian dialect of Arabic. The monosyllabic lists and half-lists are contained on the CD entitled *Brigham Young University Arabic Speech Audiometry Materials (Disc 1.0)*. The CD's contents are included in Appendix B.

References

- Al-Masri, M. (2003). *Early detection and habilitation of hearing loss*. MEHA Projects 4. Amman, Jordan.
- Altoma, S. (1969). *The problem of diglossia in Arabic, a comparative study of classical and Iraqi Arabic*. Cambridge, MA: Harvard University Press.
- Alusi, H. A., Hinchcliffe, R., Ingham, B., Knight, J. J., & North C. (1974). Arabic speech audiometry. *Audiology*, 13, 212-230.
- American Academy of Audiology (2004). *Facts about hearing loss*. Retrieved October 21, 2005, from <http://www.audiology.org/professional/gov/hlfacts.pdf>
- American Association of Teachers of Arabic (2003). *All about the Arabic language*. Retrieved February 17, 2006, from <http://www.wm.edu/aata/arabic.php>
- American National Standards Institute (1999). *Maximum permissible ambient noise levels for audiometric test rooms*. ANSI S3.1-1999. New York: ANSI.
- American National Standards Institute (2004). *Specification for audiometers*. ANSI S3.6-2004. New York: ANSI.
- American Speech-Language-Hearing Association (1988). Guidelines for determining threshold level for speech. *ASHA*, 30, 85-89.
- American Speech-Language-Hearing Association (1990). Guidelines for screening for hearing impairments and middle-ear disorders. *ASHA*, 32, 17-24.
- Ashoor, A. A., & Prochazka, T. Jr. (1982). Saudi Arabic audiometry. *Audiology*, 21, 493-508.

- Beattie, R. C., Edgerton, B. J., & Svihovec, D. V. (1977). A comparison of the Auditec of St. Louis cassette recordings of NU-6 and CID W-22 on normal-hearing population. *Journal of Speech and Hearing Disorders, 42*, 60-64.
- Beattie, R. C., Svihovec, D. V., & Edgerton, B. J. (1975). Relative intelligibility of the CID spondees as presented via monitored live voice. *Journal of Speech and Hearing Disorders, 40*, 84-91.
- Berger, K. W. (1971). Speech audiometry. In D. E. Rose (Ed.), *Audiological assessment* (pp. 207-234). Englewood Cliffs, NJ: Prentice-Hall.
- Carhart, R. (1951). Basic principles of speech audiometry. *Acta Oto-Laryngologica, 40*, 62-71.
- Carhart, R. (1965). Problems in the measurement of speech discrimination. *Archives of Otolaryngology, 82*, 253-260.
- Carhart, R., & Porter, L. S. (1971). Audiometric configuration and prediction of threshold spondees. *Journal of Speech and Hearing Research, 14*, 86-95.
- Cramer, K. D., & Erber, N. P. (1974). A spondee recognition test for young hearing-impaired children. *Journal of Speech and Hearing Disorders, 39*, 304-311.
- Davis, H. (1948). The articulation area and the social adequacy index for hearing. *Laryngoscope, 58*, 761-778.
- Doyne, M. P., & Steer, M. D. (1951). Studies in speech reception testing. *Journal of Speech and Hearing Disorders, 16*, 132-139.
- Dukes, A. (2006). Psychometrically equivalent bisyllabic words for word recognition testing in Taiwan Mandarin. Unpublished master's thesis, Brigham Young University, Provo, Utah, USA.

- Egan, J. (1948). Articulation testing methods. *Laryngoscope*, 58, 955-991.
- Elliot, L., & Katz, D. R. (1980). *Northwestern university children's perception of speech (NU-CHIPS)*. Auditec of St. Louis, St. Louis, Missouri.
- Erber, N. P. (1974). Pure-tone thresholds and word recognition abilities of hearing impaired children. *Journal of Speech and Hearing Research*, 17, 194-202.
- Fatihi, A. R. (2001). *Aspects of Arabic phonology*. Delhi: Kalinga Publication.
- Ferguson, C. (1959). Diglossia. *Word*; 15, 325-340.
- Fletcher, H. (1929). *Speech and hearing*. New York: Van Nostrand.
- Fletcher, H. (1950). A method of calculating hearing loss for speech from an audiogram. *Journal of the Acoustical Society of America*, 22, 1-10.
- Fletcher, H. (1953). *Speech and hearing in communication*. New York: Van Nostrand.
- Fry, D. B. (1964). Modifications to speech audiometry. *International Audiology*, 3, 226-236.
- Haskins, H. A. (1949). *A phonetically balanced test of speech discrimination for children*. Master's thesis, Northwestern University, Evanston, IL.
- Hirsh, I. J., Davis, H., Silverman, R., Reynolds, E. G., Eldert, E., & Benson, R. (1952). Development of materials for speech audiometry. *Journal of Speech and Hearing Disorders*, 17, 321-337.
- Hirsh, I. J., Reynolds, E. G., & Joseph, M. (1954). Intelligibility of different speech materials. *The Journal of the Acoustical Society of America*, 26, 530-538.
- Hughson, W., & Thompson, E. A. (1942). Correlation of hearing acuity for speech with discrete frequency audiograms. *Archives of Otolaryngologica*, 36, 526-540.

- Jauhiainen, T. (2001). Hearing impaired children in developing countries. *Scandinavian Audiology*, 30(Suppl. 53), 83-84.
- Kamm, C. A., Carterette, E. C., Morgan, D. E., & Dirks, D. D. (1980). Use of digitized speech materials in audiological research. *Journal of Speech and Hearing Research*, 23, 709-721.
- Knight, J. J., & Littler, T. S. (1953). The technique of speech audiometry and a simple speech audiometer with masking generator for clinical use. *Journal of Laryngology*, 67, 248-265.
- Kruel, E. J., Bell, D. W., & Nixon, J. C. (1969). Factors affecting speech discrimination test difficulty. *Journal of Speech and Hearing Research*, 12, 281-287.
- Landau, J. (1959). *A word count of modern Arabic prose*. New York: American Council of Learned Societies.
- Lyregaard, P. (1997). Towards a theory of speech audiometry tests. In M. Martin (Ed.), *Speech Audiometry 2nd edition* (pp. 34 – 62). San Diego, California: Singular Publishing Group, Inc.
- Mangum, T. C. (2005). *Performance intensity functions for digitally recorded Japanese speech audiometry materials*. Unpublished master's thesis, Brigham Young University, Provo, Utah, USA.
- McArdle, R. A., & Wilson, R. H. (2006). Homogeneity of the 18 QuickSin lists. *Journal of the American Academy of Audiology*, 17, 157-167.
- Nissen, S. L., Harris, R. W., Jennings, L. J., Eggett, D. L., & Buck, H. (2005). Psychometrically equivalent Mandarin bisyllabic speech discrimination material

- spoken by male and female talkers. *International Journal of Audiology*, 44, 379-390.
- Quitregard, D. (1994). Arabic key words, the basic 2000-word vocabulary arranged by frequency in hundred units. New York: The Oleander Press.
- Ridgeway, J. (1986). Compact disks – A revolution in the making. *Canadian Library Journal*, 43, 23-29.
- Roup, C. M., Wiley, T. L., Safady, S. H., & Stoppenbach, D. T. (1998). Tympanometric screening norms for adults. *American Journal of Audiology*, 7, 55-60.
- Smith, A. W. (2001). WHO activities for prevention of deafness and hearing impairment in children. *Scandinavian Audiology*, 30(Suppl. 53), 93-100.
- Sony. (1991). *Compact disc player operating instructions* [Brochure]. Japan: Sony Corporation.
- Tillman, T. W., & Carhart, R. (1966). An expanded test for speech discrimination utilizing CNC monosyllabic words (Northwestern University Auditory Test No.6). Technical Report No. SAM-TR-66-55, USAF School of Aerospace Medicine, Brooks Air Force Base, Texas.
- Wilson, R. H., & Carter, A. S. (2001). Relation between slopes of word recognition psychometric functions and homogeneity of the stimulus materials. *Journal of the American Academy of Audiology*, 12, 7-14.
- Wilson, R. H., Morgan, D. E., & Dirks, D. D. (1973). A proposed SRT procedure and its statistical precedent. *Journal of Speech and Hearing Disorders*, 38, 184-191.

Wilson, R. H., & Oyler, A. L. (1997). Psychometric function for the CID-W22 and NU Auditory Test No.6. materials spoken by the same speaker. *Ear and Hearing* 18, 430-433.

World Health Organization (1998). Report of the International Workshop on Primary Care and Hearing Care. Cape Town, South Africa, March 12-14.

Appendix A

INFORMED CONSENT
RESEARCH PARTICIPATION FORM

Participant: _____ Age: _____

You are asked to participate in a research study sponsored by the Department of Audiology and Speech Language Pathology at Brigham Young University, Provo, Utah. The faculty director of this research is Richard W. Harris, Ph.D. Students in the Audiology and Speech-Language Pathology program may assist in data collection.

This research project is designed to evaluate a word list recorded using improved digital techniques. You will be presented with this list of words at varying levels of intensity. Many will be very soft, but none will be uncomfortably loud to you. You may also be presented with this list of words in the presence of a background noise. The level of this noise will be audible but never uncomfortably loud to you. This testing will require you to listen carefully and repeat what is heard through earphones or loudspeakers. Before listening to the word lists, you will be administered a routine hearing test to determine that your hearing is normal and that you are qualified for this study.

It will take approximately two hours to complete the test. Testing will be broken up into 2 or 3 one hour blocks. Each subject will be required to be present for the entire time, unless prior arrangements are made with the tester. You are free to make inquiries at any time during testing and expect those inquiries to be answered.

As the testing will be carried out in standard clinical conditions, there are no known risks involved. Standard clinical test protocol will be followed to ensure that you will not be exposed to any unduly loud signals.

Names of all subjects will be kept confidential to the investigators involved in the study. Participation in the study is a voluntary service and no payment of monetary reward of any kind is possible or implied.

You are free to withdraw from the study at any time without any penalty, including penalty to future care you may desire to receive from this clinic.

If you have any questions regarding this research project you may contact Dr. Richard W. Harris, 131 TLRB, Brigham Young University, Provo, Utah 84602; phone (801) 422-6460. If you have any questions regarding your rights as a participant in a research project you may contact Dr. Renea Beckstrand, Chair of the Institutional Review Board, 422 SWKT, Brigham Young University, Provo, UT 84602; phone (801) 422-3873, email: renea_beckstrand@byu.edu.

YES: I agree to participate in the Brigham Young University research study mentioned above. I confirm that I have read the preceding information and disclosure. I hereby give my informed consent for participation as described.

Signature of Participant_____
Date_____
Signature of Witness_____
Date

Appendix B

Description of BYU Arabic Speech Audiometry Materials CD

- Track 1 1 kHz calibration tone.
- Track 2 Bisyllabic words for use in measuring the SRT in alphabetical order for familiarization purposes.
- Track 3 Bisyllabic words for use in measuring the SRT in random order, repeated in blocks for a total duration of 5 minutes.
- Track 4 Speech Discrimination List 1 – 50 monosyllabic words in random order.
- Track 5 Speech Discrimination List 2 – 50 monosyllabic words in random order.
- Track 6 Speech Discrimination List 3 – 50 monosyllabic words in random order.
- Track 7 Speech Discrimination List 4 – 50 monosyllabic words in random order.
- Track 8 Speech Discrimination List 1A – 25 monosyllabic words in random order.
- Track 9 Speech Discrimination List 1B – 25 monosyllabic words in random order.
- Track 10 Speech Discrimination List 2A – 25 monosyllabic words in random order.
- Track 11 Speech Discrimination List 2B – 25 monosyllabic words in random order.
- Track 12 Speech Discrimination List 3A – 25 monosyllabic words in random order.
- Track 13 Speech Discrimination List 3B – 25 monosyllabic words in random order.
- Track 14 Speech Discrimination List 4A – 25 monosyllabic words in random order.
- Track 15 Speech Discrimination List 4B – 25 monosyllabic words in random order.
- Tracks 16-24 contain Arabic recordings of routine instructions for various audiometric tests.
- Track 16 Instructions for speech reception threshold-verbal response:
 "الهدف من هذا الفحص هو قياس أقل مستوى من الصوت الذي يمكنك عنده سماع الكلمات و تكرارها. سوف تستمع إلى مجموعة من الكلمات قد تختلف في قوة الصوت. كل مرة تسمع كلمة الرجاء تكرارها. قم بتكرار الكلمات حتى و إن كان الصوت منخفض جدا، و يمكنك أن تحزر. إذا لم تفهم الكلمة و لا تستطيع أن تحزر الرجاء البقاء صامتا وانتظر الكلمة التالية."

- Track 17** Instructions for speech discrimination- verbal response:
"الهدف من هذا الفحص هو قياس مدى قدرتك على فهم وتكرار الكلمات عند تقديمها على مستوى صوت ثابت. الرجاء تكرار الكلمات عند سماعها، إذا لم تكن متأكد يمكنك أن تحزر. إذا لم تفهم الكلمة ولا تستطيع أن تحزر الرجاء البقاء صامتاً و انتظر الكلمة التالية."
- Track 18** Instructions speech audiometry-masking in non-test ear- verbal response:
" خلال هذا الفحص سوف تسمع تشويش في إحدى الأذنين و كلمات في الأذن الأخرى، الرجاء التركيز على الكلمات و تجاهل التشويش. الرجاء تكرار كل كلمة تسمعها، إذا لم تكن متأكد يمكنك أن تحزر. إذا لم تفهم الكلمة ولا تستطيع أن تحزر الرجاء البقاء صامتاً و انتظر الكلمة التالية."
- Track 19** Instructions for speech audiometry – written response:
"الهدف من هذا الفحص هو قياس مدى قدرتك على فهم وتكرار الكلمات عند تقديمها على مستوى صوت ثابت. عند سماعك الكلمة الرجاء كتابتها على الورقة المخصصة للفحص، إذا لم تكن متأكد يمكنك أن تحزر. إذا لم تفهم الكلمة ولا تستطيع أن تحزر الرجاء وضع خط في الفراغ المخصص للكلمة وانتظر الكلمة التالية."
- Track 20** Instructions for speech audiometry-masking in non-test ear – written response:
" خلال هذا الفحص سوف تسمع تشويش في إحدى الأذنين و كلمات في الأذن الأخرى، الرجاء التركيز على الكلمات و تجاهل التشويش. عند سماعك الكلمة الرجاء كتابتها على الورقة المخصصة للفحص، إذا لم تكن متأكد يمكنك أن تحزر. إذا لم تفهم الكلمة ولا تستطيع أن تحزر الرجاء وضع خط في الفراغ المخصص للكلمة و انتظر الكلمة التالية."
- Track 21** Instructions for pure-tone audiometry- hand raising:
" الهدف من هذا الفحص هو قياس أقل مستوى من الصوت الذي يمكنك سماعه. سوف تسمع عدة رنات في إحدى الأذنين ومن ثم في الأذن الأخرى. بعض الرنات سوف تكون سهلة السمع ولكن معظمها سوف تكون خفيفة و صعبة السمع وبعضها تكون في غاية الصعوبة. الرجاء رفع اليد كل مرة تسمع فيها رنة وابقائها مرفوعة طالما تسمع الرنة. عند توقف الرنة أنزل يدك بسرعة. تذكر، ارفع يدك كل مرة تسمع فيها الرنة مهما كان الصوت منخفض."
- Track 22** Instructions for pure-tone audiometry – masking in non-test ear – hand raising:
"خلال هذا الفحص سوف تسمع رنات مع وجود تشويش في الصوت. الرجاء تجاهل التشويش و ركز في الاستماع الى الرنات. التشويش سوف يتغير في الشدة الرجاء الاستماع الى الرنات. الرجاء رفع اليد كل مرة تسمع فيها رنة وابقائها مرفوعة طالما تسمع الرنة. عند توقف الرنة أنزل يدك بسرعة. تذكر، ارفع يدك كل مرة تسمع فيها الرنة مهما كان الصوت منخفض."
- Track 23** Instructions for pure-tone audiometry – button pressing:
" الهدف من هذا الفحص هو قياس أقل مستوى من الصوت الذي يمكنك سماعه. سوف تسمع عدة رنات في إحدى الأذنين ومن ثم في الأذن الأخرى. بعض الرنات سوف تكون سهلة السمع ولكن معظمها سوف تكون خفيفة و صعبة السمع وبعضها تكون في غاية الصعوبة. الرجاء الضغط على الزر كل مرة تسمع فيها رنة والاستمرار في الضغط طالما تسمع الرنة. عند توقف الرنة ارفع اصبعك عن الزر بسرعة. تذكر، اضغط الزر كل مرة تسمع فيها الرنة مهما كان الصوت منخفض."
- Track 24** Instruction for pure-tone audiometry – masking in non-test ear –button pressing:
"خلال هذا الفحص سوف تسمع رنات مع وجود تشويش في الصوت. الرجاء تجاهل التشويش و ركز في الاستماع الى الرنات. التشويش سوف يتغير في الشدة الرجاء الاستماع الى الرنات. الرجاء اضغط على الزر كل مرة تسمع فيها رنة واستمر بالضغط طالما تسمع الرنة. عند توقف الرنة ارفع اصبعك عن الزر بسرعة. تذكر، اضغط على الزر كل مرة تسمع فيها الرنة مهما كان الصوت منخفض."