Considerations in Speech Recognition Testing of Bilingual and Spanish-Speaking Patients Part I: Older Children and Adults

Laura Gaeta, BHS, BA
University of Oklahoma Health Sciences Center, Oklahoma City, OK

Andrew B. John, PhD
University of Oklahoma Health Sciences Center, Oklahoma City, OK

The rapid growth of the Spanish-speaking population of the United States presents challenges for all healthcare providers to develop linguistically- and culturally- appropriate best practices. A significant need for all audiologists is language-appropriate stimuli for speech recognition testing. Unfortunately, few well-validated tests exist for this purpose. We review the timeline of development of Spanish-language speech recognition test materials and address issues facing the audiologist in evaluating accurately the hearing abilities of both older children and adults who use Spanish as their primary or only language of communication.

Introduction

Hearing loss is the third most prevalent physical condition following arthritis and heart disease (Collins, 1997). The World Health Organization (2014) reports that over 328 million adults globally have a hearing loss of 40 dB HL or greater in their better ear. Lin and colleagues (2011) estimate the prevalence of unilateral or bilateral hearing loss greater than 25 dB HL in the United States as about 20% of Americans over the age of 12, or about 48 million people. These statistics have a significant impact on the U.S. Hispanic population given the high number of health disparities observed in this population (Centers for Disease Control, 2014). This population comprises 50.5 million or 16.3% of the national population, making it the largest American ethnic minority group (U.S. Census Bureau, 2011). Though there are non-Hispanics who speak Spanish, most of the people who speak Spanish are Hispanic (Ortman & Shin, 2011). However, Ortman and Shin noted that the number of English-speaking Hispanics would soon surpass Spanish-speaking Hispanics. The Hispanic-American population is projected to rise to 132.8 million by the year 2050. Despite this growing number, Spanish-speaking children and adults in the United States have limited access to healthcare, leading to health disparities in part from lack of access to language-appropriate care.

Audiology is one of many disciplines that must consider changes in diagnostic and intervention practices to account for population changes; indeed, given the primacy of hearing and listening ability in verbal communication, changes in language and culture in the patient population are of particular interest to the audiologist among healthcare providers. In this and the accompanying paper, we discuss important factors in speech-recognition testing for Spanish-English bilingual and primarily-Spanish-speaking patients. In this Part I manuscript, considerations for older children and adults are discussed; the Part II manuscript focuses on factors of concern for younger children.

Audiology and the Spanish-Speaking Population

Hispanic Americans encounter significant social and economic barriers that can decrease the likelihood of receiving timely and appropriate health care. Escarce and Kapur (2006) note that this population's access to health care is affected by a degree of acculturation (adopting or modifying the behaviors and belief systems of another culture), language, and immigration status as more than 40% of Hispanic individuals living in the United States were born in another country. A shortage of Hispanic physicians also contributes to barriers to health care access (Association of American Medical Colleges, 2010). Saha, Taggart, Komaromy, and Bindman (2000) have reported that 40% of Hispanic patients consider a physician's knowledge of Spanish when choosing a provider.

The field of audiology is not as culturally diverse as other health professions, such as physicians and physical therapists. According to a 2008 American Speech-Language-Hearing Association (ASHA) survey, 95.4% of the audiologists surveyed were Caucasian and less than 2% identified themselves as Hispanic or Latino. A 2008 survey of physicians revealed that 75% of physicians over the age of 40 identified themselves as Caucasian and 66% of those younger than 40 years old (Boukus, Cassil, & O'Malley, 2009). A 2013 survey from The U.S. Health Workforce Chartbook (U.S. Department of Health and Human Services, 2013) reported a similar distribution, reporting 79.9% of physical therapists who identified themselves as Caucasian. Whatever the ethnic composition of the healthcare provider population, the projected increase of the Hispanic population in the United States will require all health professionals, including audiologists, to be culturally sensitive and diverse to meet these growing demands.

Effective communication is a key part of an audiologist's role as a health care provider. When working with Spanish-speaking patients, Morrison (2008) suggested repetition of important information to avoid miscommunication between the health care professional and the Hispanic patient. Interpreters can be utilized to minimize misunderstandings. An interpreter relays information from the audiologist to the patient in a manner easily understood by the patient. However, interpreters may also present barriers of their own, including availability, cost, linguistic and regional differences, and knowledge of audiology vocabulary (Talamantes, Lindeman, & Mouton, 2001). It follows then that audiologists and other health care professionals need to be acutely aware of the cultural differences in the Hispanic population to make informed decisions regarding the need for an interpreter, as well as the linguistic and cultural background of the interpreter that is most appropriate for a given patient.

Bilingualism

The growing Hispanic population in the United States has led to an increase of the number of bilingual Americans. Bilingualism is the ability to use two languages. In 1933, Bloomfield defined bilingualism as "native-like control of two or more languages" (Baker, 2011). The National Association for Language Development in the Curriculum (NALDIC, 2009), however, revises that definition to include varying degrees of proficiency and communication. For example, a person may identify as bilingual but may only be able to communicate orally in one language. Likewise, a person may have a proficiency in reading in two languages, but may be unable to converse orally in one of the languages. The separation of these two abilities draws from the four language domains: listening, speaking, reading, and writing. These four abilities can be further categorized as receptive and expressive language skills, affecting degree of language proficiency.

Proficiency in two languages also is reflected in the dominance of the languages. A dominant bilingual is a person who is dominant in one language with the less dominant language referred to as the subordinate language. Chin and Wigglesworth (2007), however, argued that dominance may not be applicable to all four domains.

Degrees of proficiency in receptive and productive language skills have led to categorization of bilingual speakers into groups, such as incipient bilinguals and balanced bilingual speakers (Baker, 2011). The term "incipient bilingual" is used to describe a person with minimal competence of a second language, such as a tourist who learns a few survival phrases. Baker warns this inclusion may be perceived as too exclusive because almost every adult in the world has knowledge of a few words in another language. A balanced bilingual is someone who is essentially equally fluent in two languages. However, true equal fluency is rare as most bilingual speakers use each language for different situations, such as at home or at work. In addition to the groups introduced by Baker, in 1994, Valdes and Figueroa (as cited in von Hapsburg & Peña, 2002) included elective and circumstantial bilinguals as other groups. Elective bilinguals are people who have chosen to learn a second language, but may not necessarily use that language everyday (e.g., tourists, study abroad). Circumstantial bilinguals are those who are required to use the language every day, requiring them to learn a second language in order to communicate (e.g., immigrants). Valdes and Figueroa also stated that bilingualism is based on a "situational continuum," as exposure and dominance varies with the situation. Given its fluidity, it is possible for a bilingual person to be considered for any of the above categories during his or her life. Soares and Grosjean (1984) explored this continuum in their study to determine how bilingual speakers on both ends utilize the lexicons for both languages for a word recognition task. The researchers found that, depending on where the person falls on the continuum, he or she may function like a monolingual or as a bilingual. However, a person who functions as a bilingual will use one language more than the other. Shi (2014a) also noted from Weiss and Dempsey's 2008 study that even though a person may speak Spanish as a native language, it is possible that English has now become the dominant language due to the age of acquisition and increased use of the second language.

Disuse of one language may lead to eventual loss of competence in that language. This is called passive bilingualism. Chin and Wigglesworth (2007) explained that it is common for a bilingual person to understand a language but not be able to speak the same language, especially after undergoing a shift in languages. Passive bilingualism is usually seen in the children or grandchildren of immigrants who have gradually replaced the primary language with a second language based on their community and education.

Another factor influencing the nature of an individual's bilingualism is the age of language acquisition. Tabors' 1997 study (as cited in Goodman, 2007) defined two types of language acquisition: simultaneous and sequential. Simultaneous bilingualism occurs when a child is exposed to two languages from an early age, whereas sequential bilingualism occurs when a child learns the second language (L2) after the first language (L1) is partially established. Typically, children in the United States develop the first language at home before learning a second language at school. Most bilingual children in the United States, therefore, are considered sequential bilinguals (Bedore & Peña, 2008). A person is considered an early bilingual if both the L1 and L2 have been mastered similarly before the age of six; a late bilingual is someone who mastered the L2 after the age of 12 (Knapp & Seidlhofer, 2009).

Speech Audiometry

Communication is the basis for interaction, and clear speech is critical to understanding what we hear. Although pure-tone audiometry provides information of a patient's hearing status, it does not assess a person's ability to understand and hear the sounds used in everyday communication. Speech audiometry uses stimuli, such as words or sentences, often in the presence of noise or other simulated distortion, as a presumably more ecologically valid assessment of a patient's hearing (Gelfand, 2009).

Speech recognition can be assessed with words or sentences presented either via recorded material or monitored live voice. Generally speaking, sentences are considered to be a more realistic simulation of everyday communication, having high face validity, but may place additional cognitive demands on the listener. These cognitive demands are the result of repeating multiple words instead of a single word, which relies on working memory. The demands can be magnified when the sentence consists of words that are not meaningful or does not follow syntax rules (McArdle, Wilson, & Burks, 2005). Words minimize the cognitive demands placed on working memory and are the most popular stimuli, but are not a good representation of every day speech.

Testing using recorded materials tends to result in better control over the intensity and quality of the speech material, whereas monitored live voice may be needed for patients who need extra time to respond. Mendel and Owen (2011) determined the test administration times for monitored live voice and recorded word recognition lists. Mendel and Owen concluded no statistically significant differences in administration time between the two methods. Examples of audibility measures of speech include the Northwestern University Auditory List 6 (NU-6) (Tillman & Carhart, 1966), Auditory Test W-22 (Hirsh, Davis, Silverman, Reynolds, Eldert,, & Benson, 1952), and Phonetically Balanced (PB-50) lists (Egan, 1948).

Speech recognition tests also can be performed in the presence of noise. Speech testing with noise was first used in the 1960s as a way to determine the amount of distortion (McArdle, n.d.). Distortion is a term used to describe some undesired change in the signal and can be the result of reverberation, echo, or changes during transmission (Vaseghi, 2000). In 1970, Carhart and Tillman encouraged the use of speech-in-noise testing as part of a test battery. However, a 2003 survey by Strom (as cited in Taylor, 2007) found that only 42% of dispensing audiologists use speech-in-noise testing as part of a standard test battery. Most patients complain of difficulty understanding speech in the presence of background noise, so speech-in-noise measures are useful to address this concern. Ease of administration and duration of the test are factors to consider when selecting a speech-in-noise test. Examples of speech-in-noise tests include the Speech Perception in Noise (SPIN; Kalikow, Stevens, & Elliott, 1977), Hearing in Noise Test (HINT; Nilsson, Soli, & Sullivan, 1994), QuickSIN (Sentences in Noise; Killion, Niquette, Gudmundsen, Revit, & Banerjee, 2004), Bamford-Kowal-Bench Speech in Noise test (BKB-SIN; Bench, Kowal, & Bamford, 1979), and Words in Noise (WIN; Wilson, 2003).

History of the Development and Validation of Spanish-Language Materials Available for Audiometric Testing

In spite of growing demands, the audiology test materials currently available in Spanish are substandard (Tye-Murray, 2014). As a result, testing for the multilingual population has posed significant practical challenges for audiologists. In addition, many audiologists report a low level of knowledge and confidence in selecting Spanish-language speech-recognition tests. In order to select from the current tests available, the progression of the development and validation must first be reviewed. At the time of the first publication of Spanish-language material, Hispanics comprised 3.5% of the U.S. population (Passel & D'Vera, 2008). As the Hispanic population has grown, the development of testing materials has not kept pace with this growth. In fact, Passel and D'Vera project the Hispanic population to comprise 29% of the U.S. population by 2050, only strengthening the need for well-studied and validated Spanish-language materials. Although tests were developed for research purposes, not all have been validated for clinical use. This poses a challenge for audiologists who seek these measures for clinical use, but are unable to find normative data or supporting research. The following sections provide a chronological historical review of test material development, as well as a discussion of what reliability and validity studies (if any) have been conducted for these test materials. Tables 1 and 2 provide a summary of test materials.

Interest in speech perception in the Spanish language began in the middle of the 20th century. In 1949, Tato published "Lecciones de Audiometria," in which he studied Spanish phonology and created three lists of words based on the composition of the Spanish language. He concluded that Spanish words were typically comprised of two syllables, as very few Spanish words are monosyllabic, and were tetraphonemic, consisting of four phonemes. The three tests developed by Tato (1949) were comprised of 1) 12 phonetically-balanced lists of 25 trochaic words, 2) five lists of 15 trochaic, bisyllabic words that were not phonetically balanced, and 3) three lists of 50 monosyllabic words that were not phonetically balanced. Based on these lists, Tato defined the Spanish articulation curve, which is a function of percent words correct to presentation intensity. He found that Spanish-speaking subjects tested using Spanish stimuli required 10 dB less intensity to obtain the same percentage correct as they obtained using stimuli in English.

Table 1. Summary of speech recognition materials developed for use with Spanish-speaking patients.

Test Name	Author	Stimulus Type	Number of Lists / Stimuli	Example Stimulus
Lecciones de Audiometria	Tato (1948)	trochaic bisyllabic words	12 lists of 25 words (5 lists of 15 phonetically balanced; 3 lists of 50 not phonetically balanced)	n/a
(no title)	Ferrer (1960)	nonsense CVC syllables	4 lists of 50 words	ses, ard, nes, lat, sel
(no title)	Cancel (1965)	1000 bisyllabic *grave* words	20 lists of 50 words	[Casa, taza, masa, raza]; [dama, llama, cama, lama]
(no title)	Tosi (1966)	bisyllabic *grave* words commonly used in Spain and Latin America	12 lists of 2 forms of 648 words	n/a
(no title)	Berruecos and Rodriguez (1967)	phonetically-balanced trochaic words	Sueña, suena, tierra, venta, gesta	n/a
(no title)	Benitez and Speaks (1968)	third-order synthetic sentences with competing message	n/a	n/a
Spanish Multiple Choice Rhyme Test (Spanish MRT)	Tosi (1969)	similar to English Modified Rhyme Test (MRT) - monosyllabic words in sets of six differing by initial or final consonant	n/a	n/a
(no title)	Connery (1977)	Non-phonetically-balanced multisyllabic words ending in a vowel	20-26 words per list; number of lists not available	n/a
(no title)	Spitzer (1980)	words chosen from lists of common Spanish words (objects, animals, body parts, etc)	51 bisyllabic words	niño, toro, perro, suéter, sofá
Boston College Auditory Test	Zubick et al (1983)	bisyllabic and trisyllabic *grave* words	8 lists of 50 trisyllabic words; 7 lists of 50 bisyllabic words	precioso, respeto, espalda, completo, afecto

In 1960, Ferrer reviewed Tato's "Lecciones de Audiometria" and underlined some shortcomings in the work. First, Tato did not establish the clinical application for the lists. Second, Ferrer explained that the slope of the articulation curve depended on the syllables in the words and speaker intelligibility. This caused the articulation curve to vary depending on the type of speech presented and lucidity of the speaker. Lastly, Ferrer (1960) noted an unpublished 1952 study by Berruecos, Faria, and Fernandez in which the researchers used Tato's lists to establish thresholds for intelligibility. There was a 1 dB difference between Berruecos and colleagues' and Tato's work, confirming the lists' use for obtaining speech thresholds. However, Ferrer felt a greater degree of difficulty was needed for a speech discrimination test.

Ferrer (1960) next sought to develop a Spanish language speech discrimination test using nonsense syllables, noting that these stimuli did not depend on the listener's vocabulary and could be limited to certain phonemes. Maintaining Tato's phonetic patterns, Ferrer constructed four lists of 50 nonsense syllables in consonant-vowel-consonant configurations. A pilot study of 11 Spanish-speaking participants with normal hearing found that performance on the test was consistent across participants and lists, supporting the clinical use of the nonsense syllable corpus for speech discrimination testing.

In 1965, Cancel compiled a list of *grave* words from Spanish language newspapers for a multiple choice intelligibility test in Spanish. *Grave* is a term used in Spanish to describe words that are stressed on the penultimate, or second to last, syllable (also known as paroxytone words). Cancel chose *grave* words because they were most similar to spondaic words in English, are a very common word structure in the Spanish language, and are more intelligible in Spanish than single-syllable words. Cancel hypothesized that common *grave* words should be adequate for obtaining reliable scores for assessing speech intelligibility.

In developing the *grave* word lists, Cancel noted the lack of homogeneity in Ferrer's nonsense syllables resulting from Ferrer's prerequisite for a phonetically balanced list. Cancel highlighted factors that should be considered when developing a Spanish speech reception and discrimination test, including intensity levels, degree of difficulty, equivalent measurements in both English and Spanish, phonetic length, position of the phonemes within the word, and presence of nearby sounds. In 1968, Cancel constructed a list of phonetically balanced bisyllabic *grave* words taken from the 1965 lists and developed them into a picture-naming task for use with Spanish-speaking children. Few published data using these lists could be identified; however, it is notable that this appears to be the first of many tests employing common *grave* words as stimuli.

One of the first formal adaptations of an English-language speech recognition test was the Spanish Multiple Choice Rhyme Test (Tosi, 1969). This test was based on the Modified Rhyme Test (MRT; Kruel, Nixon, Kryter, Bell, Land, & Schubert, 1968), which uses rhyming monosyllabic words that differ by either the initial or final consonant. For this test, Tosi constructed a 12-list multiple-choice test using 648 bisyllabic *grave* Spanish words commonly spoken in Latin America and Spain, as well as a list of 1,944 error words to be used as foils.

In 1978, Cooper and Langley evaluated Tosi's Spanish-language MRT for diagnostic use. Sixty native speakers of American English and 60 native speakers of Spanish were assessed by the translated MRT with either monaural auditory only or monaural audiovisual presentations with varying signal-to-noise-ratios (SNR). Based on the number of correct items for each test, Cooper and Langley concluded that the MRT is useful for auditory, visual, or audiovisual performance measurements.

Around the same time, Connery (1977; as cited in Weisleder [1987] and Taylor [2009]) outlined the use of a word list that was used with Spanish-speaking patients at the Chicago Hearing Society. The lists consisted of 20-26 non-phonetically balanced common Spanish words. Connery deemed these lists as not sensitive enough for diagnostic use, but viable for obtaining speech recognition thresholds. Weisleder (1987) noted that an audiologist who was "moderately fluent in Spanish" read the words.

Martin and Hart (1978) also recognized the need for speech audiometry materials in Spanish for children that could be administered and interpreted by a non-Spanish- speaking clinician. To accomplish this, Martin and Hart developed lists of simple English and Spanish words that could be represented visually on illustrated cards and evaluated them in a group of young children. Based on the findings of the study, Martin and Hart concluded that both the English and Spanish lists had high degrees of homogeneity, a quickly upward sloping performance-intensity function within a limited range of intensity, and good interlist equivalency, reliability, and stability. The authors suggested that these materials may be useful not only for children, but also for Spanish-speaking older patients who have little or no knowledge of English.

In 1980, Spitzer noted problems with existing test materials, including difficulty in administering the test by non-Spanish-speaking audiologists. In addition, Spitzer acknowledged the work of Martin and Hart (1978) as feasible and reliable, but noted that its rationale for selection of words may have yielded words unsuitable for clinical use, due to regional variations of the Spanish language, even within the United States. In response, Spitzer created a tape-recorded speech reception threshold (SRT) test to be administered by a non-Spanish-speaking audiologist using a picture-identification task. Test stimuli were selected from Spanish words for people, body parts, clothing, food, animals, and common objects, which were matched to pictures. Spitzer reported good correspondence (within + 10 dB) between the SRT obtained with the test and the pure-tone average, concluding it is a feasible method for obtaining an SRT in Spanish-speaking patients.

Table 2. Summary of speech recognition materials developed for use with Spanish-speaking patients, continued.

Test Name	Author	Stimulus Type	Number of Lists / Stimuli	Example Stimulus
Auditec Spanish Speech Discrimination Lists	Weisleder (1987)	bisyllabic words, most *grave*	4 lists of 50 words	mucho, compra (*grave*); salud, ayer (second-syllable accent)
Spanish Picture Identification Task	McCullough et al (1994)	bisyllabic words selected for easy illustration	2 lists of 50 words	roca, zorro, risa, tasa, sala
Digit SRT (D-SRT)	Ramkissoon et al. (2002)	pairs of digits (monosyllabic numbers between 1 and 9)	56 digit pairs	2-4; 9-3;6-8
Hearing in Noise Test (HINT) - Latin American Spanish	Barón de Otero et al (2008)	high and low context sentences	12 lists of 20 sentences	n/a
Hearing in Noise Test (HINT) - Castilian Spanish	Huarte (2008)	high and low context sentences	24 lists of 10 sentences	n/a
(no title)	Keller (2009)	homogeneous trisyllabic words	1 list of 28 words	apenas, apoyo, comprender, derecho
(no title)	Taylor (2009)	Bisyllabic/trochaic words spoken by male and female speakers	Four lists of 50 words or eight half-lists of 25 words	abrir, ahí, algo, allá, alma
Spanish Speech Perception in Noise Test (SPIN)	Cervera and Gonzalez-Alvarez (2011)	high and low context sentences, similar to the SPIN	6 lists of high-predictability sentences and 6 lists of low-predictability sentences	En el castillo se alza la TORRE (high context); Ha estado pronunciado TORRE (low context)
Spanish Language MRT (Modified Rhyme Test)	Ball (2011)	Bisyllabic words	Six 50-word lists	Olla, papa, abril, tomo, alma Ola, patio, aquí, topo, algo
HearCom Matrix Test - Spanish	Hochmuth et al (2012)	consistently-structured sentences (name, verb, number, object, adjective); closed- and open-set presentations	Twelve triple-lists (3 test lists combined to lists of 30 sentences (part 1); 6 lists of 20 sentences (part 2)	Claudia tiene DOS libros grandes. Carmen hace tres barcos VIEJOS. ELENA toma doce platos nuevos.

In 1983, Zubick, Irizarry, Rosen, Feudo, Kelly, and Strome (1983) developed the Boston College Auditory Test, using *grave*-stressed bisyllabic and trisyllabic words. At the time of publication, Zubick and colleagues noted that field-testing and validation were pending. No further published studies on the Boston College Auditory Lists from this group could be located; however, the psychometric response function of these lists was evaluated in a 2008 study discussed later in this section.

Two significant weaknesses of early Spanish-language speech recognition tests were (1) a lack of standardization in recording and (2) limited information on the effect of presentation level on performance. These factors were assessed in two studies by Weisleder and colleagues (Weisleder, 1987; Weisleder & Hodgson, 1989) for the "Spanish Speech Discrimination Lists 1-4" by Auditec of St. Louis (an original citation for the development of these tests prior to Weisleder's evaluations could not be located).

First, Weisleder (1987) examined the performance intensity functions for the Auditec lists with native speakers of Spanish. His findings showed the /s/ phoneme as the most common source of erroneous responses, which may have been a result of the variants of the /s/ sound in the Spanish language. This is evident for the phonemes /s/, /z/, and /c/, which can be pronounced as /s/ in various dialects. Words that had a plural /s/ phoneme in the final position were also commonly missed. However, the errors did not affect the word's meaning even if the /s/ phoneme was deleted. Weisleder also reported the substitution of /k/ for /t/ phonemes, likely due to the lack of aspiration in Spanish for unvoiced plosive phonemes. He concluded the performance on the word recognition ability tasks was not related to the list but to the presentation level.

Second, Weisleder and Hodgson (1989) evaluated list equivalency of the four Auditec lists. Results suggested that List 3 was statistically significantly less intelligible than the other lists. The authors also noted that study participants of Mexican origin seemed to be at an advantage due to regional variations of the native Mexican speaker on the recording. While Weisleder and Hodgson acknowledged separate lists for each Spanish-speaking region as impractical, they advised audiologists to use the most adequate test for a patient's place of origin. Based on their findings, the authors found the slope of the performance-intensity function to be comparable to that of English lists and, with the exception of List 3, considered the Auditec lists adequate for assessing word recognition abilities in Spanish speakers.

Noting that clinicians who do not speak the same language as their patients may have difficulty understanding and scoring their responses, McCullough and colleagues (1994) developed a Spanish picture-identification task that utilized audio-visual presentation. Test items which could be identified using images were selected from the Department of Veterans Affairs' Picture Identification Tasks (Wilson & Antablin, 1980) and translated to Spanish. The multimedia approach came from a computer connected to two monitors, one of which was in the control room with the audiologist and the other in the test room with the patient. Test items were presented in closed set (a grid of pictures) allowing the clinician to administer and score the test without knowing the language of the test stimuli. In an initial evaluation of the test, English-speaking audiologists were able to administer the Spanish Picture-Identification Task to Spanish speakers successfully.

Ramkissoon and colleagues (2002) took a different approach to the problem of a non-Spanish-speaking tester evaluating a Spanish-speaking patient. Instead of developing Spanish-language stimuli, these researchers created an SRT procedure using pairs of English-language monosyllabic digits between one and nine. This test presumed that even a patient with very limited English proficiency would have some knowledge of the first ten digits. The digit SRT (D-SRT) procedure was evaluated with both native- and non-native-English speakers with normal hearing who underwent testing with the D-SRT and the CID W-1 SRT stimuli. Ramkissoon et al., reported that both measures yielded accurate hearing thresholds for all participants, but the D-SRT was more sensitive than the CID W-1 stimuli for obtaining an SRT. Based on these results, the authors concluded that the D-SRT was effective for obtaining an SRT due to the familiarity of the stimuli (spondaic pairs of digits) rather than the words typically used in an SRT measure, and that English-speaking audiologists should use the D-SRT to obtain an SRT on non-native speakers of English. This approach should allow the audiologist to discern audiometric results for a non-native speaker of English who may be limited by vocabulary, proficiency in English, and educational level from hearing sensitivity.

In 2008, Flores and Ayoama compared the psychometric function of four existing Spanish word recognition tests (the Auditec of St Louis lists, the Boston College Auditory Test, the Comm Tech monosyllabic word test, and the trochaic word lists developed by Berruecos and Rodriguez [1967]). The authors found similar performance for the Auditec of St. Louis and Boston College Auditory test measures in Spanish-speaking patients; however, these results differed from the monosyllabic words from the Comm Tech test and the trochaic word lists from Berruecos and Rodriguez (1967). Flores and Ayoama also found that bilingual speakers who learned English as a second language performed significantly better than bilingual speakers who learned both languages simultaneously, suggesting the effects of linguistic background, such as balance between the participants' first and second languages and pattern of acquisition.

The Hearing in Noise Test (HINT) is a measure of repetition of simple sentences in a background of noise (Nilsson, Soli, & Sullivan, 1994). The HINT is commonly used in clinical settings and has been developed in several other languages, including Latin American Spanish (Barón de Otero, Brik, Flores, Ortiz, & Abdala, 2008) and Castilian (Huarte, 2008).

Given the linguistically diverse countries of Latin America, the authors of the Latin American version of the HINT were challenged to create a test that could be used in several of these countries while avoiding dialectal differences. Using a general dialect of Latin American Spanish typically used by newscasters, the HINT sentences were shared among 14 Latin American countries such as Argentina, Chile, Cuba, Perú, and Venezuela. After considering each country's idiomatic usage, the words were divided into 12 lists of 20 sentences (Barón de Otero et al., 2008). Evaluations of the performance-intensity function of the test were conducted in Mexico, Colombia, and Argentina, and yielded almost identical performance-intensity functions for each list and SNR condition (-7, -4, and -2 dB). Because normative data had not yet been collected at the time of publication, the normative values from the American English HINT were used until norms had been established in Spanish. At the time of this publication, normative data for this test were unavailable.

Castilian is a variation of Spanish and is the official language of Spain. Huarte (2008) developed a Castilian Spanish version of the HINT from translated and adapted sentences from the American English version of the HINT. The phonemes in the Castilian Spanish HINT are typical of those present in conversation. Similar to the procedure described above by Barón de Otero and colleagues, a performance-intensity function was estimated for 24 lists of 10 sentences in the same SNR conditions and an initial set of normative data was collected. Based on this initial evaluation, Huarte recommended the use of the Castilian Spanish HINT for evaluations of adults using hearing aids or cochlear implants.

In 2009, Keller developed and evaluated a speech reception threshold test that used 90 Spanish trisyllabic words selected from a list of the 2,000 most commonly used words by Davies (2006). Test words were recorded by male and female speakers of Spanish. Participants, who were native speakers of Mexican Spanish and had normal hearing sensitivity, listened to and repeated the trisyllabic words, which were then scored by a native Spanish speaker. Keller selected a list of 28 words with the steepest performance-intensity function (10.1% dB for the male talker and 8.7% dB for the female talker) and recommended the use of this list for obtaining SRT from individuals with hearing loss.

Also in 2009, Taylor developed a Spanish word recognition measure with more modern vocabulary and language than the words used in older tests. Male and female adults who were native speakers of Mexican Spanish recorded four lists of 50 words or eight lists of 25 words. The highest ranked female and male speakers were chosen for the recordings. In an initial study of 20 participants with normal hearing, the lists were determined to be homogenous in audibility and psychometric function.

While the SPIN test (Kalikow, Stevens, & Elliot, 1977), a clinical speech perception measure using sentence stimuli, is not available in Spanish, Cervera and González-Alvarez (2011) used it as the basis for developing an intelligibility measure using Spanish sentences in noise. Similar to the English SPIN, the test consisted of high predictability and low predictability sentences presented with three different SNR conditions (0 dB, +5 dB, and +10 dB). Cervera and González-Alvarez (2011) highlighted the advantages of the measure, including ease of administration, simple listener response, and short duration of test. In addition, the test was designed to control for phonetic content, final word stress and frequency, and sentence length. However, to date, there have been no further published studies using these lists.

In 2011, Ball created a Spanish-language version of the Modified Rhyme Test, based on previous work by Tato (1949) and Aguilar (1991). Six lists of 50 words were developed and recorded, and normative data were collected from 44 native Spanish speakers with normal hearing. Although two of the lists produced more errors than the other lists, Ball recommended validation of the words and further use of the lists with Spanish-speaking adults with hearing loss.

In 2012, Hochmuth and colleagues developed a matrix sentence test in Spanish to obtain an SRT. The authors constructed the test as part of the HearCom project, a research project to develop and validate tests into other languages like British English, French, Spanish, Russian, and Greek (see Zokoll, Hochmuth, Warzybok, Wagener, Buschermöhle, & Kollmeier, 2013). Hochmuth and colleagues used the Spanish matrix sentence test to compare the SRT obtained with other matrix tests in other languages, the variability between lists, differences between closed and open-set versions, and performance between subjects from different Spanish-speaking countries. Test lists were generated from the most frequently used words in Spanish (Davies, 2006) to form a sentence that included a name, verb, number, object, and an adjective. Competition noise was created from superimposing all sentences, generating the same long-term average speech spectrum as the sentences for optimal masking.

The Spanish matrix test was then evaluated for practice effects (an effect of 1.1 dB SNR was observed between the first and second measurements) and compared across lists and between open- and closed-set formats. Hochmuth and her co-investigators found that the lists could be used interchangeably, as there were no significant differences between SRTs on the 10 lists. Performance on the open- and closed-set format was also similar. In addition, there were no performance differences between Spanish and Latin American subjects, nor were there regional differences between participants from Tenerife and the Spanish mainland. These findings support the use of the Spanish matrix test for Spanish speakers from different origins.

A few general findings are notable from this review of test material development in Spanish. First, while several attempts have been made to develop speech recognition materials for Spanish-speaking patients, the majority of these have not been validated adequately for clinical use. For many of these materials, no validity studies could be identified at all. Second, it is clear that both the dialect of the patient (i.e., Weisleder & Hodgson, 1989) and the dialect of the audiologist, if materials are presented via live voice (i.e., Weiselder, 1987), are likely to affect scores obtained in speech recognition testing. This presents a challenge for the audiologist to identify regionally-appropriate materials for Spanish-speaking patients and to present those materials in such a way that the tester's knowledge of Spanish and/or dialect have a minimal influence on scoring. Finally, for bilingual patients, selection of most appropriate test materials may be complicated by the nature of each patient's language knowledge and order of

Table 3. Comparison of stimuli for word-recognition and speech-reception-threshold testing in English and Spanish.

Test	English-Language Stimulus	Example Stimulus	Spanish-Language Approximate	Example Stimulus	Rationale
Speech Reception Threshold (SRT)	Spondaic Words (Spondees)	Baseball, Toothbrush, Airplane	Bisyllabic *Grave* (Trochaic) Words	Casa, Puerta, Mono	Spondaic forms (equal stress on both syllables) are uncommon in Spanish. More than half of all Spanish words are *grave* (having stress on the penultimate syllable).
			Trisyllabic *Grave* Words	Cincuenta, Manzana, Hamaca	
Word Recognition Score (WRS)	Consonant-Nucleus-Consonant (CNC) Words	Knock, Tape, Gaze	Bisyllabic *Grave* (Trochaic) Words	Casa, Puerta, Mono	Few concrete words (nouns, simple verbs and adjectives) are monosyllabic in Spanish. The consonant-nucleus-consonant construction of English word-recognition stimuli is uncommon in Spanish.

languages learned (i.e., Flores & Ayoama, 2008).

Factors Affecting Speech Recognition Testing with Bilingual and Spanish-Speaking Patients

Speech recognition ability in all listeners is affected by numerous patient, stimulus, and environmental test factors. It is useful to discuss the research on the effect of some of these factors on speech recognition of Spanish-speaking listeners in particular.

Patient factors. Performance on word recognition tests may be influenced by several characteristics of the patient, including the age of acquisition of the second language and proficiency in the second language.

The age of language acquisition impacts speech perception in noisy and reverberant environments. A 1997 study by Mayo, Florentine, and Buus assessed the performance of Mexican-Spanish-speaking early bilinguals and late bilinguals on speech perception tests. The SPIN test was presented at varying SNR. The results indicated early bilinguals performed better in noise than late bilinguals, but both groups performed equally in quiet conditions. Also, the authors noted the possibility of the first language interfering with an early bilingual's perception of their second language in noise.

A study by Von Hapsburg and Peña (2002) supported the findings of Mayo, Florentine, and Buus (1997) concluding that bilingual listeners did not perform as well as monolingual listeners in the presence of background noise. Von Hapsburg and Peña also noted longer processing times for bilingual listeners, highlighting the effects a timed test might have on a bilingual patient. In 2003, Febo studied the effects of speech perception in early bilinguals compared to monolinguals. The speech perception abilities of monolingual English speakers and early bilingual speakers (who had acquired Spanish and English prior to six years of age) were assessed with varying levels in noisy anechoic and noisy reverberant environments. Febo (2003) learned that the early bilingual participants experienced adverse effects on their speech perception abilities in the noisy environments and scored poorer than monolingual speakers in all noise levels. Both monolingual and bilingual speakers performed similarly in quiet. Results of this study support the idea that bilingual listeners, regardless of proficiency, do not perform as well as monolingual speakers on speech recognition measures in adverse listening conditions.

Von Hapsburg, Champlin, and Shetty (2004) also investigated age of acquisition in bilingual speakers completing a speech perception task. A homogeneous group of bilingual speakers was created based on age of L2 acquisition, language function, language competency, and language history. The reception threshold for sentences (RTS) was found for each participant on the HINT with two speakers at 0 degrees azimuth and 90 degrees azimuth presenting noise. Group comparisons showed equal performance for both late bilinguals and monolinguals in noise, and that bilingual speakers needed a signal-to-noise ratio (SNR) of about 4 dB more than monolingual speakers for the HINT test. The HINT manual states that an SNR difference of 1 dB is equal to nine percentage points for sentence intelligibility, corresponding to a 36% poorer score for bilingual speakers than monolingual speakers when the L2 is used as the stimulus language. Hanks and Johnson (1998) investigated the list equivalency of the HINT for older adult listeners between the ages of 60 and 70 with a mild sensorineural hearing loss. The RTS from their study was about 10 dB greater than the RTS from von Hapsburg, Champlin, and Shetty. Based on these conclusions, von Hapsburg and colleagues suggested that bilingual listeners with normal hearing perform equally or worse than an older adult with a mild hearing loss. However, depending on azimuth of the noise and the presence of background noise, bilingual speakers and individuals with a mild hearing loss showed no differences. The authors attribute the similar scores to additional auditory processing requirements of late bilinguals.

In 2008, Weiss and Dempsey used the Latin American Spanish and English versions of the HINT to compare bilingual speakers' performance. The participants were divided into two groups based on age of second-language acquisition (early bilinguals and late bilinguals) as past studies have indicated that incomplete linguistic profiles make comparisons among studies and subjects difficult (Von Hapsburg & Peña, 2002). Weiss and Dempsey found that all bilingual participants had higher scores on the Latin American Spanish version of the HINT than on the English version in both quiet and noise conditions. The authors also reported higher scores for the late bilingual group, echoing the findings of von Hapsburg and Peña (2002). Although the explanation for these findings was inconclusive, Weiss and Dempsey caution audiologists when choosing the appropriate version of speech perception tests and interpreting the test results due to differences in performance based on the participant's L1 and L2.

Other studies have also addressed the issue of patient's English proficiency (which may be distinct from age of language acquisition) as a determinant in selection of speech audiometry test materials. Although a bilingual listener may use English daily at work or in the community, it may not be prudent to administer English-only speech perception measures during an audiologic evaluation. Shi and Sánchez (2010) recommended speech recognition testing in Spanish or in both languages; however, testing in both languages may not be practical due to busy clinician schedules and patient fatigue. The authors sought to predict the dominant language to administer speech perception tests to bilingual Spanish/English participants. Linguistic variables, such as age of acquisition and use of language were noted for each participant. The English word recognition test came from the NU-6 lists and the Spanish test material was taken from Lists 1, 2, and 4 of the bisyllabic words from Weisleder and Hodgson (1989). Shi and Sánchez learned that the age of acquisition of English, duration of immersion in the English language, self-reported Spanish listening proficiency, and language dominance had the largest impact on bilingual speakers' performance. Performance on one measure did not correlate with performance on the other, and performance may not be predicted by linguistic variables. Shi and Sánchez (2010) recommended using age of acquisition or language dominance to determine the optimal language for word recognition testing instead.

Shi (2014b) sought to replicate results of the 2010 study on predicting success on word recognition measures with bilingual subjects. Comparable results were found, validating the findings of the previous study (Shi & Sánchez, 2010). The proposed models included language dominance, language proficiency, and age of acquisition. Shi recommended the use of these models for audiologists employed in urban settings who work with large Hispanic populations.

Proficiency in a second language is highly influenced by the aforementioned linguistic variables, compelling clinicians to rely on subjective measures of language proficiency. Shi (2011) identified a method to assess a bilingual listener's proficiency in English reliably and efficiently. In the study, 125 bilingual adults were administered the NU-6 word recognition test and were asked to rate their own proficiency in listening, reading, and speaking on the Language Experience and Proficiency Questionnaire (LEAP-Q). Shi noted high sensitivity when the self-reported proficiency in listener was used as the only predictor, but also reported low specificity from overrating their listening proficiency instead of reading and speaking. About 90% of the bilingual listeners reported at least a "good" proficiency in all three domains of English. However, only 68.8% scored a 90% or better on the NU-6 test. Prediction specificity improved when language dominance and age of acquisition of English were factored in self-reported proficiency ratings. Shi concluded that, although the self-rated proficiency was convenient, it had limitations when used with more difficult measures, such as word recognition in noise and with less lenient scoring.

Table 4. Factors to consider in testing word recognition of Spanish-speaking patients.

Patient Factors	Stimulus Factors	Environmental Factors	Talker Factors
Age of second language acquisition (Mayo et al., 1997; von Hapsburg and Peña, 2002; Weiss and Dempsey, 2008)	Item complexity (Cervera and González-Alvarez, 2010	Reverberation (Rogers et al, 2006)	Clear speech (Bradlow and Bent, 2002; Smiljanić & Bradlow, 2008)
Second language proficiency (Shi, 2011, 2014b; Shi and Sánchez, 2010)	Item familiarity (Shi, 2014a; Shi and Sánchez, 2011	Noise (Cooke et al., 2008; Kilman et al., 2014)	
	Item dialect (Rogers et al., 2006; Shi and Canizalez, 2013)		
	Ease of administration by non-language-proficient clinician (Cokely and Yager, 1993)		

Stimulus factors. Characteristics of the stimulus are also likely to influence performance on word-recognition tasks by bilingual patients. These include the familiarity and complexity of test items, dialectical characteristics of test items, and ease of administration of the test by a clinician with limited language proficiency.

Cervera and González-Alvarez (2010) compiled a list of Spanish sentences that have been used in cognition and speech-processing research to study the effects context has on recognition of words, such as with elderly listeners. For example, tests like the SPIN have low-predictability and high-predictability sentences, and are useful for testing elderly patients, as those patients can present with age-related cognition changes. This cognitive decline is independent of hearing sensitivity and may result in higher performance on the high-predictability sentences than the low-predictability sentences (Pichora-Fuller, 2003). If there is no difference in performance between the high- and low-predictability sentences, cognitive processing deficits may be present. The lists Cervera and González-Alvarez chose for this compilation were controlled for length, predictability, and final word frequency. The authors chose six lists of 25 high-predictability sentences and six lists of 25 low-predictability sentences that were equivalent in all of the aforementioned properties. Cervera and González-Alvarez intended these sentences to be used in psycholinguistics, as no equivalent lists had previously existed in the Spanish language. There has been no further testing using these sentence lists.

Word familiarity should also be considered when administering speech recognition measures to bilingual participants. Unfamiliar words can lead to greater perceptual errors than familiar words when administered to both native and non-native listeners. This fact underlines the importance of familiarity in speech recognition in English and Spanish, independently. Shi and Sánchez (2011) explored the role word familiarity played on bilingual participants' performance on the English NU-6 monosyllabic words and Spanish bisyllabic words from Weisleder and Hodgson (1989). Shi and Sánchez learned that there was no difference between familiarity and word recognition scores in quiet and noise conditions. Participants also reported more unfamiliar words in their less-dominant language than in their dominant language, scoring lower on the unfamiliar words than the familiar words. Based on these findings, it is important that participants be tested in their more-dominant language and be familiar with the words used for testing. Shi (2014a) recommended further research to determine if testing should be completed in the more dominant language or in both languages, as well as conducting the measures in either language, given the varying language status in bilingual listeners.

Dialectal differences also have a significant impact on the scoring of word recognition measures. Shi and Canizales (2013) explored the effects of listeners' dialects and their variations on Spanish word recognition tests. The study's subjects included 40 native Spanish speakers with normal hearing who originated from either the Highland region (which includes the Andean regions of South America), the Caribbean, or coastal countries. The subjects were also further divided by dominant language, either English or Spanish. Canizales administered the Auditec bisyllabic Spanish word lists to the subjects in different signal-to-noise ratios (SNR+6, +3, and 0 dB). The authors found significant effects of dialect and language dominance, along with the SNR. However, it should be noted the effects of dialect were independent of those from SNR and language dominance. The results from this study are important for clinicians scoring word recognition measures, as the phonology of the Spanish language and its various dialects can affect results (Rogers, Lister, Febo, Besing, & Abrams, 2006).

Ease of administration is important for Spanish speech audiometry materials, as most clinicians are not bilingual and feel less confident scoring a test for phonemes correctly, particularly if the responses were oral. A clinician may incorrectly score a patient's speech recognition, due to lack of knowledge of the area or linguistic competency of Spanish phonemes. In 1993, Cokely and Yager assessed the scoring of two groups of judges, one group of 15 native English speakers with no knowledge of Spanish and another group of 15 native English speakers who spoke Spanish. Oral responses from the Auditec 1-4 lists were recorded and scored by both groups. Cokely and Yager (1993) found no significant differences between the groups, with both groups of judges obtaining similar word recognition scores (WRS) from oral and written responses. This difference was not deemed clinically significant, suggesting that the language of the scorer did not have an effect on the WRS in the other language. The authors also echoed the findings of Weisleder and Hodgson (1989) after observing statistically significant differences for the Spanish speakers on the Auditec lists; however, they found List 1 to be the outlier, with 13-22% higher than the means of Lists 2, 3, and 4. They suggested the need for further research of the equivalency between the Auditec lists.

Environmental factors. Environmental factors, particularly noise and reverberation as competition for the test stimuli, are also likely to influence performance. Reverberation refers to the reflected sounds from surfaces. If excessive, reverberation from the environment and noise can have adverse effects on speech understanding (Nabelek & Mason, 1981). A 2006 study by Rogers and colleagues compared the performance of monolingual English speakers and Spanish/English bilingual speakers who had learned English before the age of 6. They reported poorer scores for the bilingual participants in noise and reverberation, but equal scores for both groups in the quiet condition. Overall, all participants had lower scores for the noisy and reverberant environments (Rogers, Lister, Febo, Besing & Abrams, 2006). These results indicate that early bilingual listeners have less tolerance for acoustic degradations than monolingual listeners. This may be attributed to increased cognitive demands to process and attend to the active language and isolate the phonemes needed for speech understanding in each language.

Masking can be added to a speech perception measure to ensure the participation of only the test-ear or to simulate a real-world listening environment. There are two types of masking: energetic and informational. Energetic masking is typically used in clinical settings, where the masker's amplitude fluctuates and the stimuli can still be heard during these oscillations. Examples of energetic maskers include multi-talker babble and stationary noise. Informational masking refers to the use of sentences or words that are meaningful. These words can be heard and understood by the patient, and are therefore likely to interfere with the stimuli. Past research has demonstrated that similarity between the masker and stimuli leads to increased effort to separate them (Van Engen, 2010).

The independent contributions of energetic and informational masking in difficult listening environments may be dependent on proficiency in a non-native language. Kilman, Zekveld, Hällgren, and Rönnberg (2014) utilized energetic and informational masking to determine the influence proficiency in a non-native language had on speech perception abilities in noise. The maskers used were stationary noise, fluctuating noise, two-talker babble in Swedish, and two-talker babble in English. Twenty-three native Swedish participants between the ages of 28 years and 64 years who had normal hearing underwent speech recognition testing in the presence of background noise. Participants also underwent a test of working memory capacity, non-verbal reasoning, and English proficiency. Participants had better SRTs when the target speech was in their native language (Swedish). This improvement was also noted for target speech in the non-native language (English) for participants who reported high levels of English proficiency. However, when the masker and target speech were in the same language (i.e. Swedish masker and Swedish target speech), participants experienced more interference and lower SRTs than when the target speech was different from the language of the masker. This highlights the degree to which experience in a non-native language influences difficult speech perception.

The presence of background noise may create additional demands on attention and processing which may be ameliorated with the use of a slower rate of speech. In 2008, Cooke, Lecumberri, and Barker explored the performance of English and

Spanish speakers on the identification of keywords of sentences that were spoken by native speakers of English in the quiet and noise conditions. The noise conditions involved either stationary speech-shaped or competing noise (energetic and informational). Non-native listeners found the task more difficult when the masker level increased, especially when the masking was stationary noise. Compared to the native-speakers, non-native speakers performed worse in both noise conditions. However, when the keywords were produced slowly, the non-native speakers were able to identify more utterances.

Talker factors. Use of clear speech can improve intelligibility of spoken messages for people with hearing loss (Picheny et al., 1985 as cited in Schum, 1996). Clear speech requires the talker to speak louder and slower while decreasing his or her rate of speech, distinguishing phonemes, and increasing phoneme length. A talker may use clear speech when speaking to someone who has a hearing loss or is not native speaker of the talker's language (Smiljanić & Bradlow, 2008).

The benefit of clear speech when listening in a non-native language may be limited. Bradlow and Bent (2002) evaluated clear speech benefit in 32 non-native English speakers and 32 native English speakers with normal hearing. Sentences from a modified version of the Revised Bamford-Kowal-Bench Standard Sentence Test were read by two native speakers of American English (one female and one male), first using a conversational style of speaking and then with clear speech. The sentences were also presented in varying SNR of -4 to -8 dB. Results of the study revealed that the non-native listeners experienced a smaller benefit from clear speech, did not experience negative effects when the noise level was increased, and demographic variables did not appear to be related to speech perception ability. The effects of clear speech were greater for the female talker. Interestingly, Bradlow and Bent (2002) found that clear speech is only fully beneficial for listeners who are familiar with the phonemes and phonology of the language spoken, thereby referring to it as "native-listener oriented."

Summary

Although many tests have been developed, it is evident that further measures of validity and reliability are needed to assess those tests' clinical application. Furthermore, the complexity of the bilingual population sheds light on the need for culturally and linguistically competent clinicians to be aware of these differences when developing and administering these tests. The phonetic and semantic nuances of Spanish and English complicate the effects of sensorineural hearing loss in both pediatric and adult populations. Clinicians should be aware of what test materials exist for primarily Spanish-speaking patients and consider the effects of language dominance, age of second-language acquisition, and language used in the home when evaluating speech recognition test results.

Despite the paucity of well-validated test materials and procedures for testing the Spanish-speaking population, audiologists in all settings must be prepared to appropriately diagnose patients who cannot be assessed using standard English-language materials. While it is difficult to make broad recommendations for testing such a heterogeneous population as would be described by the term "Spanish-speaking patients," we offer the following suggestions. First, it is useful to understand that creation of Spanish-language (or other language) test stimuli directly analogous in form to English-language test stimuli presents challenges. The consonant-nucleus-consonant form of words commonly used in word recognition testing, for example, does not occur in Spanish. Spondaic words are also uncommon in Spanish; most words feature penultimate stress. The reader is referred to Table 3 for a comparison of English and Spanish stimulus types, which may inform comparison between SRT and word recognition scores obtained in both languages. Second, as with all word recognition testing, recorded stimuli are preferable to stimuli presented via live voice. This limits the potential distortion of stimuli introduced by the talker's dialect and knowledge of the language of the stimulus. Picture-pointing tasks may also help overcome these potential problems. Third, dialectical differences should be considered in the response. That is, stimulus items on most of the tests reviewed here are based on Mexican Spanish. Speakers of other Spanish dialects may make a greater number of errors than Mexican Spanish speakers based on relative unfamiliarity of words presented. Finally, conservative interpretation of scores obtained by any of the tests reviewed here is indicated until further reliability and validity studies can be conducted on these measures.

References

Aguilar, L. (1991). Propuesta de un test de rimas modificado para el español (Unpublished master's thesis). Univeristat Autònoma de Barcelona, Barcelona, Spain.

American Speech-Language-Hearing Association. (2008). Audiology survey report: Survey methodology, respondent demographics, and glossary. Retrieved from http://www.asha.org/uploadedFiles/research/memberdata/08AudSurveyMethod.pdf

Association of American Medical Colleges (2010). Diversity in the physician workforce. Retrieved from https://members.aamc.org/eweb/upload/Diversity in the Physician Workforce Facts and Figures 2010.pdf

Baker, C. (2011). *Foundations of bilingual education and bilingualism* (5th ed.). Tonawanda, NY: Multilingual Matters.

Barón de Otero, C., Brik, G., Flores, L., Ortiz, S., & Abdala, C. (2008). The Latin American Spanish hearing in noise test. *International Journal of Audiology, 47*, 362-363. doi: 10.1080/14992020802060888

Bedore, L.M., & Peña, E.D. (2008). Assessment of bilingual children for identification of language impairment: Current findings and implications for practice. *International Journal of Bilingual Education and Bilingualism, 11*(1), 1-29.

Bench, J., Kowal, A., & Bamford, J. (1979). The BKB (Bamford-Kowal-Bench) Sentence Lists for partially-hearing children. *British Journal of Audiology, 13*, 108-112.

Berruecos, T., & Rodriguez, J. (1967). Determination of the phonetic percent in the Spanish language spoken in Mexico City, and the formation of PB lists of trochaic words. *International Journal of Audiology, 6*(2), 211-216.

Boukus, E.R., Cassil, A., & O'Malley, A.S. (2009). A snapshot of U.S. physicians: Key findings from the 2008 Health Tracking Physician Survey. Retrieved from http://www.hschange.com/CONTENT/1078/1078.pdf

Bradlow, A.R., & Bent, T. (2002). The clear speech effect for non-native listeners. *Journal of the Acoustical Society of America, 112*(1), 272-284.

Cancel, C.A. (1965). Multiple-choice intelligibility lists for Spanish speech audiometry. *International Journal of Audiology, 4*(2), 91-93.

Cancel, C.A. (1968). Spanish speech audiometry. *International Journal of Audiology, 7*(2), 206-208.

Carhart R., & Tillman, T.W. (1970). Interaction of competing speech signals with hearing losses. *Archives of Otolaryngology, 91*, 273-279.

Centers for Disease Control (2014). Minority health: Hispanic or Latino populations. Retrieved from http://www.cdc.gov/minorityhealth/populations/REMP/hispanic.html

Cervera, T., & González-Alvarez, J. (2010). Lists of Spanish sentences with equivalent predictability, phonetic content, length, and frequency of the last word. *Perceptual and Motor Skills, 111*(2), 1-13.

Cervera, T., & González-Alvarez, J. (2011). Test of Spanish sentences to measure speech intelligibility in noise conditions. *Behavioral Research, 43*, 459-467. doi:10.3758/s13428-011-0063-2

Chin, N.B., & Wigglesworth, G. (2007). Bilingualism: An advanced resource book. London: Routledge.

Cokely, J.A., & Yager, C.R. (1993). Scoring Spanish word-recognition measures. *Ear and Hearing, 14*(6), 395-400.

Collins, J.G. (1997). Prevalence of selected chronic conditions: United States, 1990-1992. *Vital Health Statistics, 194*, 1-89.

Cooke, M., Lecumberri, M.L.G., & Barker, J. (2008). The foreign language cocktail party problem: Energetic and informational masking effects in non-native speech perception. *Journal of the Acoustical Society of America, 123*(1), 414-427.

Cooper, J.C., & Langley, L.R. (1978). Multiple choice speech discrimination tests for both diagnostic and rehabilitative evaluation: English and Spanish. *Journal of the Academy of Rehabilitative Audiology, XI*(1), 132-141.

Davies, M. (2006). *A frequency dictionary of Spanish: Core vocabulary for learners*. New York, NY: Routledge.

Egan, J. (1948). Articulation testing methods. *The Laryngoscope, 58*, 955-991.

Escarce, J.J., & Kapur, K. (2006). Access to and quality of health care. In M. Tienda & F. Mitchell (Eds.), *Hispanics and the Future of America* (410-455). Washington, D.C.: The National Academies Press.

Febo, D. (2003). Effects of bilingualism, noise, and reverberation on speech perception by listeners with normal hearing. *Graduate Theses and Dissertations*. http://scholarcommons.usf.edu/etd/1364

Ferrer, O. (1960). Speech audiometry: A discrimination test for Spanish language. *Laryngoscope, 58*, 1541-1551.

Flores, L., & Aoyama, K. (2008). A comparison of psychometric performance on four modified Spanish word recognition tests. *Texas Journal of Audiology and Speech-Language Pathology, 31*, 64-70.

Gelfand, S.A. (2009). Essentials of audiology. New York: Thieme.

Goodman, D. (2007). Performance phenomena in simultaneous and sequential bilinguals: A case study of two Chilean bilingual children. *Literatura y Lingüística*, 219-232. doi:10.4067/S0716-58112007000100013

Hanks, W.D., & Johnson, G.D. (1998). HINT list equivalency using older listeners. *Journal of Speech and Hearing Research, 41*, 1335-1340.

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.

Hochmuth, S., Brand, T., Zokoll, M.A., Castro, F.Z., Wargenga, N., & Kollmeier, B. (2012). A Spanish matrix sentence test for assessing speech reception thresholds in noise. *International Journal of Audiology, 51*, 536-544.

Huarte, A. (2008). The Castilian Spanish hearing in noise test. *International Journal of Audiology, 47*, 369-370.

Kalikow, D.N., Stevens, K.N., & Elliot, L.L. (1977). Development of a test of speech intelligibility in noise using sentence materials with controlled word predictability. *Journal of the Acoustical Society of America, 61*, 1337-1351.

Keller, L.A. (2009). Psychometrically equivalent trisyllabic words for speech reception threshold testing in Spanish. Retrieved from http://scholarsarchive.byu.edu/cgi/viewcontent.cgi?article=2948&context=etd

Killion, M.C., Niquete, P.A., Gudmundsen, G.I., Revit, L.J., & Banerjee, S. (2004). Development of a quick speech-in-noise test for measuring signal-to-noise ratio loss in normal-hearing and hearing-impaired listeners. *Journal of the Acoustical Society of America, 116*(4), 2395-2405.

Kilman, L., Zekveld, A., Hällgren, M., & Rönnberg, J. (2014). The influence of non-native language proficiency on speech perception performance. *Frontiers in Psychology, 5*, 1-9.

Knapp, K., & Seidlhofer, B. (2009). *Handbook of foreign language communication and learning*. Berlin: Walter de Gruyter & Co.

Kruel, E.J., Nixon, J.C., Kryter, K.D., Bell, D.W., Land, J.S., & Schubert, E.D. (1968). A proposed clinical test of speech discrimination. *Journal of Speech and Hearing Research, 11*, 536-552.

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.

Mayo, L.H., Florentine, M., & Buus, S. (1997). Age of second-language acquisition and perception of speech in noise. *Journal of Speech, Language, and Hearing Research, 40*, 686-693.

McArdle, R. (n.d.). Speech recognition testing: The basics. [PDF document]. Retrieved from http://www.afaslp.org/AVAA conferences/McArdle_Speech in Noise.pdf

McArdle, R.A., Wilson, R.H., & Burks, C.A. (2005). Speech recognition in multitalker babble using digits, words, and sentences. *Journal of the American Academy of Audiology, 16*, 726-739.

McCullough, J.A., Wilson, R.H., Birck, J.D., & Anderson, L.G. (1994). A multimedia approach for estimating speech recognition of multilingual clients. *American Journal of Audiology, 3*(1), 19-22.

Mendel, L.L., & Owen, S.R. (2011). A study of recorded versus live voice word recognition. *International Journal of Audiology, 50*(10), 688-693.

Morrison, A.M. (2008). Providing culturally effective audiological services to the Hispanic pediatric population. Retrieved from http://kb.osu.edu/dspace/handle/1811/49164

Nabelek, A.K., & Mason, D. (1981). Effect of noise and reverberation on binaural and monaural word identification by subjects with various audiograms. *Journal of Speech, Language, and Hearing Research, 24*(3), 375-383.

National Association for Language Development in the Curriculum (2009). What is bilingualism? Retrieved from http://www.naldic.org.uk/Resources/NALDIC/Initial Teacher Education/Documents/B1.pdf

Nilsson, M., Soli, S.D., & Sullivan, J.A. (1994). Development of the Hearing in Noise Test for the measurement of speech reception thresholds in quiet and in noise. *Journal of the Acoustical Society of America, 95*(2), 1085-1099.

Ortman, J.M., & Shin, H.B. (2011). Language projections: 2010 to 2020. Paper presented at the Annual Meeting of the American Sociological Association, Las Vegas, NV, August 20-23, 2011. Retrieved from http://www.census.gov/hhes/socdemo/language/data/acs/Ortman_Shin_ASA2011_paper.pdf

Passel, J., & D'Vera, C. (2008). U.S. population projections: 2005-2050. Retrieved from http://www.pewhispanic.org/files/reports/85.pdf

Picheny, M., Durlach, N., & Braida, L. (1985). Speaking clearly for the hard of hearing I: Intelligibility differences between clear and conversational speech. *Journal of Speech, Language, and Hearing Research, 28*, 96-103.

Pichora-Fuller, M.K. (2003). Processing speed and timing in aging adults: Psychoacoustics, speech perception, and comprehension. *International Journal of Audiology, 42*, S59-S67.

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

Rogers, C., Lister, J., Febo, D., Besing, J., & Abrams, H. (2006). Effects of bilingualism, noise, and reverberation on speech perception by listeners with normal hearing. *Applied Psycholinguistics, 27*, 465-485.

Saha, S., Taggart, S.H., Komaromy, M., & Bindman, A.B. (2000). Do patients choose physicians of their own race? *Health Affairs*, 19(4), 76-83. doi:10.1377/hlthaff.19.4.76

Schum, D.J. (1996). Intelligibility of clear and conversational speech of young and elderly talkers. *Journal of the American Academy of Audiology*, 7, 212-218.

Shi, L.F. (2014a). Speech audiometry and Spanish-English bilinguals: Challenges in clinical practice. *American Journal of Audiology, 23*, 243-259.

Shi, L.F. (2014b). Validating models of clinical word recognition tests for Spanish/English bilinguals. *Journal of Speech, Language, and Hearing Research, 57*, 1896-1907.

Shi, L.F. (2011). How "proficient" is proficient? Subjective proficiency as a predictor of bilingual listeners' recognition of English words. *American Journal of Audiology, 20*, 19-32.

Shi, L.F., & Canizales, L.A. (2013). Dialectal effects on a clinical Spanish word recognition test. *American Journal of Audiology, 22*, 74-83

Shi, L.F., & Sánchez, 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.

Shi, L.F., & Sánchez, D. (2011). The role of word familiarity in Spanish/English bilingual word recognition. *International Journal of Audiology, 50*, 66-76.

Smiljanić, R., & Bradlow, A.R. (2008). Temporal organization of English clear and conversational speech. *Journal of the Acoustical Society of America, 124*(5), 3171-3182.

Soares, C., & Grosjean, F. (1984). Bilinguals in a monolingual and a bilingual speech mode: The effect on lexical access. *Memory & Cognition, 12*(4), 380-386.

Spitzer, J.B. (1980). The development of picture speech reception threshold test in Spanish for use with urban U.S. residents of Hispanic background. *Journal of Communication Disorders, 13*, 147-151.

Talamantes, M., Lindeman, R., & Mouton, C. (2001). Health and health care of Hispanic/Latino American elders. *Curriculum in Ethnogeriatrics*. Retrieved from http://www.stanford.edu/group/ethnoger/

Tato, J. (1949). Lecciones de audiometría. Buenos Aires: El Ateneo.

Taylor, A.M. (2009). Psychometrically equivalent bisyllabic word-lists for word recognition testing in Spanish. Retrieved from http://scholarsarchive.byu.edu/cgi/viewcontent.cgi?article=3100&context=etd

Taylor, B. (2007). Predicting real world hearing aid benefit with speech audiometry: An evidence-based review. Retrieved from http://www.audiologyonline.com/articles/predicting-real-world-hearing-aid-946

Tillman, T.W., & Carhart, R. (1966). *An expanded test for speech discrimination utilizing CNC monosyllabic words: Northwestern University Auditory Test No. 6. Technical report SAM-TR-66-55*. Brooks Air Force Base, TX: USAF School of Aerospace Medicine.

Tosi, O.E. (1969). Estudio experimental sobre la inteligibilidad de un test de multiple elección en idioma español. *Fonoaud, 15*, 28-35.

Tye-Murray, N. (2014). *Foundations of aural rehabilitation: Children, adults, and their family members*. Stamford, CT: Cengage Learning.

U.S. Census Bureau. (2011). Profile America: Facts for features. Retrieved from https://www.census.gov/newsroom/releases/archives/facts_for_features_special_editions/cb11-ff18.htmlhttps://www.census.gov/newsroom/releases/archives/facts_for_features_special_editions/cb11-ff18.html

U.S. Department of Health and Human Services. (2013). The U.S. Health Workforce Chartbook: Part IV. Retrieved from http://bhpr.hrsa.gov/healthworkforce/supplydemand/usworkforce/chartbook/chartbookpart4.pdf

Van Engen, K.J. (2010). Similarity and familiarity: Second language sentence recognition in first- and second-language multi-talker babble. *Speech Communication*. doi:10.1016/j.specom.2010.05.002

Vaseghi, S.V. (2000). Noise and distortion. In *Advanced Digital Signal Processing and Noise Reduction* (29-43). Retrieved from http://dsp-book.narod.ru/295.pdf

von Hapsburg, D., & Peña, E.D. (2002). Understanding bilingualism and its impact on speech audiometry. *Journal of Speech, Language, and Hearing Research, 45*, 202-213.

von Hapsburg, D., Champlin, C.A., & Shetty, S.R. (2004). Reception thresholds for sentences in bilingual (Spanish/English) and monolingual (English) listeners. *Journal of the American Academy of Audiology, 15*, 88-98.

Weisleder, P. (1987). Comparative intelligibility functions and some normative data of four Spanish word recognition ability lists. http://arizona.openrepository.com/arizona/handle/10150/276520

Weisleder, P., & Hodgson, W.R. (1989). Evaluation of four Spanish word-recognition ability lists. *Ear and Hearing, 10*(6), 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), 5-17. doi:10.3766/jaaa.19.1.2

Wilson, R.H. (2003). Development of a speech in multitalker babble paradigm to assess word-recognition performance. *Journal of the American Academy of Audiology, 14*, 453-470.

Wilson, R.H., & Antablin, J.K. (1980). A picture-identification task as an estimate of the word-recognition performance of nonverbal adults. *Journal of Speech and Hearing Disorders, 45*, 223-248.

World Health Organization. (2014). Deafness and hearing loss. Retrieved from http://www.who.int/mediacentre/factsheets/fs300/en/

Zokoll, M.A., Hochmuth, S., Warzybok, A., Wagener, K.C., Buschermöhle, M., & Kollmeier, B. (2013). Speech-in-noise tests for multilingual hearing screening and diagnostics. *American Journal of Audiology, 22*, 175-178.

Zubick, H., Irizarry, L., M., Rosen, L., Feudo, P., Kelly, J.H., & Strome, M. (1983). Development of speech audiometry materials for native Spanish-speaking adults. *Audiology, 22*, 88-102.