2012-07-11

Development of Psychometrically Equivalent Speech Audiometry Materials for Measuring Speech Recognition Thresholds in Native Tagalog Speakers

Brandon Brian Taylor
Brigham Young University - Provo

Development of Psychometrically Equivalent Speech Audiometry Materials for Measuring Speech Recognition Thresholds in Native Tagalog Speakers

Brandon B. Taylor

A thesis submitted to the faculty of
Brigham Young University
in partial fulfillment of the requirements for the degree of

Master of Science

Richard Harris, Chair
Shawn L. Nissen
Ron Channell

Department of Communication Disorders

Brigham Young University

August 2012

ABSTRACT

Development of Psychometrically Equivalent Speech Audiometry
Materials for Measuring Speech Recognition
Thresholds in Native Tagalog Speakers

Brandon B. Taylor
Department of Communication Disorders
Master of Science

In addition to the use of pure-tones for testing hearing, speech signals are highly valuable diagnostic tools for identifying and evaluating hearing impairment. Speech audiometry involves the implementation of such signals in the measurement of hearing acuity. One aspect of speech audiometry involves assessment of the speech recognition threshold (SRT) which evaluates an individual's ability to hear and understand speech. While live speech has been used in the past to assess SRT, recorded materials are preferred and have been shown to be advantageous over live speech. High-quality digitally recorded speech audiometry materials have been available in English for some time, but assessment of individuals using speech materials from a language that they do not speak natively has been shown to be both inadequate and inaccurate. Speech audiometry materials have recently become available in many languages. Currently, however, there are no known published recordings for assessment of SRT in the Tagalog language. The goal of this study was to develop psychometrically equivalent speech audiometry materials for measuring speech recognition threshold in Tagalog. During this study Tagalog words were initially recorded by a native speaker selected for accent and vocal quality. The words were reduced down to 90 words to be evaluated in the study. Each of the 90 trisyllabic words were evaluated at 2 dB increments from -10 to 16 dB HL by 20 native Tagalog speakers, all having normal hearing. Based on the results, 34 trisyllabic Tagalog words were selected based on their familiarity to native listeners, relative homogeneity with regards to audibility and psychometric function slope. Each word was then adjusted to make the 50% performance threshold equal to the mean PTA of the 20 research participants (4.3 dB HL). The final edited words were then digitally recorded onto compact disc for distribution and for use in assessing SRT in native Tagalog speakers worldwide.

Keywords: speech reception threshold, SRT, speech audiometry, Tagalog, Filipino, psychometrically equivalent, performance-intensity function, word lists, languages

ACKNOWLEDGMENTS

The completion of this project would not have been possible without the help, sacrifice, and support of many great individuals who have offered much encouragement, as well as their time, and talents towards this project. I am grateful foremost to my Father in heaven who has supported me and my family through this undertaking, and made the completion of this project possible. I am grateful to my wife Tonya and my children for their patience and never-ending love and support, and for the encouragement and confidence of loving family members and friends. I would like to thank my interpreter, Stephanie Surio, and my speaker, Norman Villaroman as well as all of those who participated in the project for their long hours spent in the lab, and for their willingness to help with this project. I would like to thank my committee members, Dr. Richard Harris, Dr. Shawn L. Nissen, and Dr. Ron Channell for their invaluable advice and assistance. I would also like to thank the talented faculty and staff of the department of Communication Disorders who have supported me and offered guidance as well as their knowledge in providing me with such an excellent educational experience. To all those mentioned, as well as any others who have directly or indirectly offered support, assistance and encouragement I offer heartfelt thanks and gratitude.

Table of Contents

List of Tables

List of Figures

Introduction

In the assessment and diagnosis of hearing impairment, a series of tests are used to evaluate an individual's hearing thresholds. The use of these tests to determine the presence and severity of hearing impairment is referred to as audiometry. One of the standard tests in audiometry involves the use of pure-tones to ascertain frequency specific hearing thresholds—the level at which an individual can hear a given tone 50% of the time. In pure-tone audiometry, sinusoidal tones across a range of frequencies and at varying levels of intensity are presented sequentially to the listener. The listener's indication that he or she perceives the tone constitutes the basis for establishing pure-tone thresholds (ASHA, 1978).

Since sinusoidal tones do not adequately represent the acoustic signals used in verbal communication, a second type of test must be used to evaluate the individual's ability to hear and comprehend speech (Egan, 1979). This type of test is referred to as speech audiometry and involves the presentation of speech signals to the listener. One of the specific goals of this type of testing is to determine the individual's speech recognition threshold (SRT). SRT is defined as the lowest intensity at which a listener can correctly repeat spondaic words—bisyllabic words with equal stress on each syllable—50% of the time (Brandy, 2002). Of necessity when performing assessments of the SRT is an appropriately conceived list of spondaic words that are familiar to the listener (Hudgins, Hawkins, Karlin, & Stevens, 1947). Digitally recording word lists for playback to the listener is currently a common practice (ASHA, 1988; Egan, 1979). High-quality digital recordings for assessing the SRT have been available for many years in English. More recently, high-quality digital recordings have become available in a variety of other languages, since the ability to reliably determine the SRT is dependent upon the use of materials in the native language of the listener (Carhart, 1951). While SRT materials are currently available in many languages, high-quality digital recordings for assessing the SRT are not yet available in Tagalog, the major language of the Philippines. Therefore, the objective of this study was to develop, and digitally record, linguistically appropriate materials for the evaluation of SRT in Tagalog.

Review of Literature

Speech Audiometry

While audiologists are able to test hearing using pure-tones, the ability to create and use measures that can quantitatively evaluate hearing thresholds using speech serves as an indicator of how the individual's hearing performs when given a more complex signal, specifically speech (Egan, 1979). Since a common complaint among individuals with hearing impairment is the difficulty in hearing and understanding speech (Martin & Clark, 2009), the use of speech stimuli to assess hearing seems only logical. Speech is a more complex signal than the simple tones used in pure-tone audiometry and understanding speech involves a more complex process than detecting pure-tones. For this reason it is important to use speech stimuli to evaluate hearing thresholds in order to more closely mimic and assess the complex process involved in hearing and comprehending speech (Egan, 1979; Fry, 1964). In addition to evaluating the listener's ability to hear and understand speech, speech audiometry assessments also serve as a verification of an individual's pure-tone thresholds (ASHA, 1988). Speech recognition scores may play an important role as an early indication of a variety of conditions including pseudohypacusis, central auditory disorder, etc. (ASHA, 1988), and in determining site of lesion and development of rehabilitation strategies (Thibodeau, 2007). Fry (1964) framed the importance of speech audiometry by stating that, "there clearly cannot be any substitute for speech audiometry…it is not possible to arrive at a patient's ability to take in speech simply from a pure-tone threshold audiogram" (1964, p. 227).

While the importance of speech audiometry is widely accepted by professionals, the best method and stimuli to be used for testing speech thresholds has not been so readily agreed upon and has undergone many changes in recent history. Before the current practice of assessing speech thresholds using digitally captured and standardized recordings, speech tests were whispered or spoken at a measured distance in order to evaluate the listener's hearing acuity for speech. This type of assessment offered a rough estimation of an individual's hearing ability, but did not yield consistent results nor quantifiable data (ASHA, 1988).

An early recorded version using speech signals to test hearing was developed in 1904 (Bryant). This early test involved the use of a phonograph which was enclosed in a sound-proof box. The sound signals, which were played from recordings made on wax drums, could be varied in their intensity by the test administrator and were directed to the listener's ear via a tube (Bryant, 1904). While this test claimed to overcome many of the shortcomings of using live voice, recorded tests for determining hearing losses for speech were not widely used until the development of the Western Electric 4A (which was later renamed the Western Electric 4C) by the Bell Telephone Laboratories (with the cooperation of the American Federation of Organizations of the Hard of Hearing) in 1926 (Hudgins et al., 1947).

In many of these early tests, items incorporated as stimuli (syllables, words, sentences, etc.) were arbitrarily chosen (Fletcher & Steinberg, 1930). More recently, lists of specific words have been chosen based on specific criteria, carefully arranged, and recorded in order to establish balanced and standardized tools for measuring threshold levels for speech. Several criteria used for selection of appropriate test items have been outlined which include (a) familiarity,

(b) phonetic dissimilarity, (c) normal sampling of English sounds, and (d) homogeneity with respect to basic audibility (Hudgins et al., 1947). In 1988, guidelines for determining the threshold level for speech were released by the American Speech-Language-Hearing Association (ASHA). These guidelines revised similarly aimed guidelines released 9 years earlier. Among the guidelines delineated in this release was the recommendation of the use of spondee words as speech stimulus, the descending method preferred for assessing SRT, and the advantage and preference of digitally recorded materials over live voice and older types of recorded materials such as phonograph records or tape recording. These recommendations, along with other guidelines set out by ASHA in the same release, have become widely accepted standards for measuring SRT and remain relevant for implementation in modern audiological practice.

Necessity of Recorded Materials

Historically, monitored live voice presentation was used to administer items when performing a speech audiometry assessment (ASHA, 1988). This type of testing has obvious limitations, primarily involving consistency. When using monitored live voice, many sources of variability exist that affect the reliability of the results of speech audiometry. Because differences exist among speakers, results may be unreliable due to individual speaker characteristics such as pitch and individual articulation. Additionally, a given speaker’s voice and speech vary from day to day limiting the possibility of consistent test administration over time (Creston, Gillespie, & Krohn, 1966). In order to overcome these disadvantages, recorded materials have been created and their utility is well documented (ASHA, 1988; Brandy, 1966; Thibodeau, 2007). The implementation of recorded materials for assessing hearing thresholds for speech allows the administration of test items to be acoustically consistent, regardless of the test administrator. The

American Speech-Language-Hearing Association (ASHA) recommends the use of recorded materials in speech audiometry:

> Recorded presentation of the test material is the preferred procedure. The use of recorded material standardizes the composition and presentation of the test list. It allows for better control of the intensity of the test items and ensures that the speech pattern of the recorded talker will be consistent to each client. (1988, p. 86)

Although monitored live voice may still be used to administer speech audiometry tests when more preferred methods are not available, recorded materials are greatly preferred (Martin & Clark, 2009).

Due to the advantages of recorded materials, their creation has been attempted for many years including an early set of recorded test materials created in 1904, and a more successful subsequent attempt in 1926 (Hudgins et al., 1947). The development of recorded materials still continues. Today speech audiometry materials are created in a digital format allowing for high fidelity recordings, and computerized manipulation. This type of materials allows for greater consistency in test administration, as well as more reliable results (ASHA, 1988), and is therefore, highly preferred.

Speech Tests and Non-Native English Speakers

The process of selecting appropriate items to be used in any speech audiometry test is of the utmost importance and warrants careful consideration. A set of essential criteria have been outlined for the careful selection of suitable test items for speech audiometry. These criteria include (a) familiarity, (b) phonetic dissimilarity, (c) normal sampling of English speech sounds, and (d) homogeneity with respect to basic audibility (Hudgins et al., 1947). Of these, a normal sampling of English speech sounds is noted to be of lesser importance. Hudgins et al. (1947)

explain that “A normal representation of English speech sounds is of less importance than the other criteria…there is ample evidence that a complete representation of English sounds is not essential in threshold measurement” (p. 59). In light of this, only the other three criteria (familiarity, phonetic dissimilarity, and homogeneity) are discussed further. These criteria are significant when considering the application of speech tests to non-native English speakers, and highlight the insufficiency of using English materials in the assessment of these individuals.

Familiarity. Familiarity refers to the inclusion of words that are relatively easy for the listener to recognize due to previous exposure. This is an essential criterion for selection of test stimuli because it helps ensure test validity (Nissen, Harris, Jennings, Eggett, & Buck, 2005). Since the goal of SRT testing is to measure an individual’s threshold of intelligibility for speech, stimulus words to be used in testing should be as familiar as possible in order to reduce the influence of other factors in the test results, such as limitations in vocabulary knowledge (Hudgins et al., 1947; Ramkissoon, 2001). Additionally, it has been shown that the speed and accuracy with which one is able to recognize spoken words is influenced by various lexical factors including the number and frequency of words that are phonemically similar to the target, the level of similarity between these words and the target, as well as the frequency of occurrence of the target word in the lexicon (Dirks, Takayana, & Moshfegh, 2001; Luce, 1986).

Achieving familiarity is an obvious challenge when using English speech materials to assess individuals who have a native language other than English. According to the most recent census data reported in 2010, more than 55 million American residents reported speaking a language other than English at home (Shin, Kominski, & U.S. Census Bureau., 2010). Of these, more than 24 million reported that they speak English less than “very well” (Shin et al., 2010). When testing individuals who speak English as a second language, especially those who don’t speak English well, achieving the goal of familiarity in the items presented to the listener may not be possible when using English speech audiometry testing materials. One solution to this problem when testing SRT has been to modify the testing procedure. Often this has been accomplished by reducing the number of words to be administered in order to ensure familiarity (Ramkissoon, Proctor, Lansing, & Bilger, 2002). One problem found with this methodology is that when set size is substantially reduced, there is a measured improvement in the SRT that is clinically significant (Punch & Howard, 1985). If a systematic improvement of the SRT results from reducing the set size of the stimuli, then making this type of modification would result in an inaccurate and unreliable measurement of the SRT that overestimates the ability of the listener. Reducing set size is therefore not an acceptable modification for individuals with limited English proficiency.

Another early adaptation made when assessing an individual's SRT involved the use of English digits because it was believed that digits would be more familiar to individuals learning English (Ramkissoon et al., 2002). Several more recent studies support the use of digits in speech reception threshold testing due to their high level of intelligibility and familiarity, as well as a high correlation between the resultant SRT when using digits, and the calculated PTA (pure-tone average) for non-native English speakers (Miller, Heise, & Lichten, 1951; Ramkissoon et al., 2002; Rudmin, 1987). While the use of digits may have advantages over traditional English SRT tests for non-native English speakers, this type of modification may result in reduced sensitivity of the test. Dillon found that "a large number of response foils is required if the test is to be as sensitive as possible" (1983, p. 343). He goes on to explain that tests become easier and therefore lose sensitivity as the number of response foils available to the subject are reduced (Dillon, 1983). Since the use of digits allows for only a small number of available response foils for the subject, sensitivity of the test is reduced making the use of digits less desirable and less useful than spondaic words in testing the SRT.

When considering the importance of familiarity, and given the growing population of non-native English speakers within the United States, the need for speech audiometry materials in the native language of those being tested is unmistakable.

Phonetic dissimilarity. When selecting target stimuli, it is preferred, although not requisite, to include words that represent the natural phoneme distribution of the language (Dirks et al., 2001; Hudgins et al., 1947). When doing so, however, it is essential that target words selected meet the criteria of phonetic dissimilarity, which refers to the goal of selecting stimulus words that are as phonetically unique as possible. This criterion is important to consider because the inclusion of stimuli that vary only minimally would require a more acute discrimination on behalf of the listener, making the test more difficult, without adding to the test's effectiveness as an instrument in determining thresholds for speech (Hudgins et al., 1947). Additionally, the inclusion of highly similar words, such as rhyming words, could offer additional unintended auditory clues to the listener (Ramkissoon, 2001).

Lexical effects such as word frequency and neighborhood density, as discussed in the Neighborhood Activation Model (NAM), have also been shown to affect the ability of a listener to recognize and distinguish spoken words (Luce & Pisoni, 1998). The NAM explains that "words that occur frequently and have few phonemically similar neighbors (lexically 'easy' words) are recognized more accurately than [lexically hard] words that occur less frequently but have a large number of phonemically similar neighbors" (Dirks et al., 2001, p. 233). Since phonetically similar words are more difficult to accurately differentiate, especially for those with hearing impairment who have difficulty with phoneme discrimination (Dirks et al., 2001), it is important that words be included that are phonetically distinct as to not make the test unnecessarily difficult (Bell & Wilson, 2001; Hudgins et al., 1947).

Homogeneity. Achieving homogeneity with respect to basic audibility of the test items is paramount to the proper production of successful speech recognition materials (Hudgins et al., 1947). It is accomplished when all the items within the test are equal in terms of both individual difficulty and overall intelligibility (Dillon, 1983). It has been found that homogeneity of audibility of the individual test items results in increased sensitivity of the test as a whole, and that homogeneity must be achieved if the test is to be maximally sensitive (Dillon, 1983). Additionally, fewer test items are necessary to establish SRT when homogeneity is achieved (Young, Dudley, & Gunter, 1982). Another substantial benefit of homogeneity among test items is evident in the increase in steepness of slope of their psychometric functions, which is directly representative of the homogeneity of the test items and indicative of increased precision in determining threshold (Hirsh et al., 1952; Wilson & Carter, 2001; Young et al., 1982). Since the steepness of the psychometric function is directly related to the homogeneity of the test items, manipulation of the recorded words to increase the steepness of the psychometric function of the words can simultaneously increase the homogeneity of the individual test items and of the test as a whole (Wilson & Carter, 2001). While complete homogeneity is never perfectly accomplished, it should be strived for as much as possible when constructing lists for speech recognition testing (Dillon, 1983).

Given the aforementioned criteria, it is clear that speech audiometry materials used in testing must correspond with the language of the subject if the test is to be both accurate and reliable. Carhart expressed this same conclusion saying, “It is apparent that the techniques of speech audiometry must be developed independently for every language in which they are used”

(1951, p. 62) . Recently, attempts have been made to record speech audiometry materials in a variety of languages including Spanish (Christensen, 1995), Italian (Harris & Greer, 1997), Polish (Harris, Nielson, McPherson, & Skarzynski, 2004), Russian (Harris et al., 2007), Mandarin Chinese (Nissen, Harris, & Slade, 2007), Danish (Olsen, 1996), Brazilian Portuguese (Harris, Goffi, Pedalini, Gygi, & Merrill, 2001), Japanese (Mangum, 2005), French (Nelson, 2004), Greek (Iliadou, Fourakis, Vakalos, Hawks, & Kaprinis, 2006), Korean, and Cantonese (Lau & So, 1988). While these efforts constitute a considerable effort and advancement in the availability of speech audiometry materials, there remains a need for continued development of comparable materials in other languages.

Tagalog Language

Tagalog is one of the prominent languages of the Philippines, a nation estimated to have a population of more than 103 million people ("The World Factbook," 2012). Tagalog, declared the nation’s official language in 1937, currently forms the basis for Filipino, one of the country’s two official languages, along with English (Ramos, 2005; "The World Factbook," 2012). The term Filipino has been used since the 1980s to refer to the country’s national language, but is often referred to as Tagalog by those outside the Philippines. While Tagalog and Filipino are closely related, and the terms are sometimes used interchangeably by outsiders, they are not synonymous in that Filipino more readily implements words borrowed from both English and Spanish (Ledesma, 1974; Ramos, 2005). Tagalog is spoken natively by approximately fifteen million people in the Philippines, but is understood and utilized as the country’s official language by approximately 75% of the Philippine population (Ramos, 2005; Rubino & Llenado, 2002). Tagalog is an Austronesian language of the Philippine type and has served as the basis of the national language of the Philippines since the 1930s (Ramos, 2005; Rubino & Llenado, 2002).

The use of Tagalog in the media, along with its considerable role in education, has served to bolster Tagalog as the country's official language (Rubino & Llenado, 2002).

The Austronesian family of languages, of which Tagalog is a member, is believed to be the second largest language family in the world and includes anywhere from 800 to over 1200 languages (Erickson, 2000). In addition to being large in number, the Austronesian language family is also vast in region, spanning from Southeast Asia to the Pacific's westernmost islands (Blaylock, 1999).

Much like English, Tagalog pronunciation varies from place to place and across social and educational classes. These variations form a variety of dialects found throughout the Philippines.

> Among the distinctive regional dialects may be noted at least the following six: Bataan, Batangas, Bulacan, Manila, Tanay-Paete, and Tayabas. The dialect of Bulacan is sometimes considered to be the "purest", showing perhaps fewer signs than the others of the influence of Spanish, English, or neighboring Philippine languages. But it is the dialect of Manila that is now generally regarded as standard, in both pronunciation and grammar, by virtue both of its prestige as the dialect of the capital and of its overwhelming numerical superiority. (Schachter & Otanes, 1972, p. 1)

Behind the arrival of the Spanish to the Philippines, the ancient Indic syllabary previously used was replaced with Latin script based on Spanish orthography (Rubino & Llenado, 2002). Tagalog has five vowels which are phonetically similar to those of Spanish or Italian, and are commonly arranged into a basic syllable shape of the CV(C) type. Words in Tagalog, much like Spanish, generally carry stress on the penultimate syllable, unless indicated elsewhere by an accent mark. Along with the generic accent mark to indicate stress (á), Tagalog orthography also contains two other types of accent marks. The other two types, grave (à) and circumflex (â) differentiate between penultimate stress and final stress respectively, in words ending in a glottal stop. Tagalog also exhibits antepenultimate accompanying stress in some words, which bears stress on the final or penultimate syllable, but also has an additional prominent syllable discernible by vowel length and pitch prominence (Rubino & Llenado, 2002).

While Tagalog has been long established as a prominent language in the Philippines, it is also becoming increasingly more prevalent in the United States. According to the 2010 census, Tagalog ranked as the 4th most frequently spoken language at home in the United states with over 1.4 million speakers, and the 2nd most commonly spoken Asian language, behind Chinese (Shin et al., 2010). Although Filipino is regarded as one of the national languages of the Philippines, the term Tagalog will be used for this study due to its well-rooted status as a basis for Filipino, its wide use and greater familiarity in the United States, and since words borrowed from English and Spanish into the Filipino language are less desirable for this project.

Tagalog Speech Audiometry Materials

The need for speech audiometry services for native speakers of Tagalog is growing as this population expands both here and abroad. Currently there are no known recorded speech audiometry materials available in Tagalog. A previous study was conducted at Kent State University in 1974 in an effort to construct speech audiometry word lists in Pilipino—a term sometimes used interchangeably with Tagalog, but not synonymous (Ledesma, 1974). While recordings were used in that study, those recordings were not made available, nor were they standardized. Since then, the advancement of technology as well as research in the field of audiology provides the means necessary to create digitally recorded, and standardized materials that can more accurately and efficiently be utilized to assess thresholds for speech. In addition, these materials, including the digital recordings, can be made readily available worldwide. Aside from the study previously mentioned, no attempts to create Tagalog materials for speech audiometry have been found. The need for these types of materials is highlighted by the fact that an estimated 275 million people worldwide have moderate to profound hearing loss in both ears, 80% of which live in low- and middle-income countries such as the Philippines (World Health Organization, 2012). With an ever growing population of Tagalog speakers both in the Philippines and here in the United States, and the growing pervasiveness of hearing loss in underdeveloped countries such as the Philippines, the need for speech audiometry materials in Tagalog has never been greater. The goal of this project was to create such speech audiometry materials in Tagalog.

Method

Participants

Selection and screening. A total of 20 native Tagalog speakers participated in the evaluation of materials developed in this study (9 male and 11 female). The participants' ages ranged from 19 to 51 years (M = 27). They had lived in the United States from 3 to 32 years with an average U.S. residency of 9 years. Each participant underwent a hearing screening and each displayed pure-tone air-conduction thresholds of <15 dB HL at octave and mid-octave frequencies from 125-8000 Hz. Each participant also displayed static acoustic admittance between 0.3 and 1.4 mmhos with peak pressure between -100 and +50 daPa (ASHA, 1990; Roup, Wiley, Safady, & Stoppenbach, 1998). All participants possessed an ipsilateral acoustic reflex of 95 dB HL or better at 1000 Hz. Additionally, each participant signed a document indicating informed consent which had been approved by the Brigham Young University

Institutional Review Board for human subjects. The mean pure-tone average (PTA) for all participants was 4.3 dB HL. Statistical data for participant thresholds are displayed in Table 1.

Materials

Word selection. A preliminary corpus consisting of 250 trisyllabic words was drawn from an inventory of high frequency Tagalog words harvested from the internet (K. Scannell, personal communication, February 3, 2011). Each word was evaluated by three native judges who evaluated the words based on familiarity and appropriateness. From the initial corpus of 250 words, 160 of the words were excluded because (a) words were deemed unfamiliar by the native judges, (b) words were deemed as inappropriate or culturally insensitive, and (c) words were homophonous.

Speakers. Initial test recordings were created from six male native Tagalog speakers. Each speaker was native to the Philippines and self-reported speaking Tagalog daily. Following creation of the initial recordings, a judging panel consisting of seven native Tagalog speakers evaluated the speech of each recorded talker. The judges were asked to rank order all six speakers based on overall vocal quality, standard dialect, pleasantness of voice, intelligibility, and pronunciation. The speaker who received the highest overall ranking was selected as the speaker for all subsequent recordings.

Recordings. All recordings were made in a double walled sound booth located on the campus of Brigham Young University, Provo Utah, U.S.A. The sound booth maintains a signal-to-noise ratio of approximately 65 dB with the sound floor measuring 0 dB SPL. A Larson-Davis model 2541, 1.27 cm microphone covered by a 7.62 cm windscreen was used for all recordings and was positioned at 0° azimuth at a distance of approximately 15 cm from each speaker. The

Table 1

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

Frequency (Hz)	*M*	*Minimum*	*Maximum*	*SD*
125	4.3	0	15	4.1
250	2.3	-10	10	5.3
500	4.3	-5	15	6.3
750	5.3	-5	15	5.3
1000	4.0	-10	15	5.5
1500	5.0	-5	10	4.3
2000	4.5	-10	10	4.8
3000	3.3	-5	15	5.9
4000	2.3	-10	15	6.6
6000	-1.8	-10	15	6.1
8000	1.5	-10	15	8.9
PTA[a]	4.3	-8.3	10.0	4.6

[a]PTA = arithmetic average of thresholds at 500, 1000, and 2000 Hz.

microphone signal was amplified using a Larson-Davis model PRM902 microphone preamp along with a Larson-Davis model 2221 microphone preamplifier power supply. The signal was
digitized using a Benchmark ADC1 analog-to-digital converter and was stored on a hard drive for later editing. A 44.1 kHz sampling rate with 24-bit quantization was used for all recordings, and every effort was made to utilize the full range of the 24-bit analog-to-digital converter.

During the recording sessions, the speaker was instructed to use normal vocal effort and to pronounce each word to the best of his ability at least four times, with intermitting short pauses. To avoid possible list effects, the first and last repetitions of each word were excluded from selection. A native judge then rated the remaining repetitions of each word for perceived quality of production. The best production of each word was selected for inclusion in the evaluation portion of the study. Any words that were judged to be poorly recorded were rerecorded or eliminated from the study prior to listener evaluation. After the rating process, the intensity of each word to be included in the listener evaluation was edited as a single utterance using Sadie Disk Editor software to yield the same level equivalent as that of a 1 kHz calibration tone. Each word was saved as a 24-bit *wav* file.

Procedures

Custom software was used to control the randomization, presentation, and scoring of the trisyllabic words used and evaluated in this study. This software was also used to record the performance data. The signal was routed from a computer hard drive to the external input of a Grason Stadler model 1761 (GSI-61) audiometer. The stimuli were then presented via TDH-50P headphones from the audiometer to the participant, who was seated in a double-walled sound booth which conforms with ANSI S3.1 standards for maximum permissible ambient noise levels for the ears not covered condition using one-third octave-band measurements (American National Standards Institute, 1999). Prior to testing each participant, a 1 kHz calibration tone was used to calibrate the inputs to the audiometer to 0 VU. Additionally, the audiometer was calibrated weekly during, and at the conclusion of data collection. Calibration was performed in accordance with ANSI S3.6 standards (American National Standards Institute, 2004). No changes in calibration were necessary throughout the course of data collection.

Each participant attended one test session and was allowed to have several rest periods during the test session. After passing a hearing screening exam, each participant was familiarized with the list of 90 trisyllabic words at a comfortable listening level of 50 dB HL. The list of words was played in alphabetical order as each participant listened and read along using a provided copy of the word list. The following instructions were read in English to each participant prior to familiarization of the test stimuli:

> You will now hear the list of words we will use in this part of the research study. These words will be presented at a comfortable listening level. Please read the list of words silently as you hear them to make sure you are familiar with all the words. Do you have any questions?

Following familiarization, the copy of the word list was taken from the subject, and a randomized list made up of the recorded 90 trisyllabic words was presented to each listener at an initial sound level 6 dB below the listener's PTA. Correct repetition of any words on the list during the initial presentation was followed by a 2 dB decrement and subsequent presentation of the entire list. The list was played again at decrements of 2 dB until the listener repeated none correctly, or until the list had been played at -10 dB HL. After the lowest intensity had been played, the list of words was randomized and was presented at 2 dB above the initial intensity.

Randomized lists were presented at increasing intensities of 2 dB until the listener repeated all 90 words correctly, or until the list was presented at 16 dB HL. Each participant listened to the trisyllabic word recordings made by the speaker in a randomly determined sequence. Participants repeated words verbally, which were scored as being correct or incorrect by a judge who spoke Tagalog natively. The potential interference of learning effects was reduced by (a) the relatively large number of words evaluated by listeners, (b) the presentation of the stimuli from low to high intensity, and (c) the randomization of presented stimulus at each intensity level. Prior to the evaluation of the trisyllabic words, and following familiarization of the word list in alphabetical order, instructions were given to the participants in English as follows:

> You will hear Tagalog words at a number of different loudness levels. Each word is three syllables in length. At the very soft loudness levels, it may be difficult for you to hear the words. For each word, listen carefully to the word, and then repeat what you think the word was. If you are not sure, you may guess. If you have no guess simply say, “I don’t know,” or wait silently for the next word. Do you have any questions?

Results

Upon completion of data collection, regression slope and intercept were calculated for each of the 90 trisyllabic words evaluated in the study. These values were then inserted into a modified logistic regression equation that was designed to calculate the percent correct at each intensity level. The original logistic regression equation is as follows:

$$log\frac{p}{1-p} = a + b \times i \tag{1}$$

In Equation 1, p is the proportion correct at any given intensity level, a is the regression intercept, b is the regression slope, and i is the intensity level in dB HL. When Equation 1 is solved for p and multiplied by 100, Equation 2 can then be calculated where P is percent correct recognition:

$$P = \left(1 - \frac{\exp(a + b\ \times i)}{1 + \exp(a + b \times i)}\right) * 100 \tag{2}$$

In Equation 2, P is percentage of correct recognition, a is the regression intercept, b is the regression slope, and i is the presentation intensity in dB HL. By inserting the regression slope, regression intercept, and intensity level into Equation 2, it is possible to predict the percentage of correct recognition at any specified intensity level. Percentage of correct recognition was calculated for each of the trisyllabic words for a range of –10 to 16 dB HL in 1 dB increments.

In order to calculate the intensity level required for a given proportion, Equation 1 was solved for i (see Equation 3). By inserting the desired proportions into Equation 3, it is possible to calculate the threshold (the intensity required for 50% intelligibility), the slope (%/dB) at threshold, and the slope from 20 to 80% for each psychometric function. When solving for the threshold ($p = 0.5$), Equation 3 can be simplified to Equation 4.

$$i = \frac{log\frac{p}{1-p} - a}{b} \tag{3}$$

$$i = \frac{-a}{b} \tag{4}$$

Logistic regression slopes and intercepts were used to calculate threshold (intensity required for 50% correct perception), slope at 50%, and slope from 20 to 80% for each trisyllabic word. Words selected for inclusion in the final list of trisyllabic words were those that exhibited the steepest psychometric performance-intensity functions.

Thresholds of the 90 trisyllabic words ranged from 2.1 dB HL to 17.4 dB HL ($M = 9.5$). Equation 2 was used to calculate the psychometric performance-intensity functions for each trisyllabic word using the logistic regression intercept and slope values. The slopes at 50% ranged from 4.4%/dB to 18.1%/dB ($M = 9.9$). The slopes from 20 to 80% ranged from 3.8%/dB to 15.7%/dB ($M = 8.5$). The slopes at 50% threshold were steeper than the slopes at 20 to 80%. Slopes of the psychometric performance-intensity functions and 50% thresholds for all trisyllabic words are presented in Table 2.

As explained earlier, previous research has demonstrated that test time is reduced, and sensitivity and reliability improve when SRT test stimuli are relatively homogeneous with steep psychometric performance-intensity function slopes (Dillon, 1983; Wilson & Strouse, 1999). Therefore, only the words with the steepest psychometric performance-intensity function slopes of $\geq$ 7.0%/dB which had enough available headroom for adjustment were included in the final list of trisyllabic words, resulting in a total of 34 selected words. The slope of the psychometric performance-intensity functions for the complete list of 90 words (Figure 1) shows less variability when compared to the selected words (Figure 2). The intensity of each of these final 34 words was digitally adjusted to align the 50% threshold of each word with the mean PTA of the subjects (4.3 dB HL). The threshold, slope at threshold, and the slope from 20 to 80% for the 34 selected trisyllabic words are listed in Table 3. The psychometric performance-intensity functions for the selected words following digital adjustment are shown in of Figure 3. The

Table 2

Mean Performance for 90 Tagalog Male Trisyllabic SRT Words

#	Word	a[a]	b[b]	Slope at 50%[c]	Slope 20-80%[d]	Threshold[e]	ΔdB[f]
1	alagád	2.96855	-0.38697	9.7	8.4	7.7	3.4
2	asáwa	2.28870	-0.35365	8.8	7.7	6.5	2.2
3	babáe	0.69105	-0.25109	6.3	5.4	2.8	-1.5
4	binigyán	3.03776	-0.32145	8.0	7.0	9.5	5.2
5	bumalík	5.79879	-0.50876	12.7	11.0	11.4	7.1
6	dahilán	6.73864	-0.65326	16.3	14.1	10.3	6.1
7	daigdíg	3.21111	-0.34502	8.6	7.5	9.3	5.1
8	dalawá	1.84909	-0.44378	11.1	9.6	4.2	-0.1
9	damdámin	2.69946	-0.29488	7.4	6.4	9.2	4.9
10	dárating	3.15180	-0.31809	8.0	6.9	9.9	5.7
11	dumatíng	2.00441	-0.38598	9.6	8.4	5.2	0.9
12	gágawin*	7.98022	-0.46827	11.7	10.1	17.0	12.8
13	gamítin	4.64124	-0.54598	13.6	11.8	8.5	4.3
14	ganitó	4.71224	-0.39669	9.9	8.6	11.9	7.6
15	ginámit	3.76009	-0.39685	9.9	8.6	9.5	5.2
16	gumawâ	5.12874	-0.36863	9.2	8.0	13.9	9.7
17	harapán	6.43498	-0.48850	12.2	10.6	13.2	8.9
18	ilalim	2.85203	-0.30450	7.6	6.6	9.4	5.1
19	itaás	9.01993	-0.72485	18.1	15.7	12.4	8.2
20	kaáway	2.38866	-0.23056	5.8	5.0	10.4	6.1
21	kabílang	0.92259	-0.39576	9.9	8.6	2.3	-1.9
22	kagáya	4.05525	-0.42893	10.7	9.3	9.5	5.2
23	kanilá	3.52087	-0.52071	13.0	11.3	6.8	2.5
24	kapatíd	4.41559	-0.38150	9.5	8.3	11.6	7.3
25	kasama	1.18213	-0.55293	13.8	12.0	2.1	-2.1
26	katulad	3.34311	-0.43551	10.9	9.4	7.7	3.4
27	kaugnáy	3.35660	-0.32046	8.0	6.9	10.5	6.2
28	kilala	3.45323	-0.44977	11.2	9.7	7.7	3.4
29	lalake	4.32851	-0.39800	10.0	8.6	10.9	6.6
30	larawan	0.77660	-0.25274	6.3	5.5	3.1	-1.2
31	lipunan	1.56200	-0.30311	7.6	6.6	5.2	0.9
32	lumabás	4.30316	-0.25112	6.3	5.4	17.1	12.9
33	lumapit	4.27654	-0.31157	7.8	6.7	13.7	9.5
34	mabilís	4.16601	-0.43111	10.8	9.3	9.7	5.4
35	mabuti	2.98497	-0.44302	11.1	9.6	6.7	2.5
36	madalás	3.38446	-0.37801	9.5	8.2	9.0	4.7
37	magulang	3.47168	-0.37416	9.4	8.1	9.3	5.0
38	mahirap	3.52648	-0.47842	12.0	10.4	7.4	3.1
39	makita	2.32147	-0.22302	5.6	4.8	10.4	6.2
40	malakás	4.63459	-0.40610	10.2	8.8	11.4	7.2
41	maliban	2.64602	-0.27356	6.8	5.9	9.7	5.4
42	malinaw	4.79825	-0.48225	12.1	10.4	9.9	5.7

#	Word	a[a]	b[b]	Slope at 50%[c]	Slope 20-80%[d]	Threshold[e]	ΔdB[f]
43	mamatáy	1.68740	-0.22307	5.6	4.8	7.6	3.3
44	marami	3.22045	-0.29887	7.5	6.5	10.8	6.5
45	masamâ	4.88982	-0.47141	11.8	10.2	10.4	6.1
46	mataás	1.98749	-0.28489	7.1	6.2	7.0	2.7
47	matagál	2.89372	-0.38255	9.6	8.3	7.6	3.3
48	mayroón	3.61132	-0.45034	11.3	9.7	8.0	3.8
49	nabanggít	3.50326	-0.42190	10.5	9.1	8.3	4.1
50	nagawâ	3.77278	-0.51862	13.0	11.2	7.3	3.0
51	nagsabi	6.65218	-0.60835	15.2	13.2	10.9	6.7
52	namatáy	3.90881	-0.35686	8.9	7.7	11.0	6.7
53	nangyari	3.11965	-0.29205	7.3	6.3	10.7	6.4
54	paano	3.08071	-0.37546	9.4	8.1	8.2	4.0
55	pag-asa	2.30110	-0.38627	9.7	8.4	6.0	1.7
56	pagdatíng	4.19413	-0.47027	11.8	10.2	8.9	4.7
57	paggamit	4.45633	-0.35767	8.9	7.7	12.5	8.2
58	pag-ibig	2.54924	-0.37425	9.4	8.1	6.8	2.6
59	pagkain	2.76666	-0.17554	4.4	3.8	15.8	11.5
60	pagsambá	3.87276	-0.37362	9.3	8.1	10.4	6.1
61	pahayág	2.63785	-0.34895	8.7	7.6	7.6	3.3
62	panahón	2.20029	-0.34059	8.5	7.4	6.5	2.2
63	pangakò	1.34057	-0.29528	7.4	6.4	4.5	0.3
64	pangalan	3.21255	-0.32072	8.0	6.9	10.0	5.8
65	paraán	3.04219	-0.38586	9.6	8.4	7.9	3.6
66	patuloy	6.86187	-0.39418	9.9	8.5	17.4	13.2
67	pintuan	3.09946	-0.35867	9.0	7.8	8.6	4.4
68	pumasok	4.32301	-0.43324	10.8	9.4	10.0	5.7
69	pumuntá	3.32535	-0.23964	6.0	5.2	13.9	9.6
70	sabihin	5.94618	-0.47502	11.9	10.3	12.5	8.3
71	salamat	7.31359	-0.54471	13.6	11.8	13.4	9.2
72	samahán	4.56068	-0.46783	11.7	10.1	9.7	5.5
73	sapagká't	2.08263	-0.38488	9.6	8.3	5.4	1.2
74	sarili	4.97367	-0.46515	11.6	10.1	10.7	6.4
75	simbahan	2.77288	-0.35164	8.8	7.6	7.9	3.6
76	simulâ	5.14056	-0.44866	11.2	9.7	11.5	7.2
77	sinabi	5.22840	-0.37689	9.4	8.2	13.9	9.6
78	subali't	2.38551	-0.28127	7.0	6.1	8.5	4.2
79	sumunód	7.42893	-0.52942	13.2	11.5	14.0	9.8
80	sundalo	4.22374	-0.45746	11.4	9.9	9.2	5.0
81	súsunod	2.96473	-0.35643	8.9	7.7	8.3	4.1
82	tagumpáy	3.19150	-0.41587	10.4	9.0	7.7	3.4
83	tahánan	2.36451	-0.37079	9.3	8.0	6.4	2.1
84	tinapay	3.83928	-0.49973	12.5	10.8	7.7	3.4
85	tinawag	3.43679	-0.44168	11.0	9.6	7.8	3.5
86	totoó	4.92485	-0.33169	8.3	7.2	14.8	10.6
87	tumanggáp	7.63608	-0.57512	14.4	12.4	13.3	9.0
88	tungkulin	6.89685	-0.47906	12.0	10.4	14.4	10.1

#	Word	a[a]	b[b]	Slope at 50% [c]	Slope 20-80%[d]	Threshold[e]	ΔdB[f]
89	umaga	3.89301	-0.37992	9.5	8.2	10.2	6.0
90	umalís	4.29152	-0.30167	7.5	6.5	14.2	10.0
	M	3.78660	-0.39504	9.9	8.5	9.5	5.3
	Min	0.69105	-0.72485	4.4	3.8	2.1	-2.1
	Max	9.01993	-0.17554	18.1	15.7	17.4	13.2
	Range	8.32888	0.54931	13.7	11.9	15.3	15.3
	SD	1.66727	0.09870	2.5	2.1	3.2	3.2

[a]*a* = regression intercept. [b]*b* = regression slope. [c]Psychometric function slope (%/dB) at 50% was calculated from 49.999 to 50.001%. [d]Psychometric function slope (%/dB) from 20-80%. [e]Intensity required for 50% intelligibility. [f]Change in intensity required to adjust the threshold of a word to the mean PTA of the subjects (4.25 dB HL).

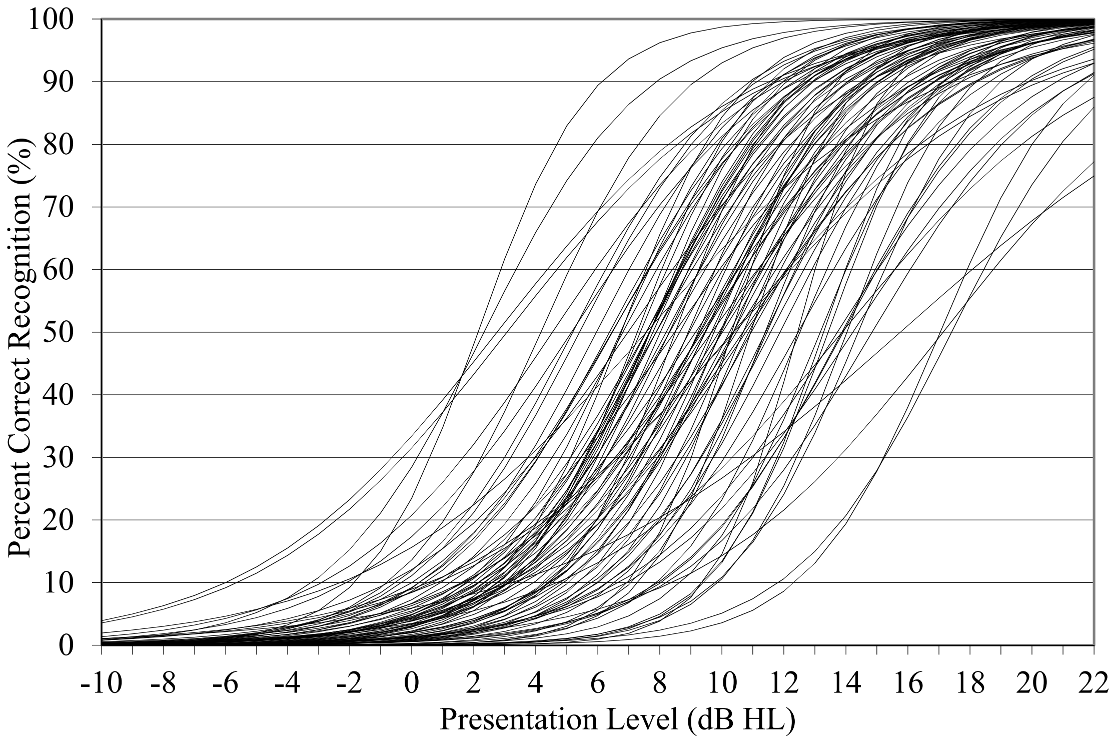

Figure 1

Predicted psychometric performance intensity functions for all 90 unadjusted trisyllabic Tagalog word recordings.

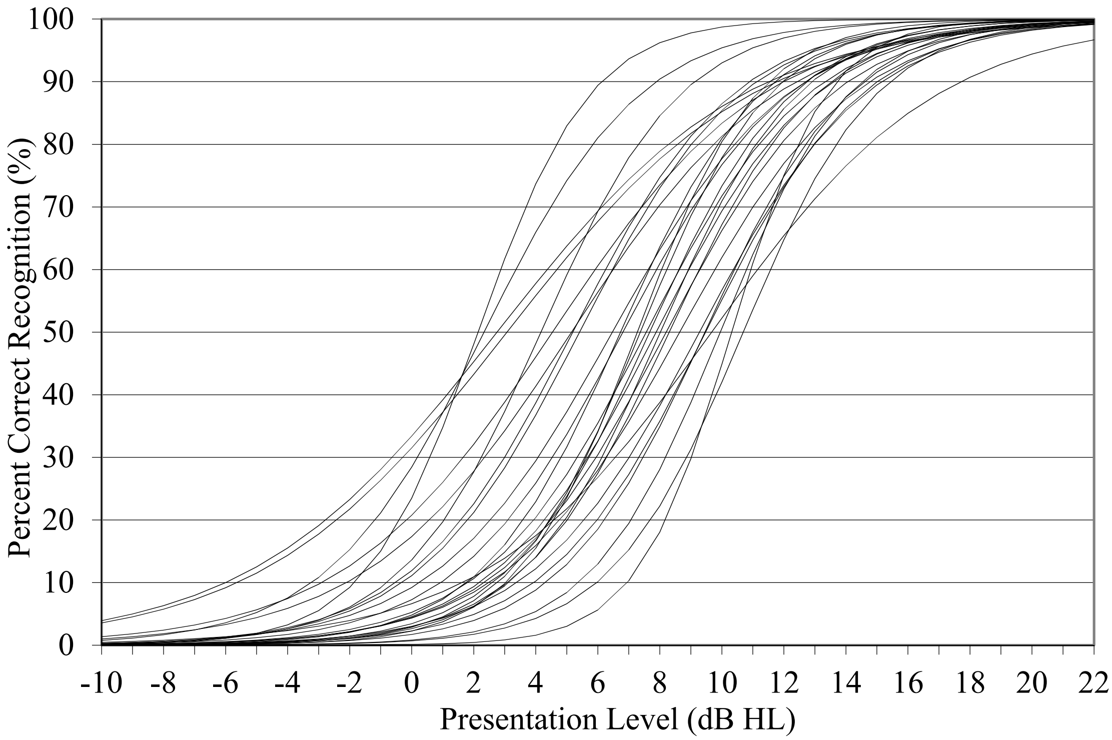

Figure 2

Predicted psychometric performance intensity functions for 34 selected trisyllabic Tagalog words before adjustment.

Table 3

Mean Performance for 34 Selected Tagalog Male Trisyllabic SRT Words

#	Word	a[a]	b[b]	Slope at 50%[c]	Slope 20-80%[d]	Threshold[e]	ΔdB[f]
1	asáwa	2.28870	-0.35365	8.8	7.7	6.5	2.2
2	babáe	0.69105	-0.25109	6.3	5.4	2.8	-1.5
3	bumalík	6.73864	-0.65326	16.3	14.1	10.3	6.1
4	daigdíg	1.84909	-0.44378	11.1	9.6	4.2	-0.1
5	dárating	2.00441	-0.38598	9.6	8.4	5.2	0.9
6	ganitó	3.76009	-0.39685	9.9	8.6	9.5	5.2
7	kaáway	0.92259	-0.39576	9.9	8.6	2.3	-1.9
8	kabílang	4.05525	-0.42893	10.7	9.3	9.5	5.2
9	kapatíd	1.18213	-0.55293	13.8	12.0	2.1	-2.1
10	kasama	3.34311	-0.43551	10.9	9.4	7.7	3.4
11	lalake	0.77660	-0.25274	6.3	5.5	3.1	-1.2
12	larawan	1.56200	-0.30311	7.6	6.6	5.2	0.9
13	mabuti	2.98497	-0.44302	11.1	9.6	6.7	2.5
14	magulang	3.47168	-0.37416	9.4	8.1	9.3	5.0
15	mahirap	3.52648	-0.47842	12.0	10.4	7.4	3.1
16	maliban	2.64602	-0.27356	6.8	5.9	9.7	5.4
17	malinaw	4.79825	-0.48225	12.1	10.4	9.9	5.7
18	matagál	2.89372	-0.38255	9.6	8.3	7.6	3.3
19	mayroón	3.61132	-0.45034	11.3	9.7	8.0	3.8
20	nabanggít	3.50326	-0.42190	10.5	9.1	8.3	4.1
21	nagawâ	3.77278	-0.51862	13.0	11.2	7.3	3.0
22	paano	3.08071	-0.37546	9.4	8.1	8.2	4.0
23	pag-ibig	2.54924	-0.37425	9.4	8.1	6.8	2.6
24	pangakò	1.34057	-0.29528	7.4	6.4	4.5	0.3
25	paraán	3.04219	-0.38586	9.6	8.4	7.9	3.6
26	pintuan	3.09946	-0.35867	9.0	7.8	8.6	4.4
27	sapagká't	2.08263	-0.38488	9.6	8.3	5.4	1.2
28	sarili	4.97367	-0.46515	11.6	10.1	10.7	6.4
29	simbahan	2.77288	-0.35164	8.8	7.6	7.9	3.6
30	súsunod	2.96473	-0.35643	8.9	7.7	8.3	4.1
31	tagumpáy	3.19150	-0.41587	10.4	9.0	7.7	3.4
32	tahánan	2.36451	-0.37079	9.3	8.0	6.4	2.1
33	tinawag	3.43679	-0.44168	11.0	9.6	7.8	3.5
34	umaga	3.89301	-0.37992	9.5	8.2	10.2	6.0
	M	2.91688	-0.40101	10.0	8.7	7.1	2.9
	Min	0.69105	-0.65326	6.3	5.4	2.1	-2.1
	Max	6.73864	-0.25109	16.3	14.1	10.7	6.4
	Range	6.04759	0.40217	10.1	8.7	8.6	8.6
	SD	1.26280	0.08203	2.1	1.8	2.4	2.4

[a]*a* = regression intercept. [b]*b* = regression slope. [c]Psychometric function slope (%/dB) at 50% was calculated from 49.999 to 50.001%. [d]Psychometric function slope (%/dB) from 20-80%. [e]Intensity required for 50% intelligibility. [f]Change in intensity required to adjust the threshold of a word to the mean PTA of the subjects (4.25 dB HL).

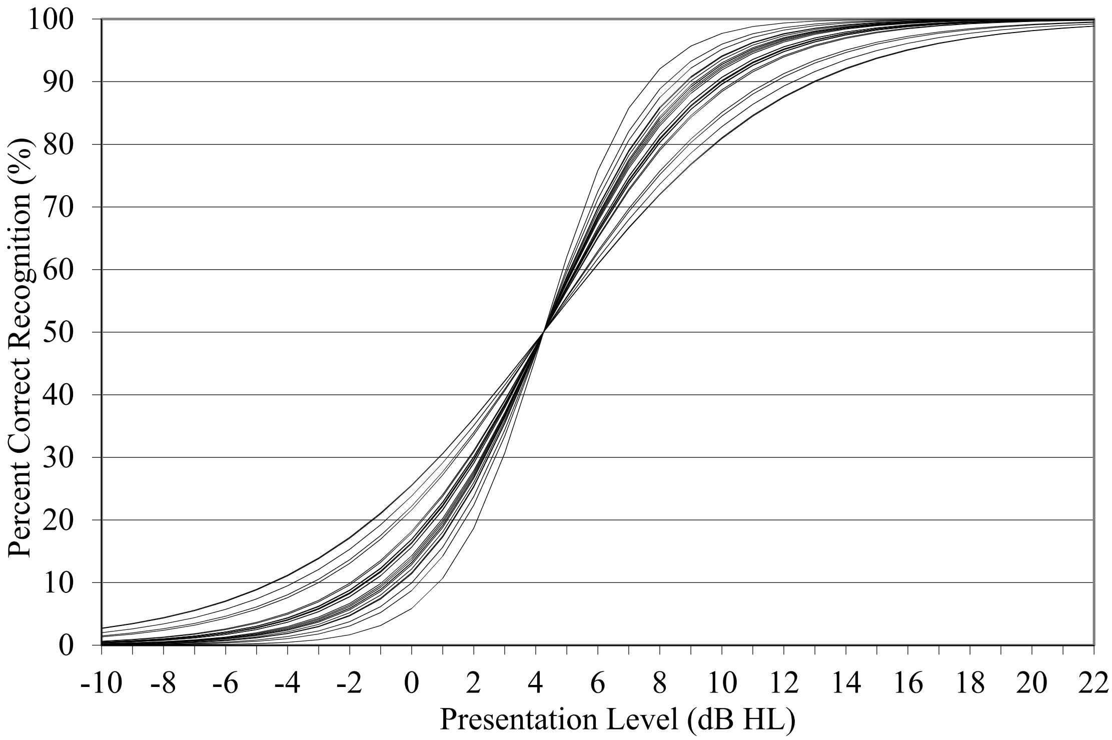

Figure 3

Predicted psychometric performance intensity functions for 34 selected trisyllabic Tagalog words digitally adjusted to equal the mean PTA of the subjects (4.3 dB HL).

psychometric performance-intensity function slopes for the 34 selected words, at 50% threshold ranged from 6.3%/dB to 16.3%/dB (*M* = 10).

Discussion

The goal of this study was to develop standardized materials for the audiometric assessment of SRT in native speakers of Tagalog and to make those materials readily available in high-fidelity digital format. The 34 trisyllabic Tagalog words contained in the final list constitute the results of efforts toward that goal. The lengthy process of creating these SRT materials was accomplished through careful selection of a preliminary body of Tagalog words, recording and evaluation of those words, and finally narrowing down the items to include a relatively homogenous list of Tagalog trisyllabic words suitable for SRT testing. This process yielded a final list of 34 trisyllabic Tagalog words found to be relatively homogenous in regard to audibility, and psychometric performance-intensity function slope. These words were spoken by a native male Tagalog speaker and recorded onto compact disc for distribution to audiologists worldwide.

The final 34 words underwent digital adjustments of intensity in order to increase the slope of the psychometric performance-intensity function of each word and homogeneity of audibility threshold for the words as a whole. The mean slopes from 20 to 80% for the psychometric performance-intensity functions of the 34 trisyllabic words ranged from 5.4%/dB to 14.1%/dB (*M* = 8.7). These values correspond with findings from other studies and materials created in other languages. SRT materials in English have mean slopes for spondaic words reported to be between 7.2%/dB and 10%/dB (Hirsh et al., 1952; Hudgins et al., 1947), but has been reported as high as 12%/dB (Beattie, Edgerton, & Svihovec, 1977; Ramkissoon, 2001). Similar slopes have resulted from 20 to 80% for the trisyllabic psychometric performance-intensity functions for both male and female talkers in materials created for other languages. For example, the mean slopes for Portuguese SRT materials were 9.1%/dB for a male talker, and 8.8%/dB for a female talker (Harris et al., 2001). For Japanese, the mean slope for SRT materials from 20 to 80% were reported at 8.9%/dB for the male speaker and 7.6%/dB for the female speaker (Mangum, 2005).

The words selected for use in this project were recorded by a male speaker only. In similar past studies recordings by male and female speakers have routinely been produced and included (Harris et al., 2001; Harris et al., 2007; Kim, 2007; Newman, 2010; Nissen et al., 2005). This convention, however, is unnecessary as studies suggest that speaker gender produces no clinically significant difference in recognition threshold scores for normal-hearing and hearing-impaired individuals (Cambron, Wilson, & Shanks, 1991; Penrod, 1979; Preece & Fowler, 1992). And since only one presentation (either male or female) can be administered to an individual at any given time, the inclusion of SRT materials by both speakers is redundant. While it is not essential to have words recorded by both male and female speakers, the inclusion of both types of recordings may be more convenient and does offer a choice when a female presentation is preferred by the audiologist.

Significant effort has been put into the process of developing and standardizing these and other SRT materials. By careful attention to the process of word selection and evaluation, and by following strict guidelines for calibrating, recording, and administering the test materials, materials can be produced that ensure an accurate and valid measure of thresholds for speech. Development of these speech audiometry materials, as well as similar past projects on behalf of other languages, represents an important advancement in the field of audiology. There remains, however, a need for further research in this area of study.

Determination of the levels of validity and reliability for the SRT materials developed in this study requires further research since no other standardized and recorded materials in Tagalog have been found. While a single study was discovered that developed word lists in Pilipino for speech threshold testing, no existing validity and reliability data for those materials have been found. Therefore, further research is warranted in order to ascertain measures of validity and reliability, such as test-retest reliability, for the materials created in this study.

The participants in this study were all shown to have normal hearing. Since the materials created in this study are intended for use with individuals who have possible hearing impairments, subsequent testing is needed to determine the efficacy of these materials in testing that population. A recent study reveals that items found to be homogeneous for speech reception testing in individuals with normal hearing may not be homogeneous for those with hearing impairment (McArdle & Wilson, 2006). Therefore, further testing should be conducted using these materials and involving native Tagalog speakers who exhibit hearing impairment.

Although the materials created in this study were intended for use with native Tagalog speakers, factors such as age, dialect, and number of languages that an individual speaks (monolingual, bilingual, multilingual) may affect the usefulness of these materials in assessing hearing thresholds for speech for any given individual. Since consideration of all such factors is beyond the scope of this study, there exists a need for further research to expand upon this project. An example of one area for potential expansion may be to develop a test suitable for use in testing children who speak Tagalog. Early identification and intervention is vitally important for children with hearing impairment. Assessing children using speech audiometry materials created for an adult population, however, would be unreliable, and would not correspond with the child’s receptive and expressive language abilities. Chronological age, as well as language exposure, needs to be considered in the development of speech audiometry materials appropriate for younger listeners. One such test was developed in German, and involved the use of sentences to assess speech discrimination in school-age children (Albrecht, 1973). Another test for use with children was developed in Arabic and involved the use of pictures as well as words that are more familiar to younger listeners (Ashoor & Prochazka, 1985). Since the materials developed in this study were not done so with the intention of testing children, the words selected for inclusion in the finished lists may not be appropriate or reliable for use in testing children.

Another need for further research involves language differences of native Filipinos. While Tagalog is widely used as one of the primary languages in the Philippines, there is a wide variety of other languages and dialects used throughout the country (Schachter & Otanes, 1972). It is clear that Tagalog materials would be inappropriate for a speaker of a different language, but the extent to which surrounding dialects differ from standard Tagalog, and the appropriateness of these materials for those who speak those dialects, need to be examined. A related issue for consideration is the number of languages spoken by potential listeners. All of the subjects who participated in this study were at a minimum, bilingual (speaking English and Tagalog). So it is unclear how performance might differ in monolingual subjects, and what effect, if any, the ability to speak multiple languages might have on testing results.

In summation, the goal of this study was to develop speech audiometry materials intended for the assessment of SRT in Tagalog. The resultant test materials are comprised of 34 trisyllabic Tagalog words recorded by a male native speaker. All words included in the final list have been carefully selected and have been found suitable for clinical testing of SRT. Each word was digitally adjusted to decrease threshold variability resulting in relative homogeneity with regards to audibility, as well as slope of the psychometric performance-intensity function. The 34 trisyllabic words are recorded on compact disc and available for clinical application worldwide.

References

Albrecht, R. (1973). Sentence speech audiometry test for young children. *Monatsschrift fur Ohrenheilkunde und Laryngo-Rhinologie, 107*(2), 65-69.

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: Acoustical Society of America.

ASHA. (1978). Guidelines for manual pure-tone threshold audiometry. *ASHA, 20*(4), 297-301.

ASHA. (1988). Guidelines for determining threshold level for speech. *American Speech-Language Hearing Association, 30*(3), 85-89.

ASHA. (1990). Guidelines for screening for hearing impairments and middle-ear disorders. *ASHA, 32*(Supplement 2), 17-24.

Ashoor, A. A., & Prochazka, T. (1985). Saudi Arabic speech audiometry for children. *British Journal of Audiology, 19*(3), 229-238.

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 a normal-hearing population. *Journal of Speech and Hearing Disorders, 42*(1), 60-64.

Bell, T. S., & Wilson, R. H. (2001). Sentence recognition materials based on frequency of word use and lexical confusability. *Journal of the American Academy of Audiology, 12*(10), 514-522.

Blaylock, N. (1999). *The Austronesian Language Family*: Brigham Young University.

Brandy, W. T. (1966). Reliability of voice tests of speech discrimination. *Journal of Speech and Hearing Research, 9*, 461-465.

Brandy, W. T. (2002). Speech audiometry. In J. Katz (Ed.), *Handbook of clinical audiology* (5th ed., pp. 96-110). Philadelphia: Lippincott Williams & Wilkins.

Bryant, W. S. (1904, July). *A Phonographic Acoumeter.* Paper presented at the Annual meeting of the American Otological Society, Atlantic City, N.J.

Cambron, N. K., Wilson, R. H., & Shanks, J. E. (1991). Spondaic word detection and recognition functions for female and male speakers. *Ear and Hearing, 12*(1), 64-70.

Carhart, R. (1951). Basic principles of speech audiometry. *Acta Oto-Laryngologica, 40*(1-2), 62-71.

Christensen, L. K. (1995). *Performance intensity functions for digitally recorded Spanish speech audiometry.* Brigham Young University, Provo.

Creston, J. E., Gillespie, M., & Krohn, C. (1966). Speech audiometry: Taped vs live voice. *Archives of Otolaryngology, 83*(1), 14-17.

Dillon, H. (1983). The effect of test difficulty on the sensitivity of speech discrimination tests. *Journal of the Acoustical Society of America, 73*(1), 336-344.

Dirks, D. D., Takayana, S., & Moshfegh, A. (2001). Effects of lexical factors on word recognition among normal-hearing and hearing-impaired listeners. *Journal of the American Academy of Audiololgy, 12*(5), 233-244.

Egan, J. J. (1979). Basic aspects of speech audiometry. *Ear, Nose, and Throat Journal, 58*(5), 190-193.

Erickson, L. J. (2000). *An Account of Austronesian*: Brigham Young University.

Fletcher, H., & Steinberg, J. C. (1930). Articulation testing methods. *Bell Telephone Systems Technical Publications, 1*(2), 17-21.

Fry, D. B. (1964). Modifications to speech audiometry. *International Audiology, 3*, 226-236.

Harris, R. W., Goffi, M. V. S., Pedalini, M. E. B., Gygi, M. A., & Merrill, A. (2001). Psychometrically equivalent Brazilian Portuguêse trisyllabic words spoken by male and female talkers. *Pr¢-Fono, 13*(1), 37-53.

Harris, R. W., & Greer, L. F. (1997). Italian speech audiometry materials [Compact Disc]. Provo, UT: Brigham Young University.

Harris, R. W., Nielson, W. S., McPherson, D. L., & Skarzynski, H. (2004). Psychometrically equivalent Polish bisyllabic words spoken by male and female talkers. *Audiofonologia, 25*, 1-15.

Harris, R. W., Nissen, S. L., Pola, M. G., McPherson, D. L., Tavartkiladze, G. A., & Eggett, D. L. (2007). Psychometrically equivalent Russian speech audiometry materials by male and female talkers. *International Journal of Audiology, 46*(1), 47-66.

Henry, T., & Museo di Santa Croce (Umbertide Italy). (2006). *Amicizia e francescanesimo : Luca Signorelli, Umbertide e la pala di Santa Croce*. Città di Castello (Perugia): Petruzzi.

Hirsh, I. J., Davis, H., Silverman, S. R., Reynolds, E. G., Eldert, E., & Benson, R. W. (1952). Development of materials for speech audiometry. *Journal of Speech and Hearing Disorders, 17*(3), 321-337.

Hudgins, C. V., Hawkins, J. E., Karlin, J. E., & Stevens, S. S. (1947). The development of recorded auditory tests for measuring hearing loss for speech. *The Laryngoscope, 57*, 57-89.

Iliadou, V., Fourakis, M., Vakalos, A., Hawks, J. W., & Kaprinis, G. (2006). Bi-syllabic, modern Greek word lists for use in word recognition tests. *International Journal of Audiology, 45*, 74-82.

Kim, M. N. (2007). *Psychometrically equivalent trisyllabic words for speech reception threshold testing in Cantonese.* Unpublished Unpublished master's thesis., Brigham Young University, Provo, UT.

Lau, C. C., & So, K. W. (1988). Material for Cantonese speech audiometry constructed by appropriate phonetic principles. *British Journal of Audiology, 22*(4), 297-304.

Ledesma, N. R. (1974). *The Development of Speech Audiometric Materials in Pilipino.* Kent State University.

Luce, P. A. (1986). A computational analysis of uniqueness points in auditory word recognition. *Perception and Psychophysics, 39*(3), 155-158.

Luce, P. A., & Pisoni, D. B. (1998). Recognizing spoken words: the neighborhood activation model. *Ear and Hearing, 19*(1), 1-36.

Mangum, T. C. (2005). *Performance intensity functions for digitally recorded Japanese speech audiometry materials.*, Brigham Young University, Provo.

Martin, F. N., & Clark, J. G. (2009). Speech Audiometry. In *Introduction to Audiology* (10th ed., pp. 126-164). Boston: Allyn and Bacon.

McArdle, R. A., & Wilson, R. H. (2006). Homogeneity of the 18 QuickSIN lists. *J Am Acad Audiol, 17*(3), 157-167.

Miller, G. A., Heise, G. A., & Lichten, W. (1951). The intelligibility of speech as a function of the context of the test materials. *J Exp Psychol, 41*(5), 329-335.

Nelson, A. (2004). *Performance intensity functions for digitally recorded French speech audiometry.* Unpublished Unpublished master's thesis., University of Colorado, Boulder, CO.

Newman, J. L. (2010). *Development of Psychometrically Equivalent Speech Recognition Threshold Materials for Native Speakers of Samoan.* Brigham Young University, Provo.

Nissen, S. L., Harris, R. W., Jennings, L., Eggett, D. L., & Buck, H. (2005). Psychometrically equivalent trisyllabic words for speech reception threshold testing in Mandarin. *International Journal of Audiology, 44*(7), 391-399.

Nissen, S. L., Harris, R. W., & Slade, K. B. (2007). Development of speech reception threshold materials for speakers of Taiwan Mandarin. *International Journal of Audiology, 46*, 449-458.

Olsen, S. O. (1996). Evaluation of the list of numerals in the Danish speech audiometry material Dantale. *Scandinavian Audiology, 25*(2), 103-107.

Penrod, J. P. (1979). Talker effects on word-discrimination scores of adults with sensorineural hearing impairment. *Journal of Speech and Hearing Disorders, 44*(3), 340-349.

Preece, J. P., & Fowler, C. G. (1992). Relationship of pure-tone averages to speech reception threshold for male and female speakers. *Journal of the American Academy of Audioilogy, 3*(3), 221-224.

Punch, J. L., & Howard, M. T. (1985). Spondee recognition threshold as a function of set size. *J Speech Hear Disord, 50*(2), 120-125.

Ramkissoon, I. (2001). Speech recognition thresholds for multilingual populations. *Communication Disorders Quarterly, 22*, 158-162.

Ramkissoon, I., Proctor, A., Lansing, C. R., & Bilger, R. C. (2002). Digit speech recognition thresholds (SRT) for non-native speakers of English. *American Journal of Audiology, 11*(1), 23-28.

Ramos, D. T. V. (2005, 6/2/2005). The Tagalog Language. Retrieved 4/23/2012, 2012, from http://www.seasite.niu.edu/Tagalog/Tagalog_Homepage99/the_tagalog_language.htm

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.

Rubino, C. R. G., & Llenado, M. G. T. (2002). *Tagalog-English, English-Tagalog dictionary = Taláhuluganang Pilipino-Ingglés, Ingglés-Pilipino Taláhuluganang* (Rev. & expanded ed.). Conshohocken, PA: Hippocrene Books.

Rudmin, F. (1987). Speech reception thresholds for digits. *Journal of Auditory Research, 27*(1), 15-21.

Schachter, P., & Otanes, F. T. (1972). *Tagalog reference grammar*. Berkeley,: University of California Press.

Shin, H. B., Kominski, R., & U.S. Census Bureau. (2010). *Language use in the United States, 2007*. from http://www.columbia.edu/cgi-bin/cul/resolve?clio8859454.

Thibodeau, L. M. (2007). Speech audiometry. In R. J. Roeser, M. Valente, & H. Hosford-Dunn (Eds.), *Audiology diagnosis* (2nd ed., pp. 288-313). New York: Thieme.

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*(1), 7-14.

Wilson, R. H., & Strouse, A. (1999). Psychometrically equivalent spondaic words spoken by a female speaker. *Journal of Speech, Language, and Hearing Research, 42*(6), 1336-1346.

The World Factbook. (2012, 4/17/2012). *The World Factbook* Retrieved 4/23/2012, 2012, from https://www.cia.gov/library/publications/the-world-factbook/geos/rp.html

World Health Organization. (2012, February 2012). Deafness and hearing impairment. Retrieved 4/23/12, 2012, from http://www.who.int/mediacentre/factsheets/fs300/en/index.html

Young, L. L., Dudley, B., & Gunter, M. B. (1982). Thresholds and psychometric functions of the individual spondaic words. *Journal of Speech and Hearing Research, 25*(4), 586-593.

Appendix - Informed Consent

Participant:______________________________________Age:__________

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

This research project is designed to evaluate speech materials recorded using improved digital techniques. You will be presented with lists of words or sentences at varying levels of intensity. Many will be very soft, but none will be uncomfortably loud to you. You may also be presented with these lists of words or sentences 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 or sentences, you will be administered a routine hearing test to determine that you qualify for this study.

It will take two to six hours to complete the test. Testing will be broken up into 2-6 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 complete your participation in this research project you will be paid the amount of $_________ for your participation.

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, Dr. Shawn L. Nissen, 138 TLRB, Brigham Young University, Provo, UT 84602, phone (801) 422-5056 or Dr. David L. McPherson, 129 TLRB, Brigham Young University, Provo, UT 84602, phone (801) 422-6458. If you have any questions regarding your rights as a participant in a research project you may contact the BYU IRB Administrator at (801) 422-1461, A-285 ASB, Brigham Young University, Provo, UT 84602, irb@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