2016-06-01

Psychometrically Equivalent Bisyllabic Word Lists for Spanish Pediatric Word Recognition Testing

Brenda Karina Peterson
Brigham Young University

Psychometrically Equivalent Bisyllabic Word Lists for Spanish Pediatric Word Recognition Testing

Brenda Karina Peterson

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

Master of Science

Richard W. Harris, Chair
David L. McPherson
Christopher D. Dromey

Department of Communication Disorders

Brigham Young University

June 2016

ABSTRACT

Psychometrically Equivalent Bisyllabic Word Lists for Spanish Pediatric Word Recognition Testing

Brenda Karina Peterson
Department of Communication Disorders, BYU
Master of Science

While recorded speech audiometry materials have been developed in Spanish for adults, currently there are no speech audiometry materials available in Spanish that are suitable for a pediatric population. Thus, the purposes of this study were (a) to analyze Spanish bisyllabic words previously developed for adults to determine the words' appropriateness when testing word recognition scores in native Spanish-speaking children and (b) to compare the logistic regressions results from the Spanish adult data obtained in a previous study with the logistic regressions from Spanish pediatric data obtained in the present study. Using common-use children's dictionaries in the Spanish language, a subset of child-appropriate words was adapted from a set of materials developed for Spanish adults. A total of 129 frequently used bisyllabic words were chosen; the words were split into five lists; four lists contained 26 words and one list contained 25 words; each was digitally recorded by both male and female Spanish talkers. Twenty native Spanish-speaking children with normal hearing between the ages of four and eight years were selected to listen to words to obtain psychometric functions. Each word was presented to the listener at 5 levels of intensity from -5 to 35 dB HL in increments of 10 dB. Custom software was used to control randomization, timing, and presentation of the words. The participants were not familiarized with the words prior to testing. The words received a ranked order based on performance to create lists and half-lists that were equivalent. Logistic regression was used to calculate psychometric functions for the lists and half-lists. Subsequently, a chi-square analysis was completed. The analysis revealed no statistical differences among the lists and half-lists for either male or female talkers. The mean bisyllabic psychometric function slopes for lists and half-lists were 5.0%/dB for the male-talker words and 5.2%/dB for the female-talker words. The 50% threshold for male and female were 16.2 dB HL and 15.5 dB HL, respectively.

Keywords: speech audiometry, word recognition testing, word lists, Spanish, pediatric, children

ACKNOWLEDGMENTS

I would like to extend my sincere gratitude to all the individuals who provided their help, sacrifice, and support throughout this project. I would like to thank my advisor, Dr. Harris, for his invaluable advice, guidance, and time devoted to this thesis. I would also like to thank my thesis committee members Dr. Dromey and Dr. McPherson for their insightful comments and advice. I wish to thank my cohort members for making these two years a fun and enjoyable experience. I am especially grateful for my parents, Diego and Estela, and my sister, Diana, who have been a constant source of support, love, and motivation, and for instilling in me the love and importance of pursuing an education. Foremost, none of these would have been possible without the love, patience, and never-ending support of my husband and best friend, Thomas, who helped me every step of the way.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Figure Page

LIST OF APPENDICES

Appendix Page

DESCRIPTION OF CONTENT

The body of this thesis is written as a manuscript suitable for submission to a peer-reviewed journal in speech-language pathology and audiology. This thesis will be suitable for submission to a peer-reviewed journal when combined with additional research concluding the age at which a child's speech thresholds approximate those of an adult. An annotated bibliography is presented in Appendix A.

Introduction

Speech Audiometry

Hearing is of great importance for the development of language, learning, and thought in children. Several aspects of language may be affected when a child cannot hear, including delays in the development of verbal expression, faults in pronunciation, and absence of oral language. If a child presents with poor listening behavior, all necessary audiological tests such as pure-tone testing and speech audiometry testing should be administered in order to reach a clear diagnosis.

Pure-tone audiometry is a tool used to measure hearing sensitivity. It aids clinicians in determining the softest sound audible to an individual at selected frequencies. Pure-tone audiometry can be done though bone and air conduction. Testing generally occurs between 250 and 8000 Hz for air-conduction testing and between 250 Hz to 4000 for bone-conduction testing (ASHA, 1978).

Speech audiometry tests are commonly used in clinical practice and are complementary to pure-tone audiometry. However, clinicians should not make assumptions about a patient's performance based solely on the pure-tone audiogram (Hamid & Brookler, 2006). Pure-tone testing uses tonal sounds which allows the clinician to identify hearing threshold levels, whereas speech audiometry uses a variety of words at selected intensities that the patient is asked to repeat and provides information on word recognition abilities (ASHA, 1978).

Speech audiometry is a valuable tool when performing audiological testing. An audiometric test would not be considered complete without a speech audiometry evaluation. Speech audiometry provides information regarding individual's sensitivity to speech stimuli and the understanding of speech at suprathreshold levels. Speech audiometry can be especially helpful when working with a pediatric population because it provides valuable information regarding overall auditory perception skills which can be of value in outlining the prognosis of speech, language, reading, and cognitive abilities of children (Mendel, 2008).

Complementary to pure-tone testing, speech audiometry can help determine the degree and type of hearing impairment of the patient, provide valuable information regarding the planning and implementation of additional forms of audiological habilitation, such as auditory training and perceptual training. Speech audiometry can also facilitate audiological rehabilitation management. It can aid audiologists in in selecting the best hearing aid fit for the individual (Hamid & Brookler, 2006). Speech audiometry also serves as a validity check for pure-tone thresholds, in other words speech audiometry is used to further clarify and interpret the data obtained during pure-tone testing (Mendel, 2008).

The two most important speech audiometry tests used in clinical practice are the speech recognition threshold (SRT) test, which uses spondaic words, and word recognition scores (WRS), which uses monosyllabic words. The SRT is defined as the lowest intensity level at which an individual can correctly recognize 50% of spondaic words from a closed-set list. In contrast, the WRS test requires an open-set list of monosyllabic words. In this test, word recognition is revealed by the percentage of correct answers or the exact repetition of the words heard (Hamid & Brookler, 2006; Martin & Clark, 2009).

The American Speech-Language-Hearing Association (ASHA, 1979) recommends the following method to obtain the SRT: First, the clinician presents five spondaic words to the patient at a previously determined starting level; then the clinician decreases the intensity level by 5 dB steps; the clinician then presents five additional words to the patient. This descending method continues until the patient misses five consecutive words at the same intensity level. The number of correct responses is totaled and the number obtained is then subtracted from the starting intensity level (in dB HL) used in the first step. The clinician then adds 2 dB as a correction factor to the result. The resulting number is now the SRT.

There is a close relationship between the SRT and the pure-tone average (PTA), which is the mean of the pure-tone thresholds at 500, 1000, and 2000 Hz (Lawson & Peterson, 2011). The SRT, recorded in dB HL, serves as a validity check for the PTA. The PTA and SRT values should agree within ± 12 dB (ASHA, 1978).

WRS testing results are the percentage of test items a person can identify correctly by ear; these results contribute to decisions regarding the location of possible lesions within the ear and the development of rehabilitation programs. The goal is to determine the patient's optimum performance for word recognition under controlled and standardized conditions (Schlauch, Anderson, & Micheyl, 2014).

Word recognition score (WRS) is the sole method available for determining speech discrimination. Word recognition score assesses an individual's ability to understand and repeat words presented at varying intensity levels. Currently, the use of this method is limited, and the benefits that it can render are not fully utilized in pediatric evaluations (Weiss & Dempsey, 2008). If a child is having difficulties in speech development and production, the evaluation of word recognition can be quite useful in determining the specific difficulties and mapping out a unique remediation plan. Word recognition score testing can also prove to be an effective way to monitor a child's progress during audiological treatment. Tracking the progression is useful with children who have a hearing impairment or an auditory processing disorder negatively affecting their speech production (Mendel, 2008).

Although several U.S. researchers have developed speech audiometric materials in Spanish (Cokely & Yager, 1993; Comstock & Martin, 1984; Ramkissoon, 2001), there are no current speech audiometry materials in Spanish that are suitable for children. As previously mentioned, the outcomes from the WRS allow us to make appropriate clinical decisions. These outcomes can reveal information about the area of damage and about the possible benefit of hearing aids or cochlear implants. Furthermore, WRS testing is used to guide a discussion with people about their degree of hearing impairment or about the possible benefit of using a sound amplification device (Schlauch et al., 2014).

The following is the methodology used to determine the WRS: for individuals with normal hearing the WRS is administered at 30 dB SL and 40 dB SL when working with individuals with a hearing impairment. The clinician presents a list of words, typically with 50 words contained in each list (Hamid & Brookler, 2006), and the patient is asked to repeat the words back to the clinician. Then, a scored percentage is calculated. The WRS is considered normal if the individual's score is at or above 90% accuracy (Hamid & Brookler, 2006).

Word Recognition Tests

The WRS test determines speech comprehension by presenting a list of words. ASHA (1978) recommends that the presentation method of all speech audiometry materials should be recorded rather than live voice to ensure standardization and accuracy when presenting the word lists. In the U.S., the traditional WRS test consists of 50 monosyllabic phonetically balanced (PB) words with a phonemic distribution that occurs with approximately the same frequency at which they occur in connected English discourse. However, a true phonetically balanced list is difficult to achieve (Martin & Clark, 2009). The idea that correct administration of a WRS test depends on the usage of phonetically balanced words is currently an area of further development and study. Some clinicians prefer to use half-lists consisting of 25-word lists instead of 50. This abbreviated version shortens examination time, but this productivity comes at the expense of precision (Schlauch et al., 2014). The original WRS word lists were the twenty PB-50 lists. Although they are still being used in some clinics and laboratories, the words contained in those lists are uncommon and difficult for unfamiliar listeners to identify, creating confusion concerning their use and interpretation.

CID W-22 Word Lists. The Central Institute for the Deaf (CID) W-22 lists were created to replace the PAL Auditory Test and the PB-list. Hirsh et al. (1952) constructed four lists of 50 phonetically balanced words, which constitute the materials for the CID W-22 test. The most important task in the creation of this test was the vocabulary selection to create a phonetically balanced word list. First, all the words needed to be monosyllabic with no repetition of words in the different lists. Second, the selected words needed to be familiar words. Third, the phonetic composition of the words needed to be very similar to that of the English language. This test has become widely used for WRS testing (Schlauch et al., 2014). Audiologists commonly use these word lists when measuring suprathreshold word recognition in quiet.

CNC Word Lists. These lists were originally designed for auditory research. Peterson and Lehiste (1962) developed ten 50-word lists. Each word within those lists contained approximately the same set of phonemes with a phonemic distribution that was comparative to the phonemic structure of CNC (consonant-nucleus-consonant) words. These CNC lists were improved and made more uniform by eliminating less-familiar words. The CNC lists do not contain any exceptionally rare words (Peterson & Lehiste, 1962).

NU-6 Word Lists. The Northwestern University Auditory Test No.6 (NU-6) word list was created by Tillman and Carhart (1966). Originally, the NU-6 test consisted of 95 words, and

104 words were later added. Currently the NU-6 consists of a comprehensive list of phonetically balanced monosyllabic words divided into four lists of fifty words each. These lists were developed with the purpose of improving the phonetic balance of the words in a previous test called the Auditory Test No. 4 and to increase the homogeneity in the set of word lists.

PBK Word Lists. The Phonetically Balanced Kindergarten (PBK) Test, developed by Haskins in 1949, has been a useful tool in assessing word recognition in children with hearing impairments (Meyer & Pisoni, 1999). Initially, the test consisted of four lists. However, only three of the lists were considered equivalent enough to be used clinically. The PBK is an open-set test and currently consists of three word lists containing 50 words each. The target population of this test is young children with profound hearing impairments. This test is typically used to identify good candidates for cochlear implants. Children who perform well on the PBK typically display high levels of performance on other perceptual tests (Meyer & Pisoni, 1999).

The most common lists used in clinical practice are the NU-6 and CID W-22 word lists (Hamid & Brookler, 2006). NU-6 word lists contain more words that are unusual or unfamiliar to most listeners than the CID W-22 lists do, making the NU-6 list more difficult to use in some cases. Audiologists make the choice of which word list is most appropriate based on the patient's needs and abilities. When selecting testing materials, it is important to consider the linguistic background of the patient so that the test results are better suited for the client being tested. Lists containing appropriate linguistic features improve familiarity, thus helping the listeners to improve their ability to discriminate between words (Comstock & Martin, 1984). This is especially true for children. When testing children, audiologists should use word lists that are familiar to the child (Ramkissoon, 2001).

Hearing Problems and Testing in the Hispanic Community

The word *Hispanic* has been used in the United States to describe Spanish speakers since 1990 (Battle, 2012). The Hispanic community is made up of a mixture of groups from many Latin American countries, including Mexico, Cuba, Puerto Rico, Dominican Republic, El Salvador, Nicaragua, Colombia, Ecuador, Peru, and Panama (Dana, 2007). According to the 2010 census, the Hispanic population at 50.5 million individuals is the largest minority group in the Unites States. In certain states such as California, Texas, and New Mexico, the number of Hispanic people has already exceeded the number of non-Hispanic white people. In 2010, people of Mexican origin comprised the largest Hispanic group, representing 63% of the total Hispanic population in the United States (U.S. Census Bureau, 2011).

The United States is currently experiencing an influx of Spanish speakers across almost all regions of the country. In general, immigrants tend to be less educated and more impoverished than the people already living in the United States. A higher proportion of children with hearing impairments tend to come from lower socioeconomic backgrounds than from middle- to high-income households. Hearing-impaired Hispanic children follow this trend and are more likely to come from economically disadvantaged families (Battle, 2012).

An estimate of the normal distribution of hearing problems within the Hispanic population indicates that an increasing number of Spanish-English bilingual persons will begin to be seen in audiology clinics in the United States. This increase is relevant to audiology clinics because Spanish tests of word recognition have not been standardized or widely adopted for use in the United States (von Hapsburg & Pena, 2002). Testing Spanish speakers using Spanish WRS tests provides more accurate results than English tests do for this population. The general definition of the term *bilingual* includes anyone who speaks two languages in their everyday life

(Baker, 2006). However, the degree to which an individual knows each language or their level of proficiency in both languages depends on many circumstances such as the number of languages learned, the age at which the languages were learned, the method used to learn the languages, and the individual's language skills such as writing, reading, and listening. How often a given individual uses the language on a daily basis is one of the main factors in influencing the bilingual proficiency of that individual. Due to the difference between linguistic performance between bilingual and monolingual Spanish speakers, it has been inferred that bilingual Spanish speakers vary significantly from monolingual Spanish speakers in an array of different language-processing tasks (Baker, 2006). Thus, we can determine that a bilingual speaker will not test equivalently to a monolingual speaker in either language.

A central concern in clinical speech audiometry research deals with how the monolingual English-speaking audiologist can obtain reliable speech audiometry measures with monolingual Spanish-speaking clients (Cokely & Yager, 1993; Comstock & Martin, 1984; Martin & Hart, 1978). Current research studies on Spanish/English bilingual listeners have determined that it is inappropriate to test all listeners with an English-only test simply because they appear to function on a day-to-day basis in English (Shi & Sanchez, 2010). Individuals who are bilingual take more time processing verbal materials in their second language than monolinguals do. When a bilingual speaker is given a limited time for auditory processing in their native language, their response time will be longer than their monolingual counterpart because the patient's two languages are always activated to some degree (von Hapsburg & Pena, 2002). Given the clear difference between English monolingual and Spanish/English bilingual perception of English speech, it seems advisable that Spanish/English bilingual listeners be evaluated in Spanish or, if necessary, in both languages (Shi & Sanchez, 2010).

Audiologists performing speech audiometry testing should take into consideration the need for tests that would be suitable to the linguistic background of their patients (Shi & Sanchez, 2010). Relatively few materials exist for the clinical assessment of clients who speak Spanish (Lopez, Martin, & Thibodeau, 1997). The initial studies on Spanish word recognition testing for audiological purposes appear to have been performed by Cancel in 1957. Within the population of individuals whose original language is Spanish, a large number of dialects exist that vary depending upon geographic location. The Puerto Rican influence exists in New York, the Cuban influence in Florida, and the Mexican influence in Texas and California. The dialectal differences are further noted within a single state, varying among communities as well as between urban and rural environments. These dialectal variations should be taken into consideration when administering Spanish word recognition tests (Weisleder & Hodgson, 1989). Mexican Spanish is the primary Spanish dialect in the United States, having increased by 54% with the largest numeric change, growing from 20.6 million in 2000 to 31.8 million in 2010 (U.S. Census Bureau, 2011). Thus, Mexican Spanish was chosen as the focus of this study.

Speech Audiometry in Children

Identifying hearing impairments and performing more audiological evaluations in newborns has led to an increase in the number of remediation cases in infants and young children. The addition of speech audiometry to the audiological evaluation of pediatric patients should be seen as vital for the development and completion of a profile of auditory function and hearing ability (Mendel, 2008). Although pure-tone threshold tests provide information about characteristics of the auditory system, they do not provide audiologists with a precise measure of a person's ability to receive and respond to speech (Weiss & Dempsey, 2008). Speech audiometry testing in children is dependent on a child's age, speech skills, hearing impairment, and ability to understand speech (Neumann et al., 2012). Speech audiometry testing can be performed in a sound field with young children, but this testing is not ear-specific. Headphones or insert phones should be used to test each ear individually. If the child cannot repeat words, then speech recognition with pictures or speech detection threshold (SDT) is used, which will help corroborate pure-tone averages to determine true hearing levels (ASHA, 1978).

For children whose language and vocabulary are undeveloped, it is not viable to perform a traditional SRT. An alternative is to administer an SDT. Modified methodologies should be used for children under the age of two or intellectually disabled children who developmentally stand in this age range. Contrary to SRTs that represent the individual's intelligibility, the SDTs represent detectability. The speech detection threshold is the lowest intensity level at which a listener may just detect the presence of a speech signal. Speech can be detected at intensity levels lower than it can be understood (Diefendorf, 1983). Speech detection threshold can be measured using headphones/insert phones or in a sound field, and the child's response to speech is evaluated. As mentioned before, SDT does not test the ability of the child to understand speech, but rather tests his awareness of speech. Different age-appropriate behavioral techniques can be used to aid in obtaining an SDT (Diefendorf, 1983). Some of these techniques are Tangible Reinforcement Operant Conditioning Audiometry and Visual Reinforcement Audiometry (Mendel, 2008).

When working with children older than two years of age, speech audiometry materials used to obtain WRS and SRT should meet the same criteria as those used with adults. The words used for testing should be highly intelligible and equally difficult (Diefendorf, 1983). Toys can be used as incentives to get WRSs and SRTs from children who are unable or unwilling to respond. Speech audiometry testing can also be performed using pictures where the child is asked to point to the picture of the word the audiologist says (Mendel, 2008).

Children are a group for whom some modifications of the traditional method of measuring WRS need to be made. These variations typically include adjusting the vocabulary to be more familiar to the younger population (Neumann et al., 2012) and using a picture-pointing activity to maintain the child's attention to the task (Mendel, 2008). Chronological age, alertness, and language development are examples of factors that must be accounted for when testing children (Weiss & Dempsey, 2008). The importance of each of these factors will lose significance over time as the child ages. It is important to select test materials within a child's receptive vocabulary competency, designate age-appropriate response tasks, utilize reinforcement, and reduce or alleviate memory load. These are all important factors that can affect the reliability and validity of pediatric measurement. It is often necessary to give the children breaks and encouragement in order to help them successfully complete WRS testing. Modifications of WRS testing are also often necessary for individuals who are not being tested in their native language. It is important that non-native listeners be presented with stimuli that have the unique linguistic characteristics of their native language (Shi & Sanchez, 2010).

In preparation for speech audiometry testing with children, the clinician should include a standardized vocabulary test. It is also possible to determine a child's level of vocabulary by parent or teacher reports or previous evaluations (Weiss & Dempsey, 2008). If these resources are not available, the clinician may use observation skills during conversation with the child to determine their approximate level of vocabulary. If a clinician tests a child using vocabulary that is above the child's level, the results will not be accurate (Madell, 1996).

Although WRS tests are often presented to Spanish-speaking adults in the United States, little emphasis has been placed on developing tests that can provide an accurate evaluation of Spanish-speaking children (Weiss & Dempsey, 2008). The need for standardized, digitally recorded speech audiometry materials for Spanish-speaking children led to the development of this research project. Thus, the purposes of this study were to analyze Spanish bisyllabic words previously developed for adults in order to evaluate their effectiveness in testing WRS in native Spanish-speaking children and to compare the logistic regressions that resulted from the Spanish adult data obtained in a previous study with the logistic regressions for Spanish pediatric data obtained in the present study.

Method

The following methodology was based on Bishop's (2009) study on psychometrically equivalent bisyllabic word lists for word recognition in Spanish. The word lists were standardized by testing native Spanish-speaking adults. All lists developed by Bishop were homogenous with respect to audibility and psychometric function slope.

Participants

Twenty native Hispanic children between the ages of four and eight years participated in this study. The ten males and ten females were of Mexican descent; six of the 20 subjects were born in Mexico, and the remaining 14 were born in the United States and had at least one parent from Mexico. It was determined that all participants speak Spanish at home on a daily basis. All participants' articulation development was typical. The Contextual Probes of Articulation Competence-Spanish (CPAC-S) test developed by Goldstein & Iglesias (2000) was administered to all participants as the initial step of the study. Children who scored in the typical range of articulation development according to the CPAC-S were considered eligible to participate in a comprehensive hearing screening. All participants had bilateral pure-tone air-conduction thresholds ≤ 15 dB HL at octave and mid-octave frequencies from 125 to 8000 Hz and static acoustic admittance between 0.3 and 1.4 mmhos with peak pressure between -10 and +50 daPa (ASHA, 1990). Table 1 displays a statistical summary of participant thresholds. All participants were recruited through word of mouth. A copy of the informed consent form can be found in Appendix B.

Materials

Words. Bisyllabic words were selected as stimuli for the WRS materials from the most commonly used words in Spanish populations as evaluated by Bishop (2009). In addition, bisyllabic words are preferred instead of the traditional monosyllabic words that are used in English WRS testing procedures, due to the relative infrequency of monosyllabic words in the Spanish language (Comstock & Martin, 1984). The performance intensity functions for the selected bisyllabic words are similar to those typically found in English monosyllabic materials (Cancel, 1968; Comstock & Martin, 1984; Weisleder & Hodgson, 1989).

A total of 250 trochaic words were rated by six native Spanish speakers. They were instructed to judge each word based on how familiar it would be to the average Spanish-speaking listener by using the following rating scale: 1 = extremely common, 2 = very common, 3 = averagely common, 4 = seldom used, 5 = rarely used. Any word receiving a score of 4 or 5 by any judge was discarded. The scores for each word were obtained by averaging the ratings from the group of judges. Words receiving a score of three or lower were included in the final selection of 250 words to be recorded. Words were also discarded if they were (a) thought to be culturally insensitive, (b) unfamiliar to native Spanish-speakers, (c) thought to represent inappropriate content, or (d) had multiple meanings for the same pronunciation. From the 250

Table 1

Pure-Tone Threshold (dB HL) Descriptive Statistics for 20 Normally Hearing Spanish-Speaking Pediatric Participants

Frequency (Hz)	*M*	*Minimum*	*Maximum*	*SD*
125	7.7	0.0	15.0	5.2
250	7.5	0.0	15.0	5.3
500	6.8	0.0	15.0	4.7
750	5.9	0.0	15.0	4.9
1000	5.3	0.0	10.0	4.7
1500	4.5	-5.0	10.0	4.7
2000	3.5	-5.0	15.0	5.6
3000	3.6	-5.0	15.0	6.7
4000	2.3	-10.0	15.0	6.0
6000	4.5	-5.0	15.0	7.2
8000	4.8	-5.0	15.0	7.0
PTA	5.2	-1.7	11.7	4.4

Note. PTA = arithmetic average of thresholds at 500, 1000, & 2000 Hz.

bisyllabic words selected, 22 were eliminated. From the 228 remaining words, 129 were found in one or more Spanish dictionaries for children: *Mi Primer Diccionario Larousse* (Girard, 2011), *Mi primer Diccionario Everest* (Gutiérrez, 2013), and *Mi Primer Diccionario Parragon* (Root, 2007).

Talkers. Native Spanish-speaking adults from Mexico recorded the WRS lists used. Eight adults (four male and four female) were recorded reading the word lists. Subsequently a panel of eight Spanish-speaking judges from Mexico evaluated the performance of each talker, ranking the talkers from best to worst based on accent, vocal quality, and pronunciation. The highest ranked male and female talkers were selected as the talkers for all subsequent recordings.

Recordings. A total of 250 trochaic words were recorded by the talkers. Recording sessions took place on the campus of Brigham Young University in the Eyring Science Center. The talkers were recorded using a Larson-Davis model 2541 microphone that was consistently placed approximately 15 cm from the mouth of each talker at a 0° azimuth. The microphone was covered with a 7.62 cm windscreen. The microphone used a Larson-Davis model 900B microphone preamp and a Larson-Davis model 2200C preamp power supply. An Apogee AD-8000 24-bit analog-to-digital converter was used to create the digital recordings, which were saved to a hard drive. 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. As the recording was performed in an anechoic chamber in the Eyring Science Center, approximately 0 dB SPL of ambient noise was present; this allowed for a signal-to-noise ratio near 65 dB during recording.

During the recording sessions, it was requested that the talker in the chamber produce each of the words a minimum of four times, pausing momentarily between each production. It was also requested that the talkers speak with their normal speech rates and intonation patterns. The first and last repetitions of each word were excluded from the study to avoid possible list effects. The words in the medial positions were evaluated by a native judge to determine the quality of each production; subsequently the production that was rated higher was chosen for testing. Productions which the judges determined were recorded poorly, mispronounced, or produced unnaturally by the talker were eliminated from the study or re-recorded.

Adobe Audition Version 2.0 (Adobe System Incorporated, 2004) was used to normalize the intensity of each word that was to be used in testing to yield the same average root mean square power as that of a 1000 Hz calibration tone in an initial attempt to equate test word threshold audibility (Harris, Nielson, McPherson, & Skarzynski, 2004; Wilson & Strouse, 1999). The edited words were then saved as individual 24-bit *wav* files.

Procedures

Custom software was created to control randomization, timing, and presentation of the words. This was accomplished by routing each of the individual files to the external input of a Grason Stadler model 1761 audiometer (Grason Stadler, Eden Prairie, MN). Each word was routed from the audiometer to the participant via a single TDH-50P headphone. A double-walled sound suite was used to carry out testing; the sound suite met ANSI S3.1 standards for maximum permissible ambient noise levels for the ear not covered, which was accomplished using one-third octave-bands (ANSI, 1999).

The external inputs to the audiometer were calibrated to 0 VU using a 1000 Hz calibration tone prior to the testing of each individual. Calibration of the audiometer was performed before and after data collection. This calibration was performed in accordance with

ANSI S3.6 specifications (ANSI, 2004). No changes in calibration were necessary throughout the course of data collections.

The participants were not familiarized with the list of bisyllabic words prior to testing. The 129 bisyllabic words were placed into five lists; four lists contained 26 words and one list contained 25 words, using randomization to assign the words to a list. After the words were grouped, the five lists were used in testing the first ten participants; these five lists were subsequently used to test the next group of ten participants. Appendix C shows the list of the 129 bisyllabic words and their definitions. In Bishop's (2009) thesis, the lists were presented to subjects in 5 dB increments. As the current study is being performed with children, it was determined that the lists could be presented in 10 dB increments in order to reduce the time of testing. A statistical analysis was performed to determine if there would be a significant difference between the slopes of these two methods. No significant variation was found in this analysis. Each list was presented at one of the five levels, which ranged from -5 to 35 dB HL and changed in increments of 10 dB. The words within the lists were randomized for each subject prior to presentation. The words were presented an equal number of times at each intensity level across the group of participants. Prior to testing, the participants were given the following instructions, both in English and Spanish to ensure the participant's understanding of the procedure:

> You will hear Spanish words, which may become louder or softer. At the very soft loudness levels, it may be difficult for you to hear the words. Please listen carefully and repeat out loud the word that you hear. If you are unsure of the word, you are encouraged to guess. If you have no guess, say, "I don't know." Do you have any questions?

A native interpreter was available to interpret any necessary information to the participants. Each of the lists was presented once per participant during each evaluation session. Participants responded to the stimuli by repeating the perceived word. Responses were determined to be either correct or incorrect as judged by a native Spanish speaker from Mexico. All data were recorded into an Excel spreadsheet.

Results

After the raw data were collected, each bisyllabic word received a difficulty ranking based on the number of times it was correctly identified across all intensity levels and participants. The 29 least-perceptible words were excluded. The remaining 100 words were divided into two lists of 50 words each. The 100 ranked words were divided between the two lists using an s-curve distribution, in order to create lists that were relatively equivalent in terms of psychometric audibility. For this, the 100 words in the male list were ranked according to number of correct responses. Words that were equally ranked were placed in random order before the s-curve distribution was completed. The first two words in the male list were assigned to separate lists (the first word to list one, the second to list two). The next two words were assigned in the same way but in reverse order (the third to list two, the fourth to list one). This distribution continued until all 100 words were divided into two lists of 50 words each. The same was done with the female words. Table 2 and Table 3 display the resulting two lists of 50 words for male and female talkers. Following the compilation of the two balanced 50-word lists, four half-lists were constructed, shown in Tables 4 and 5.

After developing the lists and half-lists, regression slopes and intercepts were calculated using logistic regression for each of the two lists and four half-lists for both the male and female recordings. These values were inserted into a modified logistic regression equation (Equation 1).

Table 2

Spanish Pediatric Male Bisyllabic Lists in Rank Order from Most Difficult to Easiest

List 1		List 2	
quince	ocho	arma	día
muerto	lista	coche	mejor
metro	jamás	duro	libro
aquí	pared	letra	error
cuenta	hija	contar	hombre
banco	ayer	para	leche
amor	once	juegan	este
ojo	calor	mezclar	rojo
algo	niño	lleno	papel
azúl	razón	color	curso
calle	correr	alto	niña
firma	hoja	falta	oro
nota	quemar	perro	igual
después	lleva	hijo	río
oír	claro	abrir	centro
blanco	noche	bajo	habla
loco	hambre	plata	guerra
fecha	mujer	entre	cara
autor	juego	querer	doble
café	doctor	listo	madre
país	mente	débil	mesa
normal	cuidar	cuadro	hacer
poder	regla	gente	cerca
nombre	cuatro	norte	llora
pronto	lejos	hora	lucha

Table 3

Spanish Pediatric Female Bisyllabic Lists in Rank Order from Most Difficult to Easiest

List 1		List 2	
arma	después	plata	normal
centro	lucha	quemar	banco
gente	hora	rojo	feliz
blanco	juegan	once	frío
listo	abrir	bebe	hambre
país	cuento	amor	cuadro
boca	lleva	hija	cama
dedo	nueve	norte	menos
doce	calle	coche	niña
dolor	calor	ojo	llenar
error	juego	cada	aquí
junto	cinco	algo	correr
lejos	querer	fecha	loco
ayer	hijo	lista	ocho
mujer	día	moda	para
nombre	papel	cuatro	lograr
madre	cara	lugar	río
nunca	guerra	noche	este
hoja	bajo	poder	nota
débil	duro	hombre	habla
comen	perro	poner	igual
color	llora	cuenta	pared
jamás	hacer	letra	mucho
contar	bien	lleno	doctor
dentro	libro	cerca	mejor

Table 4

Spanish Pediatric Male Bisyllabic Half-Lists in Rank Order from Most Difficult to Easiest

List 1A	List 1B	List 2A	List 2B
muerto	quince	arma	coche
metro	aquí	letra	duro
banco	cuenta	contar	para
amor	ojo	mezclar	juegan
azúl	algo	lleno	color
calle	firma	falta	alto
después	nota	perro	hijo
oír	blanco	bajo	abrir
fecha	loco	plata	entre
autor	café	listo	querer
normal	país	débil	cuadro
poder	nombre	norte	gente
ocho	pronto	hora	día
lista	jamás	libro	mejor
hija	pared	error	hombre
ayer	once	este	leche
niño	calor	rojo	papel
razón	correr	niña	curso
quemar	hoja	oro	igual
lleva	claro	centro	río
hambre	noche	habla	guerra
mujer	juego	doble	cara
mente	doctor	madre	mesa
cuidar	regla	cerca	hacer
lejos	cuatro	llora	lucha

Table 5

Spanish Pediatric Female Bisyllabic Half-Lists in Rank Order from Most Difficult to Easiest

List 1A	List 1B	List 2A	List 2B
centro	arma	plata	quemar
gente	blanco	once	rojo
país	listo	bebe	amor
boca	dedo	norte	hija
dolor	doce	coche	ojo
error	junto	algo	cada
ayer	lejos	fecha	lista
mujer	nombre	cuatro	moda
nunca	madre	lugar	noche
hoja	débil	hombre	poder
color	comen	poner	cuenta
jamás	contar	lleno	letra
después	dentro	cerca	normal
lucha	hora	feliz	banco
abrir	juegan	frío	hambre
cuento	lleva	cama	cuadro
calle	nueve	menos	niña
calor	juego	aquí	llenar
querer	cinco	correr	loco
hijo	día	para	ocho
cara	papel	lograr	río
guerra	bajo	nota	este
perro	duro	habla	igual
llora	hacer	mucho	pared
libro	bien	doctor	mejor

This equation was used to calculate the predicted percentage of correct performance at any specified intensity level. The resulting percent of correct values were then used to construct psychometric functions for each list and half-list. The original logistic regression equation 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 presentation level in dB HL. When Equation 1 is solved for p and multiplied by 100, Equation 2 is obtained:

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

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 presentation level into Equation 2, it is possible to predict the percentage correct at any specified intensity level. Percentage of correct recognition was calculated for each of the bisyllabic words for a range of -5 to 35 dB HL in 10 dB increments.

In order to calculate the intensity level required for a given proportion, Equation 1 was solved for i (see Equation 2). By inserting the desired proportions into Equation 2, it is possible to calculate the threshold (intensity required for 50% intelligibility), the slope (%/dB at threshold), and the slope (%/dB) from 20 to 80% for each psychometric function. A summary of the logistic regression slope and intercept values for each list and half-list for the male and female talker are presented in tables 6 and 7, respectively.

Thresholds for the pediatric bisyllabic word lists ranged from 16.1 to 16.3 dB HL (M = 16.2 dB HL) for the male talker words, and from 15.4 to 15.5 dB HL (M =15.5 = dB HL) for the female talker words. The mean psychometric function slope at 50% for the bisyllabic lists and

Table 6

Mean Performance of Pediatric Spanish Male Bisyllabic Lists and Half-Lists

List	a[a]	b[b]	Slope at 50%[c]	Slope 20%-80%[d]	Threshold[e]
1	3.08910	-0.19180	4.8	4.2	16.1
2	3.33850	-0.20450	5.1	4.4	16.3
M	3.21380	-0.19815	5.0	4.3	16.2
Minimum	3.08910	-0.20450	4.8	4.2	16.1
Maximum	3.33850	-0.19180	5.1	4.4	16.3
Range	0.24940	0.01270	0.3	0.3	0.2
SD	0.17635	0.00898	0.2	0.2	0.2
1A	3.08900	-0.18990	4.7	4.1	16.3
1B	3.09080	-0.19390	4.8	4.2	15.9
2A	3.33450	-0.20390	5.1	4.4	16.4
2B	3.34260	-0.20510	5.1	4.4	16.3
M	3.21423	-0.19820	5.0	4.3	16.2
Minimum	3.08900	-0.20510	4.7	4.1	15.9
Maximum	3.34260	-0.18990	5.1	4.4	16.4
Range	0.25360	0.01520	0.4	0.3	0.4
SD	0.14360	0.00747	0.2	0.2	0.2

[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.

Table 7

Mean Performance of Pediatric Spanish Female Bisyllabic Lists and Half-Lists

List	a[a]	b[b]	Slope at 50%[c]	Slope 20%-80%[d]	Threshold[e]
1	3.20440	-0.20660	5.2	4.5	15.5
2	3.25360	-0.21120	5.3	4.6	15.4
M	3.22900	-0.20890	5.2	4.5	15.5
Minimum	3.20440	-0.21120	5.2	4.5	15.4
Maximum	3.25360	-0.20660	5.3	4.6	15.5
Range	0.04920	0.00460	0.1	0.1	0.1
SD	0.03479	0.00325	0.1	0.1	0.1
1A	3.11120	-0.20060	5.0	4.3	15.5
1B	3.30470	-0.21310	5.3	4.6	15.5
2A	3.04420	-0.19760	4.9	4.3	15.4
2B	3.50280	-0.22740	5.7	4.9	15.4
M	3.24073	-0.20968	5.2	4.5	15.5
Minimum	3.04420	-0.22740	4.9	4.3	15.4
Maximum	3.50280	-0.19760	5.7	4.9	15.5
Range	0.45860	0.02980	0.7	0.6	0.1
SD	0.20670	0.01359	0.3	0.3	0.1

[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.

half-lists ranged from 4.8%/dB HL to 5.1%/dB HL ($M = 5.0$) for the male talker words, and from 5.2%/dB HL to 5.3%/dB HL ($M = 5.2$) for the female talker words. The average psychometric function slopes at the 20%-80%/dB for the bisyllabic lists and half-lists were 4.3%/dB HL for the male recordings and 4.5%/dB HL for the female recordings. Presented in Table 6 (male) and Table 7 (female) are the results for threshold, slope at threshold, and slope from 20%-80% for the Spanish pediatric lists and half-lists. Full lists and half-lists were created to be equivalent with regard to audibility and psychometric function slopes as shown in Figures 1 and 2.

A two-way chi-square analysis was done after compiling the WRS lists and half-lists to determine the presence of a significant difference among the two 50-word lists and four 25-word half-lists. No significant difference was found between male and female talkers for lists and half-lists. $\chi^2(1, N = 20) = 0.66, p = .4179$, $\chi^2(3, N = 20) = 0.66, p = .4181$, respectively. There were no significant differences among the two full lists or among the four half-lists, $\chi^2(1, N = 20) = 0.02, p = .8898$, $\chi^2(3, N = 20) = 0.02, p = .9990$, respectively. There was likewise no significant difference found for gender × list interactions for lists or half-lists, $\chi^2(1, N = 20) = 0.09, p = .7677$, $\chi^2(3, N = 20) = 0.13, p > .9885$, respectively.

Discussion

The purposes of this study were to test the effectiveness of previously recorded psychometrically equivalent word lists to be used for word recognition testing in the Spanish pediatric population, as well as to compare the logistic regressions obtained from the Spanish adult data in a previous study with the logistic regressions from Spanish pediatric data in the present study.

The results obtained in the present study were not equivalent with the adult data obtained from Bishop (2009). Adult mean psychometric function slope at 50% for the male bisyllabic

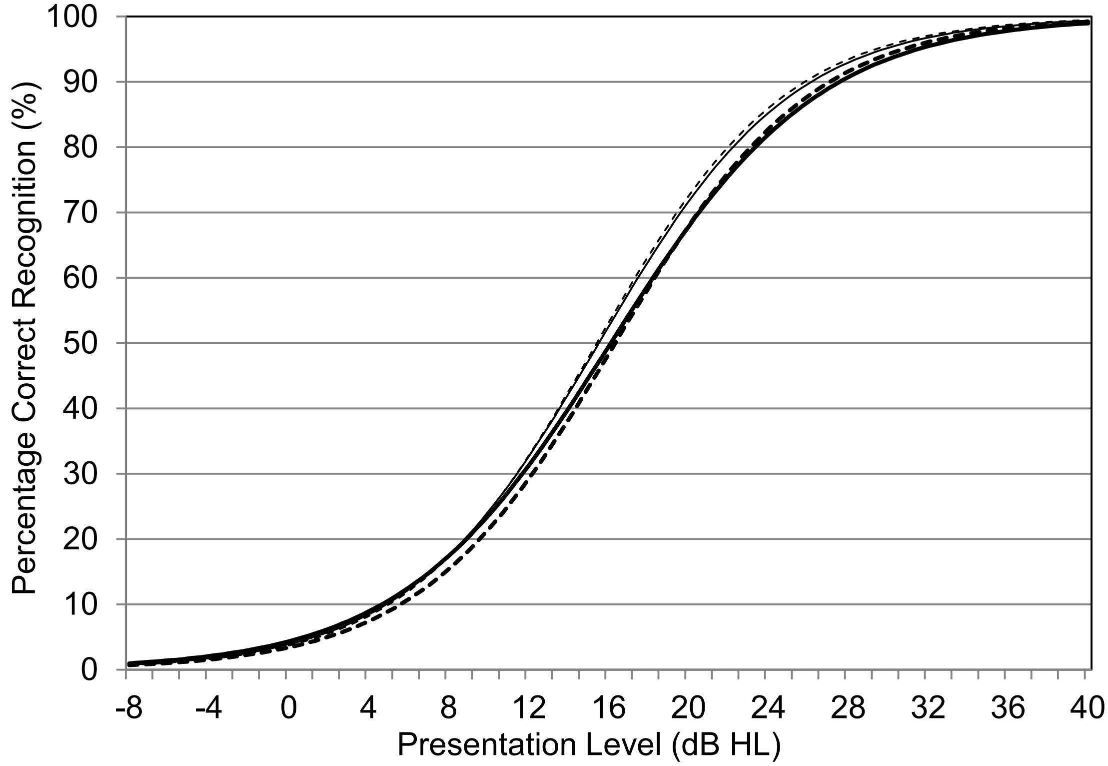

Figure 1.

Psychometric functions for two Spanish pediatric bisyllabic lists (50 words each) for male talker and female talker recordings.

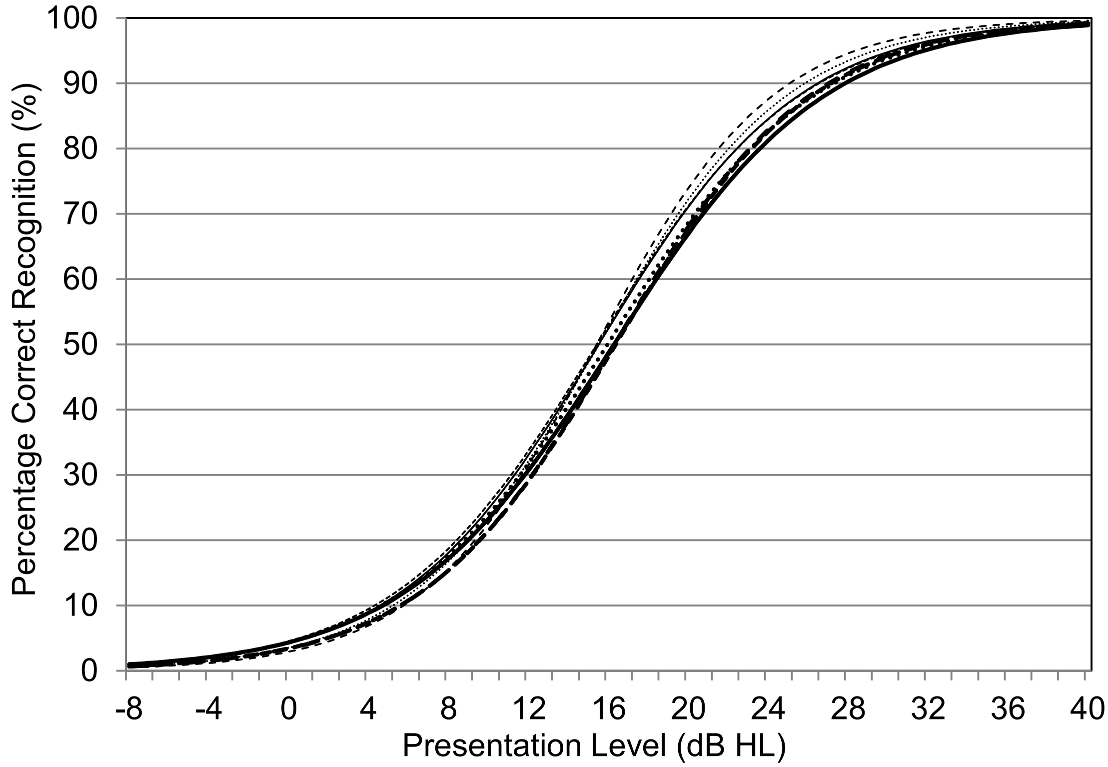

Figure 2.

Psychometric functions for four Spanish pediatric bisyllabic half-lists (25 words each) for male talker and female talker recordings.

lists and half-lists was 5.3%/dB with a mean threshold of 9.0 dB HL. Similarly, the pediatric slope for the lists and half-lists was 5.0 %/dB, showing a minimally less steep slope. However, a much higher threshold of 16.2 dB HL was obtained for the pediatric male talker words. The selected adult female talker words had a mean slope at 50% of 5.8%/dB HL with a mean threshold of 8.4 dB HL. In comparison, the total female pediatric slope was 5.2%/dB, once again similar to the adult results; however, a higher threshold of 15.5 dB HL was obtained. All adult lists underwent a digital intensity adjustment in order to increase the lists' psychometric equivalency. These adjustments ensured that the 50% threshold of all lists and half-lists would be equal to 8.73 dB HL, the average of the male and female list thresholds. Pediatric word lists cannot be adjusted due to the limited allowance in the recordings to meet the same word list's threshold audibility for either the female or male word lists obtained in Bishop's study.

The differing results show that the pediatric thresholds were higher than adult results by 7.3 dB HL for male talkers and 7.1 dB HL for female talkers; in other words, children required approximately 7 dB more intensity to perform the same as adults. This difference in performance between children and adults raises the question as to how such a contrasting outcome was obtained. Limitations of the present study may include task-related errors. Some possible errors are the subject's response time, the children's phonological development, the words used to create the word lists, and the level of motivation and focus of the participants, all of which could be possible avenues for additional research.

Research in Mongolian has yielded results similar to those found in the current study. The average for the psychometric function slopes at 50% for the pediatric WRS word lists and half-lists in Mongolian was 6.41 %/dB for the male recordings and 5.84 %/dB for the female recordings (Caldwell, 2009). Similarly, the Spanish mean psychometric function slope at 50%

for the pediatric bisyllabic lists and half-lists were 5.0%/dB for the male talker words and 5.2%/dB for the female talker words. For the pediatric monosyllabic Mongolian word lists, the mean threshold for 50% intelligibility was 11.5 dB HL for the male lists and 15.3 dB HL for the female lists (Caldwell, 2009). Comparably, thresholds for the Spanish pediatric bisyllabic word lists were 16.2 dB HL for the male talker words and 15.5 dB HL for the female talker words.

In another study, WRS word lists were developed for Modern Greek–speaking children between 6 and 12 years old. The test was designed in an open-set format and consisted of two lists of 50 bisyllabic words each. Similar to the present study, phonemically balanced bisyllabic words were chosen as stimuli due to the limited number of monosyllabic words in Modern Greek. The authors of this study tested the two word lists on children with normal hearing and children with different types of hearing loss. The Modern Greek mean psychometric function slope for the pediatric bisyllabic lists was 5.08% for the first word list and 5.24% for the second word list when testing normal-hearing children, with a mean threshold of approximately 17 dB HL for the first list and 16.5 dB HL for the second list (Trimmis, Papadeas, & Papadas, 2008). Again, these results are comparable to the ones obtained in the present study.

Conclusions and Future Research

The goals of this study were (a) to test the effectiveness of Spanish bisyllabic words, previously developed for adults for the pediatric population and (b) to compare the results of this study with the Spanish adult data obtained in a previous investigation. Children between the ages of four and eight years old with normal hearing and typical language development participated in this study. Looking at the adult data, the mean 50% for male talkers was 9.0 dB HL and 8.4 dB HL for female talkers. For pediatric subjects, the mean threshold was 16.2 dB HL for male and 15.5 dB HL for female talkers. This represents a 7.3 dB difference for male and 7.1 dB for female talkers. The discrepancy between the adult and pediatric results is likely due to the population used in this study (adults vs. children).

Even though all the words used in this study were age appropriate for children between the ages of four and eight, the children's level of exposure to the words—meaning, how frequently the word is used or heard—may have contributed to the disparity between the adults' and children's required intensity levels. It is possible that adults' increased exposure to the words enabled them to recognize and respond to the words more quickly than the children. Using words that occur more frequently in the child's daily language could potentially decrease children's response times. Also, a future extension of this research could include performing this study in the English language to compare the resulting response times of the subjects to the ones obtained in Spanish.

It has been suggested that for some speech sounds, children may require some listening experience with the sounds to show evidence of discrimination (Diefendorf, 1983). Young children's difficulty discriminating between fricatives compels us to take into account the role of phonological development in speech sound discrimination. Children's higher thresholds may have been due to the differences in phonological development between the children and adults.

Furthermore, the level of engagement of the participants during data collection may have contributed to the results obtained in this study. The monotonous nature of the task as well as the focus required may have caused children to lose interest in the activity. A more engaging methodology while performing the study should be considered in future research.

Drabman and Lahey's (1974) study has shown that praising children for following a rule, completing an assignment, or paying attention to an adult, results in an increased frequency of the desired behavior. Thus, verbal praise and more attractive rewards may be added to future testing procedures to keep the participants engaged and focused on the task. Also, simplifying and creating more child-friendly instructions may aid in accomplishing this goal.

Response time is another potential area for further research. Participants' responses could be recorded and the time between word presentation and the subjects' response could be measured. Also, children of increasing ages could be evaluated to see how response time changes as age increases.

Limitations of this study may be due primarily to task-related errors during data collection. Future research should account for the children's level of focus and motivation during testing, investigate a comparison of the subjects' response time at increasing ages, and examine the children's approximate level of exposure to the words used to create the pediatric word lists. Future research should also utilize the methods from this study in alternate languages, especially English. A study could be designed to measure response time comparing when the words are presented and when the child's oral responses are recorded. Children of increasing ages with a number of children in each age group up to young adult ages should be evaluated to identify and analyze the varying response time between the different age groups and to define the amount of time required to validly assess a child's speech recognition abilities at specific age increments. As this is the initial research in developing child-appropriate WRS materials in Spanish, further research and replication is required to determine the validity and reliability of the words chosen for the word lists as well as the performance of the judges' ratings.

Despite the limitations of this study, it is anticipated that the materials created in this study will aid in the accurate evaluation of word recognition testing in Spanish-speaking children. Two lists of 50 bisyllabic words and four half-lists of 25 bisyllabic words were recorded by both male and female talkers. All word lists are equivalent in performance and can be used for testing WRS in Spanish-speaking children. The final word lists were produced on compact disc for future clinical use. It is the hope that these word lists can be a baseline for future research into Spanish pediatric speech audiometry, as well as to assist clinicians working with children in Spanish-speaking countries and beyond.

References

Adobe Systems Incorporated. (2004). *Adobe Audition* (Version 1.5) [Computer software]. San Jose, CA: Adobe Systems Incorporated.

ANSI. (1999). *Maximum permissible ambient noise levels for audiometric test rooms*. ANSI S3.1-1999. New York: ANSI.

ANSI. (2004). *Specification for audiometers*. ANSI S3.6-2004. New York: ANSI.

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

ASHA. (1979). Guidelines for determining the threshold level for speech. *ASHA*, *21*, 353–355.

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

Baker, C. (2006). *Foundations of bilingual education and bilingualism*. Buffalo, NY: Multilingual Matters.

Battle, D. E. (2012). *Communication disorders in multicultural and international populations*. Buffalo, NY: Mosby.

Bishop, A. (2009). *Psychometrically equivalent bisyllabic word lists for word recognition in Spanish.* (Unpublished master's thesis). Brigham Young University, Provo, UT.

Caldwell, M. E. (2009). *Development of psychometrically equivalent speech audiometry materials for testing children in Mongolian.* (Unpublished master's thesis). Brigham Young University, Provo, UT.

Cancel, C. A. (1957). *Effects of selected phonetic aspects in the transmission of the Spanish language*. (Unpublished doctoral dissertation). Ohio State University, Columbus, OH.

Cancel, C. (1968). Spanish speech audiometry. *International Journal of Audiology, 7*, 206-208.

Cokely, J. A., & Yager, C. R. (1993). Scoring Spanish word-recognition measures. *Ear & Hearing, 14*, 395-400.

Comstock, C. L., & Martin, F. N. (1984). A children's Spanish word discrimination test for non-Spanish-speaking clinicians. *Ear & Hearing, 5*, 166-170.

Dana, L. P. (2007). *Handbook of research on ethnic minority entrepreneurship: A co-evolutionary view on resource management.* Cheltenham, Great Britain: Edward Elgar Publishing.

Diefendorf, A. O. (1983). Speech audiometry with infants and children. *Seminars in Hearing, 4*, 241-252. doi:10.1055/s-0028-1091428

Drabman, R. S., & Lahey, B. B. (1974). Feedback in classroom behavior modification: Effects on the target and her classmates. *Journal of Applied Behavior Analysis*, *7*, 591–598.

Girard, F. (Ed.). (2011). *Mi primer diccionario Larousse.* Barcelona, España: Larousse.

Goldstein, B. & Iglesias, A. (2000). *Contextual Probes of Articulation Competence-Spanish, First Edition*. Greenville, SC: Super Duper Publications.

Gutiérrez, A. C. L. (Ed.) (2013). *Mi primer diccionario*. León, España: Everest.

Hamid, M. A., & Brookler, K. H. (2006). Speech audiometry. *Ear, Nose, and Throat Journal, 85*, 810-812.

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*, 1-13.

Haskins, H. L. (1949). *A phonetically balanced test of speech discrimination for children.* (Unpublished master's thesis). Northwestern University, Evanston, IL.

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*, 321-337.

Lawson, G., & Peterson, M. (2011). *Speech audiometry*. San Diego, CA: Plural

Lopez, S. M., Martin, F. N., & Thibodeau, L. M. (1997). Performance of monolingual and bilingual speakers of English and Spanish on the Synthetic Sentence Identification test. *American Journal of Audiology, 6*, 33-38.

Madell, J. (1996). *Speech audiometry for children*. Washington, DC: Gallaudet University.

Martin, F. N., & Clark, J. G. (2009). *Introduction to audiology*. Boston, MA: Allyn and Bacon.

Martin, F. N., & Hart, D. B. (1978). Measurement of speech thresholds of Spanish-speaking children by non-Spanish-speaking clinicians. *Journal of Speech and Hearing Disorders, 43*, 255-262.

Mendel, L. L. (2008). Current considerations in pediatric speech audiometry. *International Journal of Audiology*, *47*, 546-553.

Meyer, T. A., & Pisoni, D. B. (1999). Some computational analyses of the PBK test: Effects of frequency and lexical density on spoken word recognition. *Ear & Hearing, 20*, 363-371.

Neumann, K., Baumeister, N., Baumann, U., Sick, U., Euler, H. A., & Weißgerber, T. (2012). Speech audiometry in quiet with the Oldenburg Sentence Test for Children. *International Journal of Audiology*, *51*, 157-163.

Peterson, G. E., & Lehiste, I. (1962). Revised CNC lists for auditory tests. *Journal of Speech and Hearing Disorders, 27*, 62-70.

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

Root, B. (Ed.). (2007). *Mi primer diccionario*. Beijing, China: Parragon.

Schlauch, R. S., Anderson, E. S., & Micheyl, C. (2014). A demonstration of improved precision of word recognition scores. *Journal of Speech, Language, and Hearing Research, 57*, 543-555. doi:10.1044/2014_JSLHR-H-13-0017

Shi, L. F., & Sanchez, D. (2010). Spanish/English bilingual listeners on clinical word recognition tests: What to expect and how to predict. *Journal of Speech, Language, and Hearing Research, 53*, 1096-1110. doi:10.1044/1092-4388(2010/09-0199)

Tillman, T., & Carhart, R. (1966). *An expanded test for speech discrimination utilizing CNC monosyllabic words: Northwestern University Auditory Test No. 6*. Brooks Air Force Base, TX: US Air Force School of Aerospace Medicine.

Trimmis, N., Papadeas, E., & Papadas, T. (2008). A Modern Greek word recognition score test designed for school-aged children. *Mediterranean Journal of Otology*, *4*, 1-8.

U.S. Census Bureau. (2011). *The Hispanic population: 2010.* Retrieved from: http://www.census.gov/prod/cen2010/briefs/c2010br-04.pdf

Von Hapsburg, D., & Pena, E. D. (2002). Understanding bilingualism and its impact on speech audiometry. *Journal of Speech, Language, and Hearing Research, 45*, 202-213. doi:10.1044/1092-4388(2002/015)

Weisleder, P., & Hodgson, W. R. (1989). Evaluation of four Spanish word-recognition-ability lists. *Ear & Hearing, 10*, 387-392.

Weiss, D., & Dempsey, J. J. (2008). Performance of bilingual speakers on the English and Spanish versions of the Hearing in Noise Test (HINT). *Journal of the American Academy of Audiology*, *19*, 5-17.

Wilson, R. H., & Strouse, A. (1999). Psychometrically equivalent spondaic words spoken by a female speaker. *Journal of Speech, Language, and Hearing Research, 42*, 1336-134

Appendix A

Annotated Bibliography

Adobe Systems Incorporated. (2004). *Adobe Audition* (Version 1.5) [Computer software]. San Jose, CA: Adobe Systems Incorporated.

Summary: The American National Standards Institute (ANSI) specifies maximum permissible ambient noise levels (MPANLs) allowed in audiometric test rooms. This ambient noise level produces negligible masking ($\leq$ 2 dB) of test signals. The MPANLs are specified from 125 HZ to 800 Hz in octave and one-third octave band intervals for two audiometric testing conditions (ears covered and ears not covered) and for three test frequency ranges: 125 to 8000 Hz, 250 to 8000 Hz, and 500 to 8000 Hz. The standard is designed to be used by all professionals testing hearing, suppliers, manufacturers of audiometric rooms, and installers.
Relevance to Current Work: The suite used for testing in this study met ANSI S3.1 standards for maximum permissible ambient noise levels for the ear not covered, which was accomplished using one-third octave bands.

ANSI. (1999). *Maximum permissible ambient noise levels for audiometric test rooms*. ANSI S3.1-1999. New York: ANSI.

Summary: This standard specifies the maximum permissible ambient noise levels (MPANL) that are allowed in an audiometric test room that will generate a negligible masking of 2 dB or less of the test signals presented at the reference equivalent threshold levels specified in ANSI S3.6-1996. The MPANLs stipulated a range from 125 to 8000 Hz in octave and one-third octave band intervals for two audiometric testing conditions—ears covered and ears not covered—and for three test frequency ranges (125 to 8000 Hz, 250 to 8000 Hz, and 500 to 8000 Hz). It is intended that the standard is used by all individuals testing hearing and for distributors, installers, designers, and manufacturers of audiometric test rooms.
Relevance to Current Work: The sound booth used for the current study exceeded the requirements set by this standard.

ANSI. (2004). *Specification for audiometers*. ANSI S3.6-2004. New York: ANSI.

Summary: The purpose of this standard is to provide consistency in testing an individual's hearing when using different audiometers. This standard provides specifications for different audiometers including specifications for recording speech test materials and for audiometers used primarily to determine hearing threshold levels with respect to standard reference threshold levels. Equipment setup, calibration, and safety criteria are also outlined in this standard.
Relevance to Current Work: The calibration of the audiometer used in this study was performed in accordance with this standard to ensure that the materials generated in this study could be useful to future clinicians in different settings.

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

Summary: The guidelines presented in this manual set standards for manual pure-tone audiometry with the purpose of minimizing variance between tests performed by audiologists and audiometric technicians in different settings. It is important to specify how one obtained the thresholds when variations in the procedure are made, to ensure that other clinicians can successfully replicate the results in the future. In order to obtain accurate measurements, the audiometric equipment must function properly, be correctly calibrated, and be used in an adequate test setting.
Relevance to Current Work: The guidelines presented in this manual enhance clinical reliability. The guidelines also provide useful information about the relationship between pure-tone threshold audiometry and speech audiometry.

ASHA. (1979). Guidelines for determining the threshold level for speech. *ASHA*, *21*, 353–355.

Summary: The guidelines presented in this manual were prepared by members of the ASHA Committee on Audiologic Evaluation. The main purpose of these guidelines is to present a standardized measurement method that has been statistically validated and should yield a speech recognition threshold that agrees closely with the pure tone average. The document includes background, scope, and a brief overview of speech threshold audiometry. The basic purpose of a speech threshold is to quantify an individual's hearing threshold level for speech. The document states that the primary purpose of a speech threshold is to serve as a validity check for the pure tone audiogram. These guidelines also include the suggested methodology for obtaining SRT.
Relevance to Current Work: This standardized process for determining speech recognition thresholds provides valuable information that will improve interclinician and interclinic comparison, which will benefit the individuals we serve.

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

Summary: The purpose of these guidelines is to identify individuals that have a potential ear disorder that could potentially affect their communication and daily activities. People identified to have a hearing problem using these guidelines should be referred to receive a medical examination or an audiology test. This inspection protocol includes four parts: case history, visual inspection, pure-tone audiometry, and tympanometry. Performing a visual inspection is an important step and is recommended to precede the tympanometry portion of the protocol. The visual inspection is expected to identify problems that are not evident using other procedures in the screening battery. When detecting structural defects, ear-canal abnormalities, and eardrum abnormalities, an immediate referral to receive medical attention must take place. When performing an audiometry screening, it is recommended that screening with pure-tone stimuli be presented at 2-dB HL at frequencies from 1000 to 4000 Hz. Individuals failing at any of these frequencies results in failure of the audiometric screening.
Relevance to Current Work: The subjects who participated in this study were screened following the guidelines presented in this protocol.

Baker, C. (2006). *Foundations of bilingual education and bilingualism*. Buffalo, NY: Multilingual Matters.

Summary: The purpose of the book is to introduce the reader to bilingualism and bilingual education. It provides detailed information concerning the development of bilingualism, languages in society, bilingual education in the U.S., minorities, and multicultural societies. Baker provides a series of definitions to important terms necessary to understand the issues discussed in the book. The author emphasizes the important distinction between individual bilingualism and the use of two or more languages in a community. The book also discusses the issue of measuring bilingualism in terms of language proficiency and use. It gives an overview on early and later development of bilingualism; Baker also includes age and language acquisition along with identity and language acquisition. He analyzes sociolinguistic topics including language shift, language maintenance, and language death; he also discusses English as a Global Language and its spread.
Relevance to Current Work: The book provided an introductory framework on bilingualism. When working with bilingual individuals in the present study, it was important to understand definitions and concepts concerning this topic.

Battle, D. E. (2012). *Communication disorders in multicultural and international populations*. Buffalo, NY: Mosby.

Summary: This book discusses the development of speech and hearing disorders and the treatments for these disorders when dealing with individuals from the diverse cultures. The author covers a wide variety of cultures and age populations. In a country as culturally diverse as the U.S., it is imperative that clinicians develop cultural sensitivity, knowledge, and respect for other cultures. A hearing impairment could be caused by many different factors, such as genetics, drugs, and infections. Even though deafness is not specific to a particular race, culture, age, or gender, there are differences between ethnic groups depending on the etiology that is being considered. For example, the main cause of deafness in the Hispanic community is rubella, whereas in the African American population the leading cause of deafness is meningitis. The book also discusses the role that socioeconomic status plays in health. On average, white Americas have more access to better health services and healthier lifestyles, and they live in healthier environments. They also have better access to preventive medical services than minorities. Hearing loss is a common disorder throughout the world, and according to the WHO it is most prevalent in developing countries.
Relevance to Current Work: The purpose of this book is to provide SLPs and audiologists with a framework for how to better serve people from different cultures. It provides informative data about the prevalence and treatment of hearing loss in varying cultures, including the Hispanic community. This information is useful in aiding professionals to provide better services when working with this population.

Bishop, A. (2009). *Psychometrically equivalent bisyllabic word lists for word recognition in Spanish.* (Unpublished master's thesis). Brigham Young University, Provo, UT.

Objective: The aim of this study was to develop, digitally record, evaluate, and psychometrically equate a set of Spanish bisyllabic word lists to be used for word recognition testing with the adult population. Frequently used bisyllabic words were selected and digitally recorded by a Spanish-speaking male and female.
Design: The word lists were created for measuring word recognition score (WRS) in the Spanish population. A total of 250 common bisyllabic words were chosen.
Study Sample: Twenty native Spanish-speaking adults with normal hearing listened to the word lists to find the percentage of words that they correctly identified. Each word was presented to the listener at 10 intensity levels, from -5 to 50 dB in increments of 10 dB.
Results: Once the data was compiled, a logistic regression was used to obtain regression slopes for all 250 words. The 200 words with the highest logistic regression slopes were then divided into four balanced lists of 50 words each and eight half-lists each containing 25 words. A two-way Chi-Square revealed no significant diffrence among lists, half-lists, and male and female recordings.
Conclusions: The purpose of this study was to create digitally recorded bisyllabic Spanish word recognition lists and half-lists of familiar words. The recorded 50-word lists and 25-word half-lists for both the male and female talkers are relatively homogenous in regard to audibility and psychometric function slope. The words were recorded on compact disc to be used for word recognition testing on native Spanish-speaking individuals.
Relevance to Current Work: This study was used as a reference for the completion of the current study as they share a similar methodology.

Caldwell, M. E. (2009). *Development of psychometrically equivalent speech audiometry materials for testing children in Mongolian.* (Unpublished master's thesis). Brigham Young University, Provo, UT.

Objective: The purpose of this study was to develop, digitally record, evaluate, and psychometrically equate a set of Mongolian monosyllabic and bisyllabic word lists to be used for word recognition testing and speech reception thresholds in native Mongolian-speaking children.
Design: Two full word lists and four half-lists were created for measuring word recognition scores in Mongolian children. Each full word list contained 50 words and each half-list contained 25 words.
Study Sample: This study followed the development of the adult speech audiometry materials. A total of eight surveys were sent to professionals in the area of child development in order to determine the appropriateness of the words for school-aged native Mongolian children. All professionals were natives from Mongolia and reported speaking Mongolian on a regular basis. The results from these surveys were used to create the speech audometry meterials. A total of 100 words were ranked according to the rate of listener identification and child appropriateness.
Results: The developed lists and half-lists were found to be homogenous in regard to audibility and psychometric function slope, with average psychometric function slopes at

50% intelligibility of 6.41 %/dB for the male recordings and 5.84 %/dB for the female recordings.
Conclusions: Two lists of 50 monosyllabic words and four half-lists of 25 words each were developed for testing the word recognition abilities of Mongolian children. SRT materials could not be developed due to the structure and limitations of the study.
Relevance to Current Work: The results obtained in the development of word recognition scores materials are similar to those obtained in the present study.

Cancel, C. A. (1957). *Effects of selected phonetic aspects in the transmission of the Spanish language*. (Unpublished doctoral dissertation). Ohio State University, Columbus, OH.

Objective: This study consisted in developing 20 word lists of 50 words each for audiometric testing of Spanish-speaking individuals. Words that ended in a consonat as opposed to a vowel were analyzed in a subsample of 100 words.
Design: Each word was selected from the 1000 most-common bisyllabic words that had the emphasis in the second-to-last syllable, also called Grave words. The words were obtained from the newspapers of 19 Latin-American countries. All words were presented to the subjects in quiet and in noise.
Study Sample: A total of 60 Spanish-American students at The Ohio State University participated in this study.
Results: The author found that words containing three and seven sounds produced low intelligibility scores. A significant difference was found between the two groups (words ending in vowels and words ending in consonants). A higher mean intelligibility value was found in the group of words ending in vowels.
Conclusions: Grave words that contain three and seven sounds can be used in discrimination tests.
Relevance to Current Work: These are the initial findings on Spanish word recognition testing for audiological purposes.

Cancel, C. (1968). Spanish speech audiometry. *International Journal of Audiology, 7*, 206-208.

Summary: In this article, the author discusses the elements necessary to the development of speech reception and discrimination tests for Spanish-speaking individuals. These elements include the following: The transmission of spoken words varies between different languages and is affected by phonetic characteristics. The levels of intensity that are necessary to obtain average thresholds and discrimination measurements also vary between languages. Monosyllabic words in Spanish are not equivalent to monosyllabic words in English used to obtained speech discrimination measurements. Monosyllables in English produce low intelligibility scores. The phonetic length of a word affects its intelligibility.
Relevance to Current Work: This article provides important guidelines to consider for the creation of speech discrimination tests in the Spanish language.

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

Summary: Speech audiometry is the most influential form of modern testing since pure-tone testing. It is a technique in which standardized samples of a given language are presented to individuals using calibrated instrumentation to measure some feature of their hearing ability. Speech audiometry gives us information concerning discomfort or tolerance of a speech stimulus, as well as information on word-recognition abilities or abilities to understand speech. Carhart presents three requirements necessary for successful speech audiometry testing. The first requirement is that the materials used for testing need to be standardized through experimentation. One should not make assumptions regarding the value and reliability of a particular list of syllables, words, phrases, or sentences. The lists can be chosen in advance. However, their effectiveness needs to be tried and tested. A second important principle to consider when testing speech audiometry is that the instrumentation used must be calibrated. The third and last principle is to remember that different tests assess specific hearing abilities.
Relevance to Current work: This article provides guidelines to consider when doing speech audiometry testing.

Carhart, R. (1951). Instruments and materials for speech audiometry. *Acta Oto-Laryngologica, 40*, 313-329.

Summary: As the title suggests, the article outlines basic principles to consider when doing speech audiometry. Speech tests are divided into two categories: threshold tests and phonemic discrimination tests. The two most common techniques when assessing speech tests are the monitored live-voice method and the recorded presentation method. The most clinically flexible of the two is the monitored live-voice technique as it may be adjusted to meet the client's needs. The basic principle of this technique is that the materials used are spoken every time for each person. The person administering the test must always be a trained talker. The talker must keep his or her voice at a uniform level by watching a decibel meter. In contrast, recorded presentations cannot be tailored for each individual. However, they can be standardized more easily. The basic feature of this technique is that the material is prepared phonographically, and it is replayed for each patient during testing. Similar to the live-voice method, monitoring is necessary when recording the material. The test materials can be presented to the patient through headphones or through speakers. Another consideration when testing is that the instrumentation needs to be of high quality and needs to be able to cover a wide range of presentation levels to meet the requirements for clinical work. Also, the instrument used must be calibrated, whether in terms of levels of intensity or in terms of the threshold necessary for speech in normal hearing patients.
Relevance to Current Work: This article provides important information and principles to consider regarding instrumentation in speech audiometry testing.

Carhart, R. (1965). Considerations in the measurement of speech discrimination. *Aeromedical Review, 3*, 1-22.

Summary: Speech discrimination scores refer to the percentage of test items that an individual correctly identifies when these items are presented by ear. The first conflict a clinician faces when doing speech audiometry testing is selecting the appropriate materials for testing. One very important element to consider in choosing a particular discrimination test is the patient's linguistic background. Also, discrimination tests must consist of relatively non-redundant items, as opposed to threshold tests. This is because if the items are duplicated, then it is harder to identify a person's inability to accurately differentiate consonants and vowels. This is the reason why monosyllabic words are used instead of sentences and multisyllabic words. Clinicians must keep in mind that guidelines for assessing auditory pathology and determining site of lesion are different from criteria used to estimate efficacy of hearing in daily life or the potential of a rehabilitative procedure such as surgery or hearing aids.
Relevance to Current Work: This author gives informative data to consider regarding speech discrimination testing.

Cokely, J. A., & Yager, C. R. (1993). Scoring Spanish word-recognition measures. *Ear & Hearing, 14*, 395-400.

Objective: The purpose of this study was to determine Spanish word recognition scores from Spanish-speaking listeners' oral and written responses and to compare the results with those obtained from English-speaking listeners.
Design: The subject's task consisted of writing and repeating words in Spanish. Their oral response was recorded on tape and the written response was scored as repeated correctly and incorrectly by an English-speaking individual that had a total of four years of experience of Spanish.
Study Sample: Ten Spanish-speaking individuals (six men and four women) between the ages of 27 and 45 years old were selected to participate in this study. Eight of the ten participants had air conduction pure-tone thresholds of 10 dB Hl or less at octave frequencies from 250-8000 Hz. The other two individuals has thresholds of 25 db hL at 4000 and 8000.
Results: The overall mean of word recognition scores from the oral responses was 59.8% when scored by a judge with knowledge in Spanish and 60.2% when scored by a judge without knowledge in Spanish. Results on the examination of individual data show that oral and written scores differed by 10 percentage points or more in <1% of the cases when judges with knowledge of Spanish scored the oral responses and in only 3% of the cases when judges without knowledge of Spanish scored the responses. Scores did not differ by more than 16 percentage points in any case. A statistical analysis revealed that the difference between the groups of judges is statistically significant. However, this difference is of little clinical significance.
Conclusions: WRSs derived from oral responses were the same when scored by English-speaking persons with knowledge of Spanish and by English-speaking individuals without knowledge of Spanish. These data suggest that an English-speaking professional is competent to judge the accuracy of Spanish-speaking listeners' oral responses to

Spanish when testing WRS.
Relevance to Current Work: This study provides information regarding English-speaking individuals testing for WRS on Spanish-speaking individuals.

Comstock, C. L., & Martin, F. N. (1984). A children's Spanish word discrimination test for non-Spanish-speaking clinicians. *Ear & Hearing, 5*, 166-170.

Objective: The purpose of this study was to develop a test for Spanish-speaking children that could be used by English-speaking clincians.
Design: A picture pointing test was created. The test contained four lists of 25 bisyllabic words each. The words were recorded in Spanish and English on channels one and two respectively, with the purpose of presenting the Spanish recording to the subjects at the same time the clinicians monitored the accuracy of the responses in English.
Study Sample: The study included 15 native Spanish speakers raised in Texas, all of whom passed a pure-tone screening test at 10 db HL at the frequencies from 250 to4000 Hz. The study also included 20 children ages three to eight whose primary langauge was Spanish and who lived in central Texas.
Results: When testing adults, the four word lists were statistically equivalent. At 32 to 40 db Sl, the list varied by less than 4% in discriminability. When testing children, an analysis showed an overall increase in discrimination scores with increase in age. Several of the words missed by the children appear to be due to their limited vocabulary.
Conclusions: It is concluded that this test is useful for assessing speech discrimination in children as long as vocabulary is taken into account. It also proved effective for the clinicians to administer this test battery to Spanish-speaking children without having to provide instructions verbally.
Relevance to Current Work: This study provides an alternative and effective method of testing that can be used when assessing word discrimination in Spanish-speaking children.

Dana, L. P. (2007). *Handbook of research on ethnic minority entrepreneurship: A co-evolutionary view on resource management.* Cheltenham, Great Britain: Edward Elgar Publishing.

Summary: This book focuses on providing the reader with a better understanding of ethnic minorities and entrepreneurship in the United States. The chapter also addresses topics such as education, economic situation, location, culture, religion, and origin of diverse ethnic groups, including the Hispanic population.
Relevance to Current Work: The chapter helps develop a framework for understanding in more depth different aspects and features of the Hispanic culture.

Diefendorf, A. O. (1983). Speech audiometry with infants and children. *Seminars in Hearing, 4*, 241-252. doi:10.1055/s-0028-1091428

Summary: This article discusses the importance of speech audiometry when testing hearing in children. The author considers this to be an essential piece to accomplish a complete and accurate diagnostic of a child with a hearing impairment. The article also

provides relevant information on speech audiometry including speech discrimination and test batteries and methods used when working with infants and young children. Speech discrimination is the ability to differentiate sounds. The main purpose of speech discrimination testing is to measure an individual's ability to understand speech under different circumstances. When assessing children, there are things the clinician needs to consider in order to get more reliable results. The clinician must consider the child's receptive vocabulary level and the child's ability to understand a task, and they should use reinforcements and reduce memory load during testing as necessary. When assessing a speech recognition threshold in children the test should be performed at a quick pace and the test should also have enough appeal to the child to keep him focused on the task. The words used must be of equal difficulty, highly intelligible, and within the child's vocabulary.

Relevance to Current Work: This article provided guidelines to consider as we assessed children's speech audiometry in order to increase the validity and reliability of our results.

Drabman, R. S., & Lahey, B. B. (1974). Feedback in classroom behavior modification: Effects on the target and her classmates. *Journal of Applied Behavior Analysis*, *7*, 591–598.

Objective: The purpose of this study was to create a behavior-management design for children within the classroom that focused mainly on feedback and no other contingencies.

Design: The target student was observed Mondays through Thursdays by two undergraduate students. The observation focused on the following categories of disruptive behavior: touching, play, noise, out of chair, non-compliance, time off task, vocalization, orienting, and aggression. The child and peers' disruptive behavior, the individual's socioeconomic status, and the positive and negative feedback to the target by her teachers and peers were taken into account.

Study Sample: A 10-year-old girl that was reported to have the most disruptive behavior in her classroom. She had no friends and was constantly teased and ignored by her classmates.

Results: The results from this investigation showed that just through feedback there was a positive and effective change in behavior and that sociometric status of the target was altered by behavioral contingencies. The target student received fewer negative comments from her teacher and more positive comments from her classmates.

Conclusions: Feedback plays an important role in behavior modification. During the study, the target student did not receive any more praise from the teacher than before. Also the feedback delivered from the teacher to the child was straightforward but with little emotion. This shows that possibly feedback alone is what made the student change her behavior.

Relevance to Current Work: The study provided useful information for behavior management when working with children.

Hamid, M. A., & Brookler, K. H. (2006). Speech audiometry. *Ear, Nose, and Throat Journal, 85*, 810-812.

Summary: This article's focus is speech audiometry. The authors state that speech audiometry aids ear specialists in determining an individual's level of hearing dysfunction and aids also in selecting the most appropriate hearing aids for a specific individual. The most common tests used in speech audiometry are word recognition using phonetically balanced words (WRS), speech recognition thresholds (SRT) using spondaic words, most comfortable loudness (MCL) and uncomfortable loudness (UCL). SRT helps determine levels of MCL and UCL. For WRS testing, ASHA requires clinicians to use standard speech lists. The most common lists used are the NU-6 and the CID-W lists. The words lists can be delivered via earphones, insert phones, or sound field using pre-recorded lists instead of using live voices to ensure reliability, consistency, and accuracy. An open set format is typically used; however, a closed set format can be used when working with the geriatric and the pediatric populations. SRT is typically 10 dB higher than the pure-tone average at 0.5 and 2 kHz. If a difference higher than 10 dB exists, the reliability of the test should be questioned. WRS is determined at 40 dB SL. Masking on the non-test ear is used if necessary. The dB SL must be below the UCL level and the MCL level should be used to determine WRS even if it is less than 40 dB.
Relevance to Current Work: This article provided necessary background knowledge on the concepts and mechanisms of testing speech audiometry, specifically concerning SRT and WRS testing.

Harris, R. W., Nielson, W. S., McPherson, D. L., & Skarzynski, H. (2004). Psychometrically equivalent Polish bisyllabic words. *Audiofonologia, 25*, 1-13.

Objective: The purpose of this study was to develop, digitally record, evaluate, and psychometrically equate Polish bisyllabic word lists for use in the measurement of speech recognition.
Design: A total of 70 familiar bysyllabic words with stress on the first syllable were selected. These words were recorded by both male and female talkers. These talkers were native to Poland. The authors calculated the psychometric functions for each of the 70 bisyllabic words. The words were presented to the subjects at 15 intensity levels (-10 to18 dB in 2dB increments).
Study Sample: A total of 26 native Polish participants were participated in this study (seven male & 19 female). Subjects ranged in age from 20 to 29 years. All participants had normal heating and had pure-tone air-conduction thresholds of 15 dB HL at octave and midoctave frequencies (125-8000 Hz) with static acoustic admittance between 0.3 and 1.4 mmhos with peak pressure between -100 and +50 daPa.
Results: The best 25 words out of the total 70 words were selected and digitally adjusted. The threshold of each word was equal to the mean PTA (2.37 dB HL) for all subjects. The mean sople was 10.1%/dB for the male talker words and 9.8%/ dB for the female talker words.
Conclusions: The Polish bisyllabic words recorded by both the male and female talkers are relatively homogenous in regard to audibility and psychometric function slope.

Relevance to Current Work: This study focused on creating homogeneous speech audiometry materials, and the methodology implemented is similar to the current study.

Haskins, H. L. (1949). *A phonetically balanced test of speech discrimination for children.* (Unpublished master's thesis). Northwestern University, Evanston, IL.

Objective: This test was created with the intent of evaluating the speech perception of kindergarten-aged children.
Design: The author developed the PBK lists with words taken from the well-known PB-50 lists for adults. Words were chosen among the 2500 words of highest frequency spoken by preschool children. The tests contained four lists of 50 words each. The author recorded a single randomization of the four lists of the PBK Test. Haskins presented the four PBK lists and one of the PB-50 lists at sequentially higher signal levels to adult subjects. The words were presented at −7.8, −2.8, 2.2, 12.2, and 27.2 dB SL (0 dB SL was determined as the mean threshold for spondaic words). Participants identifird the words in an open-set format.
Study Sample: A total of 22 adult listeners with normal hearing participated in this study.
Results: The author determined that List 2 was "easier" than the other lists at all levels of presentation. Only lists 1, 3, and 4 were considered to be equivalent enough to be used clinically with children.
Conclusion: Three of the original four word lists can be used clinically.
Relevance to Current Work: This study developed materials for testing word recognition in children.

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*, 321-337.

Summary: Speech audiometry is an important tool that has added validity to pure-tone audiometry and various speech tests. This article discusses modifications done to previously recorded auditory tests to create new ones. The intent was to satisfy needs that were not met by the old tests. These tests focused on hearing loss for speech and discrimination loss. The PAL auditory test was modified and the tests W-1, W-2, and W-22 were created. A magnetic tape recording was used, allowing the creation of different word orders of specific tests. All test items of the several versions existing are physically identical. Phonetic balance and familiarity on the W-22 test was more rigidly taken into consideration. This resulted in a more limited vocabulary on the lists. However, the lists appear to be easier and more homogeneous. The intelligibility of the new tests is presented as a function of intensity.
Relevance to Current Work: The study shows an early development of speech audiometry tests materials.

Lawson, G., & Peterson, M. (2011). *Speech audiometry*. San Diego, CA: Plural

Summary: This book gives a broad overview of speech audiometry and masking in clinical protocols. It also provides an outline of both modern and traditional assessment

tools. The aim of this book is to aid clinicians in determining differential diagnosis, assessing auditory processing ability, identifying pseudohypacusis, determining cochlear implant candidacy, predicting hearing aid benefit, and counseling.
Relevance to Current Work: The book presents valuable information about the relationship between pure-tone and speech discrimination testing.

Lopez, S. M., Martin, F. N., & Thibodeau, L. M. (1997). Performance of monolingual and bilingual speakers of English and Spanish on the Synthetic Sentence Identification test. *American Journal of Audiology, 6*, 33-38.

Objective: Adults' performance on the English and Spanish versions of the Synthetic Sentence Identification (SSI) test was compared and evaluated.
Design: The participants completed two requirements. The first one was the standard competing message (SC) and the second one was the competing message mixed with speech noise (SC+N). This was done with the purpose of removing or decreasing the effect of pauses on performance.
Study Sample: The study included 30 participants: seven were bilingual, ten were monolingual Spanish-speakers, and ten were monolingual English-speakers. All participants' hearing was within normal limits.
Results: An ANOVA test revealed that the bilingual and monolingual English speakers' scores were not significantly different. Also no significant differences were found between English and Spanish scores for the bilingual speakers in the SC condition; however, the ANOVA revealed a significant difference between the English and Spanish SSI scores in the SC+N condition.
Conclusions: Bilingual scores in Spanish were significantly better than in English, even when speech noise was used.
Relevance to Current Work: This study used bilingual English-Spanish participants for audiological testing.

Madell, J. (1996). *Speech audiometry for children*. Washington, DC: Gallaudet University.

Summary: This book provides an overview of pediatric speech audiometry. The aim of speech audiometry is to collect as much information as possible about the children's speech perception abilities. The chapter covers the evaluation of auditory perception, guidelines for testing speech audiometry, and different test batteries used when testing speech audiometry in children.
Relevance to Current Work: This source provides relevant information on the purpose of pediatric speech audiometry and procedures for testing speech audiometry in children.

Martin, F. N., & Clark, J. G. (2009). *Introduction to audiology*. Boston, MA: Allyn and Bacon.

Summary: This book gives a comprehensive introduction to audiology. It covers topics such as the physics of sound, the anatomy of the auditory system, the causes and treatment of hearing and balance disorders, and the relevant diagnostic and therapeutic techniques for these disorders. Clinical masking and the proper evaluation of hearing disorders and the treatment avenues available for these disorders are also emphasized in

this book. Instructions are given for obtaining the SRT and recording results including use of masking. The book describes the different types of phonetically balanced word lists used in word recognition testing. The book also covers topics such as the management of auditory processing disorders, the role of the audiologist in vestibular management, and the role of the audiologist in the counseling process.
Relevance to Current Work: This book provides an overview on the history, procedures, and importance of speech audiometry and audiological testing in general.

Martin, F. N., & Hart, D. B. (1978). Measurement of speech thresholds of Spanish-speaking children by non-Spanish-speaking clinicians. *Journal of Speech and Hearing Disorders, 43*, 255-262

Objective: The main purpose of the study was to evaluate the effectiveness of a picture-pointing test to assess speech thresholds of Spanish-speaking children. The test was given by non-Spanish-speaking clinicians.
Design: The study was divided into two experiments. In the first experiment, Spanish and English spondaic words were compared to determine their equivalence in terms of their threshold. In the second experiment the Spanish spondaic words were used in a picture-pointing test format with the purpose of determining the relationship between PTA and ST.
Study Sample: The experiment was divided into two parts. First, 16 Spanish-speaking adults with normal hearing participated. Eight were native English speakers and the remaining eight were native Spanish speakers. Second, eight male and eight female Spanish-speaking children from Texas and with normal hearing participated.
Results: A Spanish word list was compared to an English spondaic word list on the adult group to see if both lists were equivalent. Both lists appeared to be equivalent, reliable, and stable. A test containing these words was administered to the sixteen children. There was a strong correlation between the Spanish SRT and the PTA. The mean difference between the Spanish ST and the pure-tone average was 5dB and the greatest difference was 10 dB. These results show to be equivalent to those of other studies using children.
Conclusions: There was a good correlation between the SRT and PTA when testing the 16 Spanish-speaking children. The test studied in this investigation appeared to be practical, fast, and reliable.
Relevance to Current Work: This is an example of an earlier study that dealt with creating audiological testing materials for Spanish-speaking children, based on pre-established materials for adults.

Mendel, L. L. (2008). Current considerations in pediatric speech audiometry. *International Journal of Audiology*, *47*, 546-553.

Summary: This article highlights the current considerations in pediatric speech perception assessment. It focuses on specific test principles that must be addressed when evaluating speech perception performance in children. Not all methods used for testing speech perception are appropriate for children. Speech perception test materials must be designed so that they are appropriate for different populations of children. It is important to make sure that pediatric speech perception tests are developed with proper attention to

validity and reliability issues. Also, appropriate rules of test construction and standardization must be followed. In addition, clinicians need to use carefully controlled test methodologies to obtain accurate assessments. This article highlights that some assessments have more sensitivity and standardization information available to them than others. It recommends that tests that follow principles of psychometric theory in their development should be used in assessing children' s speech perception performance. Besides assessing performance, a test battery can help both to monitor a child's progress over time and to determine if any improvements need to be made with amplification and with intervention efforts. It is important that an ongoing speech perception assessment takes place in order for a child's progress to be monitored over time and to determine if improvements need to be made with amplification and intervention efforts. Information regarding the prognosis of speech, language, reading, and cognitive abilities of children as well as steps that need to be taken in the intervention process can be obtained from pediatric speech perception assessments

Relevance to Current Work: This article provides information, variables, and principles to take into account when assessing speech audiometry in children.

Meyer, T. A., & Pisoni, D. B. (1999). Some computational analyses of the PBK test: Effects of frequency and lexical density on spoken word recognition. *Ear & Hearing, 20*, 363-371.

Objective: The aim of this study is to determine if the lexical properties of the different PBK lists could explain any differences between three equivalent lists and the fourth PBK list that has not been used for clinical testing.

Design: A computerized database was used to measure the density and occurrence of word frequency and lexical neighborhood for all the words contained in the four PBK test lists as well as the words from a single PB-50 word list developed by Egan in 1948.

Study Sample: The four PBK lists and one of the PB-50 lists were presented to 22 adult listeners with normal hearing. The words were presented at −7.8, −2.8, 2.2, 12.2, and 27.2 dB SL. Listeners were asked to identify the words in an open-set format.

Results: The words in list two (the "easy" list) were higher in frequency than the words in the three equivalent lists. Furthermore, the lexical neighborhoods of the words on list two contained fewer words that were phonetically alike than the neighborhoods of the words on the other three equivalent lists.

Conclusions: The results obtained in this study provide additional support for the proposed idea that spoken words are recognized relationally in the context of other words that are phonetically similar in our vocabulary. Implications of using open-set word recognition tests with children with hearing impairments are discussed with regard to the specific vocabulary and information processing demands of the PBK test.

Relevance to Current Work: This study was focused on analyzing materials to be used in testing speech perception skills in children.

Neumann, K., Baumeister, N., Baumann, U., Sick, U., Euler, H. A., & Weißgerber, T. (2012). Speech audiometry in quiet with the Oldenburg Sentence Test for Children. *International Journal of Audiology*, *51*, 157-163.

Objective: The purpose of the study was the validation of the Oldenburg sentence test for children (OlKiSa) in quiet with a large sample of normal-hearing children between the ages of four and ten.
Study Sample: A total of 224 children between the ages of four and ten years old participated in this study. All participants had normal hearing abilities.
Results: The discrimination function SRT values ranged from 6.4 to 10.7 %/dB, which were steeper than those of the commonly used German single word tests.
Conclusions: The OlKiSa is a valid audiometric test to quantify speech perception in quiet in children.
Relevance to Current Work: The study analized relevant clinical and audiological materials to be used in testing speech audiometry in children. This study also tested children within the target age of the current study.

Peterson, G. E., & Lehiste, I. (1962). Revised CNC lists for auditory tests. *Journal of Speech and Hearing Disorders, 27*, 62-70

Summary: The topic of word frequency for intelligibility testing is reviewed by the authors. How frequently a word is used by an untrained listener is an important variable in their response, which must be considered. When working with untrained listeners, substantial effects have been observed related to word frequency. The authors revised the CNC list created in 1959. Uncommon words, literary words, and proper names were eliminated from the old lists. The authors present ten revised lists containing 50 words each, as well as the frequencies of the words contained in the lists.
Relevance to Current Work: The issue of phonetic balance and word frequency is addressed in this article concerning audiological/intelligibility testing.

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

Summary: Speech audiometry is a vital element in audiological testing. Testing children's and adults' hearing is incomplete without assessing speech. Despite the difficulty for standardization and specificity, speech audiometry is important for several reasons: first, speech signals represent auditory stimulation that happens in our daily life. Second, understanding speech is the basis for communication in society. Third, speech audiometry has high face validity as the patients are usually familiarized with the words used in speech audiometry. There are several components that test materials for speech audiometry should have: familiarity, phonetic dissimilarity, a representative sample of English speech sounds, and homogeneity with respect to audibility. The authors stated in this article that most speech audiometry materials used with minorities are in English. Because SRT is intended to test intelligibility for speech and not vocabulary or intelligence, familiarity is the most important aspect when selecting stimuli for SRT testing. Several studies done in the past have used digits instead of monosyllabic,

bisyllabic, or polysyllabic words in speech audiometry. Digit stimuli have been recommended for people with are unable to repeat back or imitate words, as is the case for some individuals with aphasia or dementia, for young children, etc. Also clinicians may be able to tests clients with limited English proficiency using digit stimuli to measure SRT. It is likely that measuring hearing thresholds in individuals that have limited English proficiency would be more effective if the digit pairs used are valid and reliable.
Relevance to Current Work: The article gives evidence of the need for speech audiometry materials for individuals that have limited English proficiency.

Schlauch, R. S., Anderson, E. S., & Micheyl, C. (2014). A demonstration of improved precision of word recognition scores. *Journal of Speech, Language, and Hearing Research, 57*, 543-555. doi:10.1044/2014_JSLHR-H-13-0017

Objective: The goal of this study was to document, in well-controlled conditions, how much word recognition scores (WRS) improved individual's ability to detect a small change in hearing as the number of items that contributed to the WRS was increased.
Design: Two sets of analyses were conducted; the first analysis was a group comparison of dimensionless effect sizes of the pure-tone threshold shift to that of the various permutations of the word recognition test for identifying the change in hearing. In the second analysis, the authors examined individual participants' results to establish true positive rates and false positive rates for the pure-tone and speech tests for identifying a change in hearing on retest.
Study Sample: Twenty-four adults (14 female and 10 male) between 18 and 49 years of age served as participants. All of the participants had pure-tone thresholds in the test ear that were less than or equal to 20 dB HL for audiometric frequencies between 0.25 and 4.0 kHz.
Results: Pure-tone thresholds and WRSs were measured in three levels of speech-shaped noise for the intensity levels of 50, 52, and 54 dB HL for 24 listeners with normal hearing. WRSs were obtained for half-lists and full lists presented at 48 dB HL. A resizing procedure was used to derive dimensionless effect sizes for identifying a change in hearing using the data. This allowed for direct comparison of the magnitude of shifts in WRS (%) and in the average pure-tone threshold (dB), which provided a context for interpreting the WRS.
Conclusion: The results of this study demonstrated convincingly that increasing the number of items that contributed to a WRS significantly increased the test's ability to identify a change in hearing.
Relevance to Current Work: This article provides relevant information regarding the usefulness of WRS testing.

Shi, L. F., & Sanchez, D. (2010). Spanish/English bilingual listeners on clinical word recognition tests: What to expect and how to predict. *Journal of Speech, Language, and Hearing Research, 53*, 1096-1110. doi:10.1044/1092-4388(2010/09-0199)

Objective: The purpose of this study was to provide a framework for analyzing which languages to use during speech perception testing when working with English/Spanish bilingual individuals.
Design: English-Spanish listeners were evaluated on both English and Spanish WRS tests in quiet and in speech-spectrum noise.
Study Sample: The participants were 30 individuals of varying ages and with normal hearing. All participants were given word recognitions tests in English and Spanish in quite and in speech-spectrum noise. The participants' age of language acquisition, length of immersion, daily language use, self-rated language proficiency, and language of dominance varied.
Results: There was no correlation between any conditions and the performance on the English and Spanish tests. However, English word recognition testing was significantly correlated with age of English acquisition. A logistic regression analysis showed that the English acquisition was a good predictor of how the listeners were going to do on both tests in quiet and at a +6 dB signal-to-noise ratio (SNR).
Conclusion: English/Spanish bilingual individuals whose predominant language is Spanish and who learned English at the age of 10 or older are more likely to perform better during WRS testing. Individuals who learned English between the ages of 7 and 10 would need to be evaluated in both languages.
Relevance to Current Work: This study provides information regarding WRS testing on bilingual English/Spanish-speaking children.

Tillman, T., & Carhart, R. (1966). *An expanded test for speech discrimination utilizing CNC monosyllabic words: Northwestern University Auditory Test No. 6*. Brooks Air Force Base, TX: US Air Force School of Aerospace Medicine.

Objective: The aim of this study was to create four lists of 50 consonant-nucleus-consonant monosyllabic words following the same scheme employed in the earlier development of the N.U. Test No. 4. This new test is called the N.U. Auditory Test No. 6.
Design: The first step in creating the N.U. Auditory Test No. 6 was to make a table indicating the number of times each phoneme needed to be used in a specific list if one phonemic distribution needed to be preserved. The second step was selecting four mutually exclusive groups that contained 50 words. Each word needed to be as close as possible to the distribution of phonemes in the table created in step one. The third step was to randomize each of the four lists four times and then record them on magnetic tape. Each of the 36 participants were examined twice. The four lists were administered to each subject six times starting at a presentation level of 4 dB below the individual's SRT. The intensity level was progressively increased in the subsequent presentations of the words.
Study Sample: A total of 24 individuals with normal hearing were recruited for this study. They ranged in age from 19 to 28 years. The experimental group consisted of 12 additional participants who had experienced progressive hearing loss during adulthood.

Results: Differences between normal hearing participants and participants with a hearing loss were found. The normal hearing group's psychometric function slope was 5.6%/dB. In contrast, the hearing-impaired group's slope was about 3.4%/dB. There was greater inter-subject variability in performance in the hearing impaired group than in the normal hearing group. The variability was at and above the 8-dB sensation level.
Conclusions: A good test-retest reliability that correlates with N.U. Auditory Test No. 4 was found. However, the authors found that there was a tendency for discrimination scores to slightly improve after re-testing.
Relevance to Current Work: This study provided information on the development of word recognition testing materials.

Trimmis, N., Papadeas, E., & Papadas, T. (2008). A Modern Greek word recognition score test designed for school-aged children. *Mediterranean Journal of Otology*, *4*, 1-8.

Objective: The purpose of this study was to develop a Modern Greek word recognition score test designed for school-aged children.
Design: This test was designed in an open-set format. Two lists of 50 bisyllabic words each were developed. Bisyllabic words were chosen as stimuli due to the limited number of monosyllabic words in Modern Greek. The lists are phonemically balanced. The words in each list are suitable for school-aged children ages six through 12 years. Both lists were tested in children with normal hearing, children with a sensorineural hearing loss, and children with a conductive hearing loss.
Study sample: Seventy normal-hearing children were tested, 45 boys and 45 girls. Twenty of the seventy had a conductive hearing loss, 5 boys and 5 girls and ten children had a sensorineural hearing loss children 5 boys and 5 girls. All participants ranged in age from six to 12 years and were native Modern Greek speakers. None of the subjects had a history of a speech or language disorder, learning disability, or other cognitive disorder. All normal-hearing children had pure tone thresholds of ≤ 15 dB HL at all octave frequencies ranging from 250 Hz to 8000 Hz.
Results: An average slope of 5.08% in the first and 5.24% in the second list was obtained when testing normal-hearing children. The effect of conductive hearing loss is largely one of reduced sensitivity. In sensorineural disorders, the psychometric function tends to be flatter and the maximum score is reduced. The results are reasonably equivalent for the number of subjects tested since none of the mean scores differ from the overall mean score by more than 8%.
Conclusion: The results of the lists tested in children with normal-hearing and different types of hearing loss revealed that the test is a useful tool for the audiological evaluation of Modern Greek–speaking children.
Relevance to Current Work: The data obtained in the development of word recognition scores materials are similar to those obtained in the present study.

U.S. Census Bureau. (2011). *The Hispanic population: 2010.* Retrieved from http://www.census.gov/prod/cen2010/briefs/c2010br-04.pdf

Summary: This census report analyzes population and housing data collected in 2010. It also gives an overview of the Hispanic or Latino population in the United States, and of ethnicity concepts and definitions used in the 2010 Census.
Relevance to Current Work: The census report provides important information about the rapid growth of the Hispanic community in the United States and was therefore of importance to this project.

Von Hapsburg, D., & Pena, E. D. (2002). Understanding bilingualism and its impact on speech audiometry. *Journal of Speech, Language, and Hearing Research, 45*, 202-213. doi:10.1044/1092-4388(2002/015)

Summary: The tutorial reviews previous auditory research regarding monolingual and bilingual English/Spanish speakers. It provided definitions of bilingualism and covers different methods for describing diverse research participants linguistically. It also discusses how bilingualism can affect the outcomes in auditory research. The authors recommended that researchers interested in the topic expand their knowledge regarding bilingualism before they design their projects. Other important points discussed include that early bilinguals and monolinguals' performances in speech-perception tasks are similar and that late bilinguals are different from monolinguals and early bilinguals in some speech-perception tasks. Also, when the focus is on monolingual function, bilinguals should not substitute monolingual speakers.
Relevance to Current Work: This tutorial creates awareness regarding the implications of using bilingual speakers in auditory research.

Walsh, T. E., & Goodman, A. (1955). Speech discrimination in central auditory lesions. *Laryngoscope, 65*, 1-8. doi:10.1288/00005537-195501000-00001

Summary: Studies have suggested that the relation between the hearing loss threshold of an individual and their speech discrimination performance can help in diagnosing between cochlear lesions and the auditory tract central to the end organ. Intelligibility for speech increased with a tumor on the VIII nerve and decreased with cochlear hydrops. It was suggested that distinguishing between threshold losses for speech and the discrimination score might be the differential diagnosis key between end organ lesions and central organ lesions.
Relevance to Current Work: The authors addressed the clinical purpose of speech discrimination testing.

Weisleder, P., & Hodgson, W. R. (1989). Evaluation of four Spanish word-recognition-ability lists. *Ear & Hearing, 10*, 387-392.

Objective: When testing WRS, the indiviual's linguistic background must be taken into account. Even though the Spanish language is the second most common language in the

United States, there hasn't been enough research and efforts in the development of a word recognition test in Spanish.
Design: This study took an existing Spanish language test and evaluated it. The test material included 50 bisyllabic tetraphonemic Spanish words. The words on the lists were evaluated according to the equivalence between lists, word difficulty, intelligibility of the talker, and slope of the performance/intensity function.
Study Sample: Sixteen native Spanish speakers with normal hearing participated in this study (10 males and six females). The participants' nationalities varied: nine were from Mexico, two from Panama, two from Venezuela, one from Spain, one from Honduras, and one from Colombia. The participants were asked to listen to a list of words and to repeat them back. An individual score for each participant was kept. In order to obtain a percentage score the number of words that the participants repeated correctly was multiplied by two.
Results: The lowest mean intelligibility scores came from list three. A statistical analysis demonstrated that there is a significant difference between lists three and the other lists at the 0.05 level.
Conclusion: On average, the Mexican subjects obtained better scores than participants from other nationalities at low presentation levels. The function slope of the PI-PB was comparable to the English lists. Most words missed were words that contained the /s/s sound and that do not change in meaning even when omitting the /s/ sound in the final position.
Relevance to Current Work: This study exemplifies the evaluation of speech audiometry in Spanish and gives us an idea of the impact of dialect in audiology testing.

Weiss, D., & Dempsey, J. J. (2008). Performance of bilingual speakers on the English and Spanish versions of the Hearing in Noise Test (HINT). *Journal of the American Academy of Audiology*, *19*, 5-17.

Objective: The aim of this study was to compare the performance of bilingual participants on both the English and Spanish verstions of the Hearing in Noise Test.
Design: Four test conditions of the HINT battery were used: (1) E-HINT-Q—English sentences presented in quiet at 0° azimuth, (2) S-HINT-Q—Spanish sentences presented in quiet at 0° azimuth, (3) E-HINT-N—English sentences presented with background noise, and (4) S-HINT-N—Spanish sentences presented with background noise. Participants were instructed to repeat the sentences they heard. Instructions were presented in the language of test administration.
Study Sample: The subjects were divided into two categories, early bilinguals (EB) and late bilinguals (LB), based on when they acquired their second language. To be able to participate in the EB and LB groups, subjects had to begin learning English before the age of seven years or after the age of 11 years. All participants' first language (L1) was Spanish and their second language (L2) was English, and all participants were required to demonstrate normal hearing sensitivity. The EB group was comprised of 18 participants and the LB group of seven participants.
Results: Both the EB and LB groups performed significantly better on the Spanish HINT test than on the English HINT test (EB, $t = -6.942$, $p = <0.0001$; LB, $t = -4.627$, $p = 0.004$). Early bilinguals had an average threshold score of 15.1 dB on the S-HINT

compared to a threshold of 18.0 dB on the E-HINT. Late bilinguals had an average threshold score of 11.0 dB on the S-HINT compared to a threshold of 15.2 dB on the E-HINT.
Conclusion: The results obtained in this study indicated that all bilingual participants scored better on the Spanish HINT than the English HINT for presentations in quiet and in noise, and that those in the LB group scored significantly higher.
Relevance to Current Work: This study also used Spanish-speaking children to conduct audiometric testing. Bilingual participants scored better when the test was presented to them in Spanish.

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

Objective: The main purpose of this study was to psychometrically equate thresholds of the CID W-1 spondaic words that peaked at 0 vu. According to some studies these are not equivalent.
Design: This study was divided into two experiments. Thirty-six spondaic words were recorded by both male and female talkers. Psychometric slope functions were established for the 36 words. The words spoken by the female talker were digitally adjusted in order to produce an equal intelligibility threshold to those obtained on the male taker data. In the second experiment, psychometric functions were established for the adjusted 36 spondaic words.
Study Sample: Twenty individuals with normal hearing participated in each experiment. Individuals' ages ranged from 19 to 30 years.
Results: It was found that the mean thresholds were the same for both experiments. However, the standard deviation of experiment two was found to be significantly smaller than the standard deviation of experiment one.
Conclusion: The authors concluded that the materials created in this study are psychometrically equivalent and should be used for test validity.
Relevance to Current Work: This study deals with creating homogeneity in speech audiometry materials for audiological testing.

Appendix B

Informed Consent

RESEARCH CONSENT FORM

Introduction

This research study is being conducted by Richard Harris, PhD at Brigham Young University; Brenda Peterson, BS Communication Disorders, Communication Disorders graduate student at BYU; and Jessica Graham, BS Communication Disorders, Communication Disorders graduate student at BYU to evaluate a word list recorded using improved digital techniques. You were invited to participate because you have a child who speaks Spanish and is between the ages of four and eight years.

Procedures

If you agree to participate in this research study, the following will occur:

- You will receive a hearing screening where you will hear beeps and indicate whether or not you heard them.
- You will receive an articulation screening where you will name pictures of objects.
- You will listen to Spanish words and repeat the words you hear.
- The total time commitment will be approximately 60 minutes divided into two 30-minute sessions.
- This will take in the hearing lab, room 110 of the Taylor Building on BYU Campus.

Risks/Discomforts

There are no known risks associated with this study. The researchers will be present at all times to make sure that your child is not experiencing any problems in the study conditions. If the child indicates in any way that he/she does not want to participate by crying or another behavior, we will stop immediately.

Benefits

The primary benefit to your child is finding out whether he/she has normal hearing or not throughout the course of the study. There may be benefits to society in general in that this study may result in more effective treatment methods for Spanish-speaking children participating in hearing exams.

Confidentiality

Your child's participation will be confidential. The data will be stored in locked file cabinets within locked labs within the BYU Speech & Language Clinic. Only the researchers will have access to the data. The names will be removed from research materials. Neither your or your child's name will ever be used in association with this research.

Information will be kept for three years after the study is completed. The files will remain in a locked cabinet only accessible by the researcher. Internet data will be saved as a Microsoft Excel document with no subject identifiers. Participants will be listed by number with no names or identifying referents.

Compensation

You will receive $20 for your participation; compensation will not be prorated. Your child will receive a free hearing exam and you will receive the results. Your child will also receive a sticker or other small prize after each session.

Participation

Participation in this research study is voluntary. You are free to decline to have your child participate in this research study. You may withdraw your child's participation at any point without losing the compensation.

Questions about the Research
Please direct any further questions about the study to Richard Harris at (801) 422-6460 Richard_harris@byu.edu. You may also contact Brenda Peterson or Jessica Graham at (385) 414-5022 or (949) 355-3578, bren.krog@gmail.com or jessicaleerandle@gmail.com.

Questions about Your Rights as Research Participants
If you have questions regarding your rights as a research participant contact IRB Administrator at (801) 422-1461; A-285 ASB, Brigham Young University, Provo, UT 84602; irb@byu.edu.

Statement of Consent
I have read, understood, and received a copy of the above consent and desire of my own free will to participate in this study.

Parent's Name (Printed): ____________________ Signature: _________________________ Date: ______________

Appendix C

Bisyllabic Word Definitions

Spanish	Part of Speech	English[b]	Dictionary[a]
abrir	verb	to open	1,2
algo	pronoun	something, anything	1
alto	adjective	tall	1,3
amor	noun	love	1
aqui	adverb	here	2
arma	noun	weapon	1
autor	noun	author	1
ayer	adverb	yesterday	1,2
azul	adjective	blue	2,3
bajo	adjective	short	1,3
banco	noun	bench, bank	1,2
bebé	noun	baby	1,2,3
bien	adjective	well	1
blanco	adjective	white	3
boca	noun	mouth	1,2
cada	adjective	each, every	1
café	noun	brown, coffee shop	1,2
calle	noun	street	1,2
calor	noun	heat, hot	1,2
cama	noun	Bed	1,3
cara	noun	face, look, side	1,2
centro	noun	center, downtown	1
cerca	adverb	near, close	1,2
cinco	adjective	five	2,3
claro	adjective	bright, light, pale, fair	1
coche	noun	car	1,2,3
color	noun	color	1,2,3
comen	verb	to eat	1
contar	verb	to count, to have	1,2

correr	verb	to run	1,2
cuadro	noun	painting, picture	1
cuatro	adjective	four	2,3
cuenta	noun	calculation, count, bill	1
cuento	noun	story, tale	1,2
cuidar	verb	to take care for	2
curso	noun	course	1
débil	adjective	weak, soft	1,2
dedo	noun	finger	1,3
dentro	adverb	inside	1,2,3
después	adverb	after, later, then afterwards	1,2
día	noun	day	1,2
doble	adjective	double	1,2
doce	adjective	twelve	2,3
doctor	noun	physician, doctor	1
dolor	noun	pain	2
duro	adjective	hard, tough	1,2
entre	preposition	between, among	1
error	noun	mistake, error	1
este	adjective	east, eastern	2
falta	noun	lack of, absence, offence	1,2
fecha	noun	date	1,2
feliz	noun	happy	1,2,3
firma	noun	signature	1
frío	adjective	cold	1,2,3
gente	noun/ adjective	people, kind, good	1
golpe	noun	knock	1,2
grande	adjective	big	1,2
guerra	noun	war	1,2
habla	noun	speech	1
hacer	verb	to do	1,2
hambre	noun	Hunger	1,2
hija	noun	daughter	1
hijo	noun	son	1
hoja	noun	sheet, leaf	1,2

hombre	noun	man	1
hora	noun	time	1,2
igual	adjective	equal	1,2
jamás	adverb	never	1
juegan	verb	to play	1
juego	noun	game	1,2
junto	adjective	together	1
leche	noun	milk	1,2
lejos	adverb	far	1,2
letra	noun	letter	1,2
libro	noun	book	1,2
lista	noun	list	1
listo	adjective	ready, clever, bright, smart	1,2
llenar	verb	to fill, to fill up	1
lleno	adjective	full	1,2,3
lleva	verb	to take	1
llora	verb	to cry	1
loco	adjective	crazy, mad, insane	1
lograr	verb	accomplish, achieve, attain	1
lucha	fight	fight, wrestling	1
lugar	noun	place	1,2
madre	noun	mother	1,2.3
mejor	adjective/ adverb	better, best	1,2
menos	preposition	less	1,2
mente	noun	mind	1
mesa	noun	table	1,2,3
metro	noun	meter, subway	1,2
mezclar	verb	to mix	1,3
moda	noun	fashion	1,2
mucho	adjective/ adverb/ pronoun	a lot, much, many	1,2
muerto	adejctive	dead	1
mujer	noun	woman	1
niña	noun	girl	1,2,3
niño	noun	boy	1,2,3
noche	noun	night	1,2,3
nombre	noun	name	1,2

normal	adjective	normal	1
norte	adjective	north, northern	2
nota	noun	note	1,2
nueve	adjective	nine	2,3
nunca	adverb	never	1
ocho	adjective	eight	2,3
oír	verb	to hear	1,2
ojo	noun	eye	1,2,3
once	adjective	eleven	2,3
oro	noun	gold	1,2
país	noun	country	1,2
papel	noun	paper	1,2
para	preposition	for, in order to, to	2
pared	noun	wall	1,2
perro	noun	dog	1,2,3
planta	noun	plant, first floor	1,2
plata	noun	silver	2
pobre	adjective	poor	2
poder	verb	power, to manage, can	1,2
poner	verb	to put on	1,2
pronto	adverb/adjective	soon, early, sharp, ready	1,2
pueblo	noun	town, village, people	1,2
quemar	verb	to burn	1,3
querer	verb	to want, to love	1,2,3
quince	adjective	fifteen	2,3
razón	noun	reason	1
regla	noun	rule, ruler	1,2
río	noun	river	1
rojo	adjective	red	3

[a]Dictionary: 1= Refer to Larousse, 2011; 2=Refer to Everest, 2013; 3= Refer to Root, 2007
[b]Refer to Oxford, 2015