



Published in final edited form as:

Am J Audiol. 2014 June ; 23(2): 158–172. doi:10.1044/2014_AJA-13-0055.

Development and preliminary evaluation of a pediatric Spanish/English speech perception task

Lauren Calandruccio¹, Bianca Gomez¹, Emily Buss², and Lori J. Leibold¹

¹Division of Speech and Hearing Sciences, Department of Allied Health Sciences, The University of North Carolina School of Medicine, Chapel Hill, NC 27599

²Department of Otolaryngology, The University of North Carolina School of Medicine, Chapel Hill, NC 27599

Abstract

Purpose—To develop a task to evaluate children’s English and Spanish speech perception abilities in either noise or competing speech maskers.

Methods—Eight bilingual Spanish/English and eight age matched monolingual English children (ages 4.9–16.4 years) were tested. A forced-choice, picture-pointing paradigm was selected for adaptively estimating masked speech reception thresholds. Speech stimuli were spoken by simultaneous bilingual Spanish/English talkers. The target stimuli were thirty disyllabic English and Spanish words, familiar to five-year-olds, and easily illustrated. Competing stimuli included either two-talker English or two-talker Spanish speech (corresponding to target language) and spectrally matched noise.

Results—For both groups of children, regardless of test language, performance was significantly worse for the two-talker than the noise masker. No difference in performance was found between bilingual and monolingual children. Bilingual children performed significantly better in English than in Spanish in competing speech. For all listening conditions, performance improved with increasing age.

Conclusions—Results indicate that the stimuli and task are appropriate for speech recognition testing in both languages, providing a more conventional measure of speech-in-noise perception as well as a measure of complex listening. Further research is needed to determine performance for Spanish-dominant listeners and to evaluate the feasibility of implementation into routine clinical use.

INTRODUCTION

Due to the changing linguistic demographics of the United States, specifically the large influx of Spanish speakers (Ryan, 2013), it is imperative that the audiology community considers appropriate test alternatives for the assessment of speech perception abilities of children who speak English as their second language. The purpose of the present experiment

was to develop a task to assess the speech perception abilities of Spanish/English bilingual children in both Spanish and in English in either noise or speech maskers.

A growing body of adult speech recognition literature indicates that linguistic experience can affect speech recognition (Gat & Keith, 1978; Mayo, Florentine, & Buus, 1997; Rimikis, Smiljani, & Calandruccio, 2013; Rogers, Lister, Febo, Besing, & Abrams, 2006; Shi, 2009). Though listeners attending to their second language (L2) in quiet often have good speech recognition scores, performance tends to be significantly poorer for these same listeners compared to their monolingual counterparts when they are asked to listen in the presence of competing noise (Gat & Keith, 1978; Rogers et al., 2006; Shi, 2009). Therefore, when assessing patients in the clinic who are nonnative speakers of English it is difficult to separate poor speech recognition performance due to a listener's linguistic inexperience from poor performance due to a listener's hearing loss (von Hapsburg & Pena, 2002). Testing becomes further complicated when considering bilingual children who are still in the process of not only acquiring their first language (L1), but also their L2 (Nicoladis & Genesee, 1997). Though more limited than bilingual adult speech recognition data, there is evidence to support the idea that bilingual children are also significantly disadvantaged when listening to their L2 in competing noise compared to monolingual age matched controls (Crandell & Smaldino, 1996; Nelson, Kohnert, Sabur, & Shaw, 2005).

There are many factors to consider when testing the speech perception abilities of nonnative speakers, including which language (L1 or L2) to use for testing, which test material is appropriate for the listener (due to L2 lexicon constraints and the age of the listener), and how to score listeners' responses (due to accentedness of the nonnative speakers' L2 speech production; Rimikis et al., 2013; von Hapsburg & Pena, 2002). Another consideration is the linguistic history of the tester. Only 5% of American Speech-Language-Hearing Association (ASHA) members consider themselves to be bilingual providers (ASHA, 2012). Therefore, the ability of the tester to correctly pronounce the speech tokens using monitored live voice is limited. Further, even if the tester has access to recorded audiological test materials that have been developed in other languages (e.g., the Hearing in Noise Test; Soli & Wong, 2008), correctly scoring an open-set speech recognition response may be very difficult for native English-speaking clinicians.

Several researchers have suggested that closed-set testing may be a good alternative for testing nonnative speakers of English (Jerger, Speaks, & Trammell, 1968; Rimikis et al., 2013). Rimikis and colleagues reported that, for an open-set task, inter-scorer variability was negatively correlated with English pronunciation scores as measured by the Versant (Pearson) automated speech recognition test, demonstrating the difficulty that occurs for the examiner when scoring foreign-accented speech during clinical assessment. Rimikis et al. (2013) argued that providing listeners with a closed set would alleviate scoring difficulties for the examiner; if the listener is forced to choose a response (e.g., by pointing to a picture or word), then scoring is not negatively influenced by the accentedness of the nonnative speaker's speech production. Further, Jerger and colleagues (1968) argued that providing listeners with a closed set, and therefore restricting the possible responses to a given test item, would reduce the linguistic bias inherent in open-set speech recognition testing. Though Garstecki and Wilkin (1976) reported that nonnative speakers performed more

poorly than their monolingual counterparts on a closed-set *sentence* recognition test in noise, closed-set testing with isolated *words* has been shown to limit the effects of linguistic parameters that often influence speech recognition scores, such as word frequency (Sommers, Kirk, & Pisoni, 1997).

Picture identification has been shown to be a useful means of assessing word recognition for children (Hall, Grose, Buss, & Dev, 2002; Litovsky, 2005; Ross & Lerman, 1970), and is included as a recommended behavioral assessment tool within the guidelines developed by ASHA for the audiological assessment of children (ASHA, 2004). Picture pointing as a listener response has many advantages over free response. Verbal responses can be difficult to score accurately if the child's speech production is atypical for some reason (e.g., hearing loss or linguistic history). Written responses are problematic for young children (first grade and younger; Armbruster, Lehr, & Osborn, 2001), as well as older children with minimal education or developmental delays.

Two picture-pointing tests that have gained clinical popularity include the Word Intelligibility by Picture Identification (WIPI; Ross & Lerman, 1970) and the Northwestern University Children's Perception of Speech (NU-CHIPS; Elliott & Katz, 1980) tests. These tests were developed specifically to measure the speech discrimination abilities of children with hearing loss, and continue to be recommended for pediatric audiological assessment (ASHA, 2004). Both measurements assess monosyllabic word identification using a forced choice paradigm with four test lists that are equally difficult. The tests were designed to be conducted in quiet and to evaluate differences in word recognition between populations of listeners, between ears, and between amplified listening conditions for children with hearing loss. Some have argued that when the WIPI and the NU-CHIPS have been administered with a background masker, the different lists are no longer equally difficult (Chermak, Pederson, & Bendel, 1984; Chermak, Wagner, & Bendel, 1988). This across-list variability decreases the usefulness of these tools for evaluating differences in masked word recognition between ears and between listening conditions (e.g., when evaluating with and without hearing aids).

Evaluating speech perception in the presence of a background masker is an important component of functional hearing assessment, because it more accurately resembles many natural listening conditions. However, it turns out that the content of that masker is critically important. Carhart and Tillman (1970) reported that the detrimental effects of including a background masker are larger for listeners with sensorineural hearing loss than for those with conductive hearing loss or normal hearing. This is true for steady-state noise and multi-talker babble maskers, expected to interfere with the peripheral encoding of the target speech primarily via energetic masking (Carhart, Tillman, & Greetis, 1969), but the additional masking experienced by listeners with sensorineural hearing loss can be even more evident for maskers composed of a small number of competing talkers, believed to produce both energetic and informational masking (Brungart, Chang, Simpson, & Wang, 2009; Freyman, Balakrishnan, & Helfer, 2004). At the time of their report, Carhart and Tillman argued that test tools should be developed and implemented in the clinic that included speech recognition measures in noise *and* competing speech. Such a tool would be invaluable in predicting the communication disability experienced by those with hearing

loss, which is not captured well using the two types of tests traditionally used in the clinic: pure-tone thresholds and speech recognition in quiet. Since Carhart and Tillman published their report, there has been an influx of data in support of their argument (see Grant & Walden, 2013; Nilsson, Soli, & Sullivan, 1994; Wilson, McArdle, & Smith, 2007, and others), yet few clinical tests incorporating speech maskers have gained traction among the majority of clinicians, and none include both noise *and* speech as competing maskers. This appears to be especially problematic for children, who consistently show a larger gap in performance compared to adults in competing speech than steady-state noise maskers (e.g., Bonino, Leibold, & Buss, 2013; Hall et al., 2002).

A major difference between this pediatric speech perception measure and other tools currently available is the ability to test perception in the presence of either noise or two-talker maskers. Competing speech with a small number of talkers has been shown to be a more effective masker for adults than competing steady state noise or multi-talker babble (Brungart et al., 2009; Freyman et al., 2004; Rosen, Souza, Ekelund, & Majeed, 2013; Simpson & Cooke, 2005). This increase in masker effectiveness is thought to be due to an increase in informational masking (Carhart et al., 1969). In other words, the competing speech causes energetic masking due to similar excitation patterns in the auditory periphery caused by both the target and masker stimuli, as well as confusion between which speech cues originate from the masker and which originate from the target. Complex maskers, such as those composed of two competing talkers, are thought to be more indicative of the types of background sounds that are encountered in everyday life.

The goal of the present project was to develop a closed-set picture identification task that could be used to assess monolingual English, monolingual Spanish, and Spanish/English bilingual children in the research laboratory, with the long-term goal of developing materials and procedures for clinical use. The picture-pointing response precludes any need to score production responses, and the use of recorded materials reduces the importance of the audiologist's familiarity with the test language. The task was designed so that the examiner can decide on a case-by-case basis which test language (L1 or L2) is most appropriate to use for testing, and gives the option of testing in background noise, competing speech, or both. Significant care was taken to balance the spectral and temporal acoustic properties, as well as the psycholinguistic properties of materials in the two test languages (described below in the Methods section).

Three a priori hypotheses were tested. First, we expected that children, regardless of the test language, would perform significantly better in the speech shaped noise than the two-talker maskers. This result would be consistent with previous data (e.g., Hall et al., 2002) indicating more pronounced developmental effects with increasing stimulus complexity. Second, it was hypothesized that performance on the English recognition task would be similar between the bilingual and monolingual children. This result would be consistent with data from simultaneous adult bilinguals (Calandruccio & Zhou, in press; Mayo et al., 1997; Shi, 2010). Finally, it was predicted that performance in the two-talker maskers of older children would be better than that of younger children for both monolingual and bilingual children. This finding would be consistent with the hypothesis that children become more proficient in complex acoustic environments with increasing age and listening experience.

METHODS

Test Development

Table 1 shows the 30 target words, obtained through a combination of the Dolch word list (Dolch, 1948) and a general search of children's literature in both English and Spanish. The search focused on nouns that were in the vocabulary of a typical five-year-old. Criteria for word inclusion required that the word was disyllabic when spoken in both English and Spanish (e.g., feath-er and plu-ma) and that the word could be easily illustrated in a way that was visually unambiguous. Psycholinguistic statistics were assessed for both English and Spanish target words using N-Watch (Davis, 2005) and B-Pal (Davis & Perea, 2005) software, respectively. These programs calculate many different psycholinguistic statistics including, but not limited to, word frequency, phonological neighborhoods, familiarity, and imageability, using a collection of psycholinguistic databases (see many references within Davis, 2005 and Davis & Perea, 2005). These psycholinguistic statistics were used to ensure lexical similarity between the Spanish and English word sets. These statistics appear in Table 1.

Word-frequency statistics provide information on how often a word is expected to occur within the lexicon. English word frequency values provided in Table 1 are based on the CELEX English linguistic database (Baayen, Piepenbrock, & van Rijnm, 1995), while Spanish word frequency values are based on the LEXESP database of five million Spanish words (Sebastian-Galles, Marti, Cuetos, & Carreiras, 2000). Word-frequency values for both languages are reported as the logarithm (base 10) of one plus the count per million words. The average word frequency for the English and Spanish words was 1.36 ($SD = 0.71$) and 1.30 ($SD = 0.63$), respectively.

Both software packages used to compute phonological neighbors (N-Watch and B-Pal) include neighbors based on substitutions, deletions, and additions. The average number of phonological neighbors for the English and Spanish target words was 4.40 ($SD = 3.48$; range 0 – 13 neighbors) and 5.43 ($SD = 6.66$; range 0 – 21 neighbors), respectively.

Subjective familiarity ratings are based on adult responses indicating word familiarity on a scale of 100 (very unfamiliar) to 700 (very familiar). For the English and Spanish lexicons respectively, there were eight and five words that did not have familiarity ratings available (indicated by a '-' in Table 1). The average familiarity ratings for the English and Spanish target words that did have familiarity ratings were 564 ($SD = 46.64$) and 598 ($SD = 66.99$), respectively.

Imageability indicates how easy it is to form a mental image of a word. During testing listeners were asked to point to the picture they thought they heard. Therefore, one factor in word selection was maintaining similar levels of imageability between the tokens spoken in both languages. Imageability scores are based on a scale of 100 to 700, where higher numbers indicate greater (or easier) imageability. Imageability scores for English words were obtained using N-Watch and were calculated using the MRC Psycholinguistic Database (Coltheart, 1981). If the target word was not included in the MRC database, then the imageability score was taken from norms reported by Bird, Franklin and Howard (2001;

data that can also be obtained using the N-Watch program). Spanish imageability scores were obtained using B-Pal (Davis & Perea, 2005). Cases for which imageability scores were unavailable are indicated with a '-' in Table 1. Average imageability scores for the English and Spanish target words were 596 ($SD = 45.23$) and 606 ($SD = 65.57$), respectively.

All of the psycholinguistic statistics discussed thus far (also provided in Table 1) are based on data collected from adults. There are limited cross-linguistic resources providing data based on children's linguistic experience. The Cross Linguistic Lexical Norm database (Jorgensen, Dale, Bleses, & Fenson, 2010) was examined to assess word familiarity between the two languages in children. This database provides word-learning norms from different MacArthur-Bates Communicative Development Inventories (Fenson et al., 2007) studies that have reported early language learning from parental report for infants and young children up to 18 months of age. Though we were able to assess the test words in both languages based on normative data for American English and Mexican Spanish, only approximately half of the words were found in the database; this is presumably due to the differences in age between the subjects studied in constructing the database (18 months old) and our target demographic (approximately five year old and older). Mean values obtained using the database were similar between the 17 English words and 16 Spanish words found in the database (59% and 54%, respectively, indicating that this percentage of 18 months old knew these words). In addition, the 30 English words were compared to a corpus of speech produced by kindergarten and first grade children (Storkel & Hoover, 2010) to determine if typically developing children in this age range spontaneously produced the English tokens. All 30 English words were found in the Storkel and Hoover database. A similar lexicon based on Spanish speech was not available to complete a similar analysis for the 30 Spanish words.

Phonetic inventories for the target words indicated that the 30 words in English and 30 words in Spanish had similar numbers of vowels (59 and 60, respectively), stop consonants (31 and 30, respectively), affricate/fricative consonants (13 and 15, respectively), and nasals (21 and 18, respectively). Matching stimuli in this way ensures that listeners with hearing loss would not be at a disadvantage when using one language vs. the other due to an imbalance in, for example, high-frequency speech sounds that could be hard to hear due to high-frequency hearing loss.

Stimulus Recordings

Target words and masker speech were recorded from six female Spanish/English bilingual speakers, ages 18–24 years. All talkers sat alone in a single-walled sound isolated room approximately one meter in front of a cardioid condenser Shure microphone. The microphone was connected to an M-audio D/A convertor, and stimuli were recorded using Audacity© (44.1 kHz sampling rate, 16 bit). Three of the six talkers were selected for inclusion in the test based on subjective judgments of their fluency and “native” accent in both languages by two of the authors (LC and BG). All three talkers included in the test were students at the University of North Carolina at Chapel Hill, and two were born and raised in the United States. All three talkers grew up in Spanish-English bilingual households. One of the three talkers was used as the target voice for both Spanish and

English. The other two talkers were used to create the two-talker maskers in both English and Spanish. The same voices were used to create the stimuli in both English and Spanish to minimize the spectral differences between languages. The target talker was a simultaneous bilingual whose mother spoke Venezuelan-Spanish to her, while her father spoke English. One of the talkers used to create the masker speech grew up speaking Mexican Spanish with her mother, and English with her father. The second talker used to create the masker speech was exposed to El Salvadorian Spanish at home from both of her parents. All three talkers consistently use both English and Spanish in their daily life. Subjective self-reported measures indicated that the three talkers rated their reading, speaking, and listening abilities in English and Spanish as a 10 on a scale from 1–10, with 10 indicating excellent. None of the bilingual speakers indicated that their writing in Spanish was equivalently good as their writing skills in English (with an average ranking of 7).

The target words were digitally edited to remove all silence before and after the token. Average duration of English and Spanish target words was .55 sec (range .34 – .80 sec). The speaking rates for individual words are provided in Table II. All trimmed tokens were then root-mean-square (RMS) equalized to the same pressure level using Praat (Boersma & Weenink, 2012).

English masker stimuli were composed of passages from the English story, *Jack and the Beanstalk* (Walker, 1999a), while the Spanish masker stimuli were composed of passages from the Spanish translation of this story, *Juan y Los Frijoles Magicos* (Walker, 1999b). Two English/Spanish bilingual talkers were used to create the maskers. Each talker was separately recorded reading each version of the book (English and Spanish). One of the talkers began reading both books from the first page, while the second talker began reading the books from the middle of the story. This prevented the two talkers from simultaneously reading the same segment of text at the same time. Once recorded, silent periods greater than 300 ms were digitally removed from the masker speech to prevent long pauses within the masker stimuli. All editing was completed using SoundStudio© audio software. Once each talker's speech was edited removing silences, extraneous noises, or mis-spoken words, the recordings were RMS equalized using Praat. Recordings from the two listeners were summed to create the two two-talker maskers, one composed of English productions and the other composed of Spanish productions. The endings of the two talker maskers were then trimmed so that each ended with both talkers saying a complete word (final English duration, 2 min 48 sec; final Spanish duration, 2 min 57 sec). Speaking rates of both talkers used to create the maskers for both languages are included in Table 2. A third masker condition was also created based on the two two-talker masker wav files. A 10-sec noise masker spectrally matched to the spectra of the two-talker maskers was generated in MATLAB.

Recall that the same Spanish-English bilingual talkers were used to create the English and Spanish stimuli. This was done to minimize spectral and temporal differences between conditions that have been shown to confound interpretation across listening conditions using different two-talker maskers (Calandruccio, Dhar & Bradlow, 2010). Figure 1 illustrates the long-term-average speech spectra (LTASS) for the target talker producing all 30-test words when speaking in English and Spanish, as well as the two two-talker maskers, spoken in

English and Spanish. The LTASS of the two languages were very similar. The only observable difference was an increase in energy between approximately 3500–4500 Hz for the English two-talker masker compared to the Spanish (see Figure 1). Though not perceptually noticeable, due to this slight difference in energy, the LTASS of the two maskers were normalized using MATLAB[®]. The LTASS for each of the two two-talker maskers were determined by performing a Fast Fourier Analysis on 2048-point hamming-windowed samples, and then computing the average magnitude spectrum. The resulting LTASS for the two maskers were used to compute the grand average LTASS for the two two-talker maskers. This grand average was then used to normalize the individual magnitude spectra of the two maskers (see Brouwer, Van Engen, Calandruccio, & Bradlow, 2012). The spectrum of the LTASS-normalized two-talker maskers is also shown in Figure 1.

In addition to spectral differences between the English and Spanish two-talker maskers that could affect masker effectiveness, temporal differences between the two two-talker maskers were also evaluated. This was done because when comparing between two two-talker maskers, it has been shown that differences in amplitude modulation patterns can cause differences in masker effectiveness regardless of the linguistic content of the masker speech (Calandruccio et al., 2010). Figure 1 also shows the cumulative distribution of the filtered envelope values. These values were based on the Hilbert envelopes of the two maskers, low-pass filtered using a 2nd order Butterworth filter with a 40-Hz cutoff. This filtering allowed for a quantitative assessment of the masker-envelope minima that were available to the listener. A larger proportion of relatively low envelope values would indicate a greater opportunity for dip listening (Festen & Plomp, 1990). The cumulative distributions of filtered envelope of the two maskers were nearly identical.

Listeners

Sixteen children (ages 4.9 –16.4 years) currently living in North Carolina participated in the experiment. Eight children were bilingual speakers of Spanish and English, and eight were monolingual speakers of English. The bilingual and monolingual children were age matched within 6 months, with the exception of two pairs of children (a 9 and 10 year old that were 13 months different in age, and a 15 and 16 year old that were 15 months different in age). For this sample size, the power of this test was estimated to be 0.86 based on data previously collected in our laboratory. All bilingual children identified as Hispanic, while all of the monolingual children identified as not Hispanic. All children were born and raised in the United States with the exception of Listener 7, who emigrated from Columbia at 5 years of age. A parent provided information regarding the linguistic history of the bilingual children using a questionnaire developed by Gutierrez-Clellen and Kreiter (2003); results obtained with this instrument have been shown to correlate with linguistic proficiency. The bilingual children were exposed to Spanish from birth in the home through at least one parent, with five of the eight children exposed to Spanish through both parents. Six of the children were reported to be simultaneous bilinguals, i.e., they began learning both languages before their first birthday (McLaughlin, 1978). Two of the children learned English after they acquired Spanish. Parental report indicated that the bilingual children consistently use both languages. Further detail regarding bilingual listeners' Spanish and English linguistic history is shown in Table 3.

Perceptual testing procedure

The Institutional Review Board (IRB) at the University of North Carolina at Chapel Hill approved all procedures. All listeners had hearing thresholds equal to or less than 20 dB HL bilaterally at octave frequencies between 250 and 8000 Hz (ANSI, 2010; ISO, 2012). Audiometric thresholds were measured using a Grason-Stadler 61 clinical audiometer and either TDH headphones or ER3A (Etymotic) insert phones. All speech testing was completed in a sound-isolated room. Children sat in a chair in front of a desk with a computer monitor positioned directly in front of them. Custom MATLAB software controlled the selection and presentation of test stimuli. All stimuli were mixed (TDT RZ6) and presented diotically via Sennheiser HD25 II supra-aural headphones.

Prior to testing, listeners were familiarized with the pictures in both English and Spanish for bilingual listeners and English only for the monolingual listeners. Listeners were familiarized using recorded speech spoken by a native-English speaker (for the English words) and a native-Spanish speaker (for the Spanish words). These recordings were played to the listeners using an MP3 player coupled to a portable speaker. Immediately following familiarization, listeners were presented with recorded instructions spoken by the target talker from the main experiment. Instructions were presented in English prior to testing in English, and in Spanish prior to testing in Spanish. The recorded instructions explained that the listener should try to listen to the words that she, the recorded talker, spoke. She explained that during testing other competing stimuli would be occurring, but that the listener should try to ignore the competing sounds and choose the corresponding picture of the word they heard her say.

Testing consisted of a four-alternative, forced-choice paradigm that utilized a picture-pointing response (see Figure 2). Each picture was a custom, hand-drawn color illustration of the associated word. The test paradigm was based on methods used by Hall et al. (2002). An adaptive track estimated the disyllabic word identification thresholds corresponding to 70.7% correct identification. Monolingual speakers of English completed two conditions: (1) English targets in a two-talker English masker and (2) English targets in a SSN. Bilingual speakers completed the testing using both English and Spanish target words, for a total of four conditions: (1) English targets in a two-talker English masker, (2) English targets in a SSN, (3) Spanish targets in a two-talker Spanish masker, and (4) Spanish targets in a SSN. Testing was blocked by target language so that listeners could be instructed in the test language prior to testing. The order of the test language was randomized across bilingual subjects. Within the test language, the order of the maskers (two-talker and SSN) was also randomized. The masker stimulus was presented at a fixed level of 60 dB SPL, and the signal level was adjusted based on the listener's responses. The signal level was increased via the MATLAB software after every incorrect response and decreased after two consecutive correct responses (Levitt, 1971). These level adjustments were made in steps of 4 dB until two track reversals had been obtained; steps of 2 dB were used thereafter. Each track continued until a total of eight reversals had been obtained, and threshold estimates were the average signal level at the last six reversals. Threshold estimates for each masker condition were based on the average threshold of two tracks. Each track took approximately 3.5 minutes to administer. If the thresholds of the two tracks were more than 5 dB different,

a third threshold track was obtained, and the threshold estimate was based on the two tracks that had the most similar thresholds. A third track was collected for fewer than 5% of threshold estimates. Visual feedback was provided for the children after each trial, indicating the correct response.

RESULTS

Performance Differences between SSN and Two-Talker Maskers

Individual SRTs for all listeners are shown in Table 4. In the first analysis, data from the monolingual English children were compared between the SSN and two-talker masker conditions. A regression analysis with listener as a random variable was conducted to examine differences in the signal-to-noise ratio (SNR) necessary to achieve similar performance between the two masker types (SSN and two-talker masker). The analysis indicated a significant effect of masker type ($F(1,7) = 46.847, p = .0002, \eta_p^2 = 0.87$). Parameter estimates for the regression model are shown in Table 5. As predicted, the two-talker masker was significantly more effective than the SSN masker, similar to previous reports in the literature (e.g., Hall et al., 2002).

Monolingual vs. Bilingual English Recognition

A second analysis was conducted to compare English SRTs between the bilingual and monolingual control group. Figure 3 summarizes the between-group comparison using boxplots. A mixed-effects regression analysis with listener as a random variable tested the main effect of linguistic group (bilingual and monolingual) and noise type (two-talker and SSN), and the interaction between these two main effects. The analysis indicated a significant main effect of masker type ($F(1,14) = 78.788, p < .0001, \eta_p^2 = 0.849$), but no significant effect of listener group ($F(1,14) = .047, p = .831, \eta_p^2 = 0.003$), nor a significant interaction between these two effects ($F(1,14) = 2.671, p = .124, \eta_p^2 = 0.16$). Parameter estimates for the regression model are shown in Table 6. This result indicates that the difficulty of the English recognition task was similar for the monolingual and bilingual children. The two-talker masker was significantly more challenging than the SSN masker for both groups of children.

Bilingual Children: Performance Comparison between English and Spanish

Group results for Spanish and English speech recognition for the bilingual children are also summarized in Figure 3 for both the two-talker maskers (both English and Spanish) and for the SSN masker. English and Spanish SRTs were compared for the bilingual group. Specifically, a regression analysis with listener as a random variable tested the main effect of target language (English and Spanish), noise type (two-talker and SSN), and the interaction between these two main effects. The analysis indicated a significant main effect of target language ($F(1,7) = 6.516, p = .038, \eta_p^2 = 0.482$) and masker type ($F(1,7) = 65.508, p < .0001, \eta_p^2 = 0.903$). Also, there was a significant interaction between these two effects ($F(1,7) = 6.305, p = .040, \eta_p^2 = 0.474$). Post-hoc testing indicated a simple main effect of target language for the two-talker masker ($p = 0.012$), but not the SSN masker ($p = 0.377$).

Parameter estimates for the regression model are shown in Table 7. This result indicates that for this task, the bilingual children performed significantly better when listening to English than Spanish in competing speech, but performance was similar between the two test languages when listening in competing noise. In addition, similar to the monolingual group, the two-talker maskers were significantly more challenging for the children than the SSN maskers.

The Effect of Age

There is considerable evidence that speech-in-noise recognition improves with increasing age during childhood (e.g., Elliott, Connors, Kille, Levin, Ball & Katz, 1979; McCreery & Stelmachowicz, 2011; Scollie, 2008). This time course of development appears to extend longer into childhood when the competing background sounds are made up of a small number of competing talkers than when the competing background is steady-state noise (e.g., Bonino et al., 2013; Leibold & Buss, 2013; Wightman & Kistler, 2005). That is, preschoolers and kindergarteners are often more susceptible to interference from complex sounds than older children. Consistent with these previous findings, SRTs for the bilingual children for the two-talker maskers were negatively correlated (one tailed) with age for both English ($r(6) = -.91, p = .001$) and Spanish ($r(6) = -.85, p = .004$) target words. A bivariate correlation including both monolingual and bilingual children indicated that English SRTs in the two-talker masker were significantly correlated with age ($r(14) = -.878, p < .0001$), accounting in part for the wide variability in performance scores within the two groups (see Figure 4).

DISCUSSION

The purpose of the present research was to develop and verify stimuli and test procedures that could be used to test children's English and Spanish word recognition in competing speech or noise maskers. As a first step, these stimuli and procedures have been designed for use in the research laboratory, and are available to other researchers upon request. We are in the process of developing a clinical version of this tool. The stimuli described above allow for easy assessment of English and Spanish speech recognition for a pediatric population in competing speech or competing noise utilizing a picture pointing test paradigm. The inclusion of a competing speech masker is novel, and is paramount for two main reasons. First, while it is well documented that children have greater difficulty recognizing speech in competing steady-state noise compared to adults (Elliott, Clifton & Servi, 1983; Hall et al., 2002; Litovsky, 2005; Nittrover & Boothroyd, 1990; Wightman & Kistler, 2005), greater child-adult differences are consistently observed in complex masking conditions, such as a small number of competing talkers (Bonino et al, 2013; Leibold and Buss, 2013; Hall, Buss, Grose & Roush, 2012). Therefore, conventional clinical measures obtained in steady-state noise might underestimate children's performance in everyday listening environments, which often contain competing speech maskers. Second, emerging data indicate that these complex background sounds are particularly challenging for children with hearing loss compared to their peers with normal hearing. Leibold, Hillock-Dunn, Duncan, Roush and Buss (2013) reported that children with sensorineural hearing loss were significantly poorer at identifying spondee words in competing two-talker speech than their normal-hearing

peers. In fact, the gap between the hearing-impaired and normal-hearing children increased from a 3.5-dB-SNR disadvantage in competing speech shaped noise to an 8.1-dB disadvantage in competing two-talker speech.

Another key feature of the protocol introduced in the present study is the ability to administer the test in either English or Spanish. The inclusion of both English and Spanish testing options is necessary to better serve the changing demographics of the United States so that we can 1) gain a better understanding of masked speech perception in children who are second language learners, and 2) assess the growing number of children seen in audiology clinics across the U.S. that have limited English ability. Having a tool with an option to choose the test language will facilitate audiological testing on a case-by-case basis. In addition, using recorded tokens eliminates test administration difficulties for the clinician with respect to speech production in a foreign language. Lastly, using a picture-pointing response removes confusion for scoring foreign language, accented-English, or other atypical speech production responses.

Supplemental Data – Children with hearing loss

One question of future interest is whether the stimuli used in the present experiment capture the greater susceptibility to competing speech (Leibold et al., 2013) and noise (e.g., Elliott et al., 1979, Nittrouer & Boothroyd 1990; Scollie, 2008) previously reported for children with hearing loss. Pilot data on this question were collected on two monolingual English-speaking children and one bilingual child with bilateral, sensorineural hearing loss (Listeners A, B, and C, described in Table 8). All three children were fitted with and consistently use appropriate amplification. Prior to testing, the output of the children's hearing aids were verified according to DSL 5.0a prescriptive targets for a 65-dB-SPL speech input and a maximum pure tone sweep at 90 dB SPL, using the SpeechMap function on the audioscan Verifit. Biological listening checks revealed the hearing aids were free of audible distortion.

The method of testing the three children with hearing loss was similar to that described above, with the exception that these children were tested in the sound field with their hearing aids on, positioned 1 m directly in front of a speaker (JBL Professional, Control Pro 1) and computer monitor. Though similar to the main experiment in which the listeners could see the four alternatives, to minimize listener head movement the child was instructed to either say aloud the token she perceived, or point towards the quadrant of the picture that she heard instead of using a computer mouse to choose the image. This helped to keep the child's head at an equal distance to the speaker throughout the adaptive track. An examiner clicked on the picture corresponding to the word the listener repeated. Individual SRTs for the three listeners with hearing loss are shown in Table 4. Listener C was tested in both English and Spanish. Similar to the bilingual children tested in the main experiment, Listener C was born in the U.S. and exposed to both English and Spanish from birth, mainly hearing and speaking Spanish with her Mother and English with her father. All of her education has been in English. Data from the three hearing-impaired children are included in Figure 3 represented using triangles for the monolingual children and circles for the bilingual child and their listener ID (A, B or C). Similar to results reported in Leibold et al.

(2013), children with hearing loss performed worse than their normal-hearing age matched monolingual peers; however, it is unclear, based on the limited data, whether the two-talker masker caused a greater disadvantage for the children with hearing loss than the speech-shaped noise.

Linguistic Considerations

Over 37 million people living in the U.S. report speaking Spanish in the home (Ryan, 2013). Though the Spanish speaking population in the U.S. has continued to grow in recent years (Ryan, 2013), only 5% of ASHA survey respondents report being qualified to provide bilingual services, with only half of these bilingual clinicians reporting to be Spanish language providers (ASHA, 2012). The purpose of this research was to create a simple task to assess speech recognition in monolingual and bilingual children. Because of the changing demographics of the U.S., we focused on designing the test to include both English and Spanish languages. Primary considerations for reducing linguistic bias were to eliminate the necessity for the tester to produce the speech tokens, and to eliminate the necessity for scoring either accented English speech or Spanish speech.

The eight bilingual listeners we tested in this study differed in their ability to speak English and Spanish, with some parents reporting that their child is able to have simple conversations (one child for Spanish, and one child for English), and others reporting that their child speaks the language all of the time (five children for Spanish and seven children for English). Seven of the eight children were born in the U.S. and never lived outside the country. It is reasonable to assume that, for the majority of these children, English is their dominant language (though Listener 3's mother indicated that she uses Spanish much more regularly than English). Also, these children are "heritage" speakers of Spanish, in that they use Spanish with different types of people and in different situations than when speaking English (Valdes, 2000). Table 2 shows that it is more common for these children to hear Spanish from their Mother and Father, but more common to hear English from siblings, other family members, and people outside of the home. Per parental report, four of the eight indicated speaking both languages equally well, although all but one of the eight listeners performed better in English than in Spanish for both types of maskers. Interestingly, all but one of the eight bilinguals had lower English SRTs (an average of 1.75 dB SNR) in the two-talker masker condition when compared to their age matched monolingual controls.

Six of the eight bilingual children tested in the normal-hearing listener group and Listener C (the bilingual child with hearing loss) would be categorized as simultaneous bilinguals. That is, they began to acquire both languages by the age of two (McLaughlin, 1978). Calandruccio and Zhou (in press) recently reported that sentence recognition performance in the presence of a complex speech masker was similar for a group of simultaneous bilinguals and their monolingual counterparts; a similar finding, though with limited sample sizes, was reported by Mayo et al. (1997) and Shi (2010). These data are in agreement that it is appropriate to use English test materials when evaluating simultaneous bilinguals who were raised in the US and who are heritage speakers of their other language. For these listeners, bilingualism should not negatively affect their English recognition performance in SSN and complex maskers.

Some researchers have argued that for sequential bilinguals speech recognition testing in noise should be completed in one's L1 (Carlo, 2009). Though this may be true for adult listeners with low proficiency in their L2 or those who recently emigrated from their country of origin, a dichotomous rule for which test language to use for all nonnative speakers may not be appropriate. Since children are still in the process of acquiring both languages, it may prove to be even more difficult to determine which language is most appropriate for testing in the audiology clinic for children. In fact, there may not be a definitive answer for which language to use for some children. Being able to conduct speech perception testing in both languages may provide insight for clinicians into a child's communication challenges in either or both of their languages. In addition, for Spanish L1 children this task would allow the clinician to begin their clinical assessment using Spanish and potentially change the test language to English over time as the child's dominant language evolves.

A goal for this test was for a clinician to be able to decide on a case-by-case basis the test language to use, English or Spanish, or in certain circumstances they could have the option to use both languages (e.g., children receiving English education, but predominately speaking Spanish at home). As we move forward it will be important to collect a large set of data on monolingual speakers of English, monolingual speakers of Spanish, and bilingual Spanish/English children with different levels of proficiency in both languages to obtain normative clinical data for this tool. In addition, we are presently collecting data using a hand-held touch-screen device, to reduce effects of introducing a full-sized computer monitor into the sound field.

CONCLUSIONS

1. A four-alternative forced-choice word recognition test was designed to use in both English and Spanish. A listener's SRT can be obtained in either Spanish or English, in the presence of speech shaped noise or a two-talker masker in either language.
2. Bilingual children performed better in English than in Spanish in competing speech maskers, but indicated similar performance for the two languages in competing noise. Bilingual and monolingual children's performance on English word recognition was similar.
3. As observed previously, children in all conditions required a higher SNR in the two-talker masker than in the speech-shaped noise masker. SNR was significantly correlated with age, in which younger children required a more favorable SNR.

Acknowledgments

Funded by the NIH Grant number R01 DC011038 (LJL). Portions of these data were presented at the 2013 American Auditory Society meeting in Scottsdale, AZ. We are thankful to Joan Calandruccio for the time she spent drawing the illustrations used in this project and to Dr. Barbara Rodriguez for helping us better understand language proficiency measures for bilingual children.

References

American National Standards of the Acoustical Society of America. S3.6: American national standard specification for audiometers. 2010.

- American Speech-Language-Hearing Association. Guidelines for the audiological assessment of children from birth to 5 years of age [guidelines]. 2004. Retrieved from www.asha.org/policy
- American Speech-Language-Hearing Association. 2012 audiology survey report: Clinical focus patterns. 2012. Retrieved from www.asha.org
- Armbruster, BB.; Lehr, F.; Osborn, J. Put reading first: The research building blocks for teaching children to read. Kindergarten through grade 3. Washington, DC: National Institute for Literacy; 2001. Retrieved from (available online at www.nifl.gov)
- Baayen, RH.; Piepenbrock, R.; van Rijnm, H. The CELEX lexical database. Release 2 [CD-ROM]. University of Pennsylvania; Philadelphia: 1995. Linguistic Data Consortium
- Bird H, Franklin S, Howard D. Age of acquisition and imageability ratings for a large set of words, including verbs and function words. Behavior Research Methods, Instruments, & Computers. 2001; 33(1):73–79.
- Boersma, P.; Weeknick, D. Praat: Doing phonetics by computer [computer program. 2012. (Version 5.3.23 ed.) Retrieved from <http://www.praat.org>
- Bonino AY, Leibold LJ, Buss E. Release from perceptual masking for children and adults: Benefit of a carrier phrase. Ear and Hearing. 2013; 34(1):3–14. [PubMed: 22836239]
- Brouwer S, Van Engen KJ, Calandruccio L, Bradlow AR. Linguistic contributions to speech-on-speech masking for native and non-native listeners: Language familiarity and semantic content. The Journal of the Acoustical Society of America. 2012; 131(2):1449–1464. [PubMed: 22352516]
- Brungart DS, Chang PS, Simpson BD, Wang D. Multitalker speech perception with ideal time-frequency segregation: Effects of voice characteristics and number of talkers. The Journal of the Acoustical Society of America. 2009; 125(6):4006–4022. [PubMed: 19507982]
- Calandruccio L, Zhou H. Increase in speech recognition due to linguistic mismatch between target and masker speech: Monolingual and simultaneous bilingual performance. Journal of Speech, Language, and Hearing Research. (In press).
- Calandruccio L, Dhar S, Bradlow AR. Speech-on-speech masking with variable access to the linguistic content of the masker speech. The Journal of the Acoustical Society of America. 2010; 128(2): 860–869. [PubMed: 20707455]
- Carhart R, Tillman TW. Interaction of competing speech signals with hearing losses. Archives of Otolaryngology (Chicago, Ill: 1960). 1970; 91(3):273–279.
- Carhart R, Tillman TW, Greetis ES. Perceptual masking in multiple sound backgrounds. The Journal of the Acoustical Society of America. 1969; 45(3):694–703. [PubMed: 5776931]
- Carlo, M. Spanish-English bilingual speech perception in noise. The Words-in-Noise Test (WIN): English and Spanish; ASHA Conference; 2009.
- Chermak GD, Pederson CM, Bendel RB. Equivalent forms and split-half reliability of the NU-CHIPS administered in noise. The Journal of Speech and Hearing Disorders. 1984; 49(2):196–201. [PubMed: 6716990]
- Chermak GD, Wagner DP, Bendel RB. Interlist equivalence of the word intelligibility by picture identification test administered in broad-band noise. Audiology : Official Organ of the International Society of Audiology. 1988; 27(6):324–333. [PubMed: 3240130]
- Coltheart M. The MRC psycholinguistic database. The Quarterly Journal of Experimental Psychology Section A. 1981; 33(4):497–505.
- Crandell C, Smaldino J. Speech perception in noise by children for whom English is a second language. American Journal of Audiology. 1996; 5:47–51.
- Davis N-watch: A program for deriving neighborhood size and other psycholinguistic statistics. Behavior Research Methods. 2005; 37(1):65–70. [PubMed: 16097345]
- Davis CJ, Perea M. BuscaPalabras: A program for deriving orthographic and phonological neighborhood statistics and other psycholinguistic indices in Spanish. Behavior Research Methods. 2005; 37(4):665–671. [PubMed: 16629300]
- Dolch, EW. Problems in reading. Champaign, Ill: Garrard Press; 1948. Retrieved from /z-wcorg/
- Elliot LL. Performance of children aged 9 to 17 years on a test of speech intelligibility in noise using sentence material with controlled word predictability. The Journal of the Acoustical Society of America. 1979; 66(3):651–653. [PubMed: 489836]

- Elliott LL, Clifton LA, Servi DG. Word frequency effects for a closed-set word identification task. *Audiology : Official Organ of the International Society of Audiology*. 1983; 22(3):229–240. [PubMed: 6870665]
- Elliott LL, Connors S, Kille E, Levin S, Ball K, Katz D. Children's understanding of monosyllabic nouns in quiet and in noise. *The Journal of the Acoustical Society of America*. 1979; 66(1):12–21. [PubMed: 489827]
- Elliott, LL.; Katz, D. Development of a new children's test of speech discrimination (technical manual). St. Louis, MO: Auditec; 1980.
- Fenson, L.; Marchman, VA.; Dale, PS.; Reznick, JS.; Thal, D.; Bates, E. *The MacArthur-Bates Communicative Development Inventories: User's guide and technical manual*. 2. Baltimore: Brookes; 2007.
- Festen JM, Plomp R. Effects of fluctuating noise and interfering speech on the speech-reception threshold for impaired and normal hearing. *The Journal of the Acoustical Society of America*. 1990; 88(4):1725–1736. [PubMed: 2262629]
- Freyman RL, Balakrishnan U, Helfer KS. Effect of number of masking talkers and auditory priming on informational masking in speech recognition. *The Journal of the Acoustical Society of America*. 2004; 115(5 Pt 1):2246–2256. [PubMed: 15139635]
- Garstecki DC, Wilkin MK. Linguistic background and test material considerations in assessing sentence identification ability in English- and Spanish-English-speaking adolescents. *Journal of the American Audiology Society*. 1976; 1(6):263–268. [PubMed: 931762]
- Gat IB, Keith RW. An effect of linguistic experience. auditory word discrimination by native and non-native speakers of English. *Audiology : Official Organ of the International Society of Audiology*. 1978; 17(4):339–345. [PubMed: 687239]
- Grant KW, Walden TC. Understanding excessive SNR loss in hearing-impaired listeners. *Journal of the American Academy of Audiology*. 2013; 24(4):258–73. quiz 337–338. [PubMed: 23636208]
- Gutierrez-Clellen VF, Kreiter J. Understanding child bilingual acquisition using parent and teacher reports. *Applied Psycholinguistics*. 2003; 24(02):267–288.
- Hall JW, Buss E, Grose JH, Roush PA. Effects of age and hearing impairment on the ability to benefit from temporal and spectral modulation. *Ear and Hearing*. 2012; 33(3):340–348. [PubMed: 22237164]
- Hall JW 3rd, Grose JH, Buss E, Dev MB. Spondee recognition in a two-talker masker and a speech-shaped noise masker in adults and children. *Ear and Hearing*. 2002; 23(2):159–165. [PubMed: 11951851]
- International Organization for Standardization (ISO). ISO 28961: 2012: Acoustics -- Statistical distribution of hearing thresholds of otologically normal persons in the age range from 18 years to 25 years under free-field listening conditions. 2012.
- Jerger J, Speaks C, Trammell JL. A new approach to speech audiometry. *The Journal of Speech and Hearing Disorders*. 1968; 33(4):318–328. [PubMed: 5696322]
- Jorgensen RN, Dale PS, Bleses D, Fenson L. CLEX: A cross-linguistic lexical norms database. *Journal of Child Language*. 2010; 37(2):419–428. [PubMed: 19570318]
- Leibold LJ, Buss E. Children's identification of consonants in a speech-shaped noise or a two-talker masker. *Journal of Speech, Language, and Hearing Research : JSLHR*. 2013; 56(4):1144–1155.
- Leibold LJ, Hillock-Dunn A, Duncan N, Roush PA, Buss E. Influence of hearing loss on children's identification of spondee words in a speech-shaped noise or a two-talker masker. *Ear and Hearing*. 2013; 34(5):575–584. [PubMed: 23492919]
- Levitt H. Transformed up-down methods in psychoacoustics. *The Journal of the Acoustical Society of America*. 1971; 49 Suppl 2(2):467. [PubMed: 5541744]
- Litovsky RY. Speech intelligibility and spatial release from masking in young children. *The Journal of the Acoustical Society of America*. 2005; 117(5):3091–3099. [PubMed: 15957777]
- Mayo LH, Florentine M, Buus S. Age of second-language acquisition and perception of speech in noise. *Journal of Speech, Language, and Hearing Research : JSLHR*. 1997; 40(3):686–693.
- McLaughlin, B. *Second-language acquisition in childhood*. Hillsdale, NJ: Lawrence-Erlbaum Associates; 1978.

- McCreery RW, Stelmachowicz PG. Audibility-based predictions of speech recognition for children and adults with normal hearing. *The Journal of the Acoustical Society of America*. 2011; 130(6): 4070–4081. [PubMed: 22225061]
- Nelson P, Kohnert K, Sabur S, Shaw D. Classroom noise and children learning through a second language: Double jeopardy? *Language, Speech, and Hearing Services in Schools*. 2005; 36(3): 219–229.
- Nicoladis E, Genesee F. Language development in preschool bilingual children. [L' apprentissage du langage chez les enfants bilingues d' age prescolaire]. *Journal of Speech Language-Pathology and Audiology*. 1997; 21(4):258–270.
- Nilsson M, Soli SD, Sullivan JA. Development of the hearing in noise test for the measurement of speech reception thresholds in quiet and in noise. *The Journal of the Acoustical Society of America*. 1994; 95(2):1085–1099. [PubMed: 8132902]
- Nittrouer S, Boothroyd A. Context effects in phoneme and word recognition by young children and older adults. *The Journal of the Acoustical Society of America*. 1990; 87(6):2705–2715. [PubMed: 2373804]
- Rimikis S, Smiljanic R, Calandruccio L. Nonnative English speaker performance on the basic English lexicon (BEL) sentences. *Journal of Speech, Language, and Hearing Research : JSLHR*. 2013; 56(3):792–804.
- Rogers C, Lister J, Febo D, Besing J, Abrams H. Effects of bilingualism, noise, and reverberation on speech perception by listeners with normal hearing. *Applied Psycholinguistics*. 2006; 27(3):465.
- Rosen S, Souza P, Ekelund C, Majeed AA. Listening to speech in a background of other talkers: Effects of talker number and noise vocoding. *The Journal of the Acoustical Society of America*. 2013; 133(4):2431–2443. [PubMed: 23556608]
- Ross M, Lerman J. A picture identification test for hearing-impaired children. *Journal of Speech and Hearing Research*. 1970; 13(1):44–53. [PubMed: 4192711]
- Ryan. American Community Survey Reports. Language use in the United States: 2011. U.S. Census Bureau; 2013.
- Scollie SD. Children's speech recognition scores: The Speech Intelligibility Index and proficiency factors for age and hearing level. *Ear and Hearing*. 2008; 29(4):543–556. [PubMed: 18469717]
- Sebastian-Galles, N.; Marti, MA.; Cuetos, F.; Carreiras, M. LEXESP: Léxico informatizado del español. Barcelona: Edicions de la Universitat de Barcelona; 2000.
- Shi LF. Normal-hearing English-as-a-second-language listeners' recognition of English words in competing signals. *International Journal of Audiology*. 2009; 48(5):260–270. [PubMed: 19842801]
- Shi LF. Perception of acoustically degraded sentences in bilingual listeners who differ in age of English acquisition. *Journal of Speech, Language, and Hearing Research : JSLHR*. 2010; 53(4): 821–835.
- Simpson SA, Cooke M. Consonant identification in N-talker babble is a nonmonotonic function of N. *The Journal of the Acoustical Society of America*. 2005; 118(5):2775–2778. [PubMed: 16334654]
- Soli SD, Wong LL. Assessment of speech intelligibility in noise with the hearing in noise test. *International Journal of Audiology*. 2008; 47(6):356–361. [PubMed: 18569108]
- Sommers MS, Kirk KI, Pisoni DB. Some considerations in evaluating spoken word recognition by normal-hearing, noise-masked normal-hearing, and cochlear implant listeners. I: The effects of response format. *Ear and Hearing*. 1997; 18(2):89–99. [PubMed: 9099558]
- Storkel HL, Hoover JR. An online calculator to compute phonotactic probability and neighborhood density on the basis of child corpora of spoken American English. *Behavior Research Methods*. 2010; 42(2):497–506. [PubMed: 20479181]
- Valdes, G. Spanish for native speakers: AATSP professional development series handbook for teachers K-16. New York, NY: Harcourt College Publishers; 2000. p. 1 Vol. 1 ed
- von Hapsburg D, Pena ED. Understanding bilingualism and its impact on speech audiometry. *Journal of Speech, Language, and Hearing Research : JSLHR*. 2002; 45(1):202–213.
- Walker, R. Jack and the beanstalk. Cambridge, MA: Barefoot Books; 1999a.
- Walker, R. Juan y los frijoles magicos. Sarfatti, E., translator. Cambridge, MA: Barefoot Books; 1999b.

- Wightman FL, Kistler DJ. Informational masking of speech in children: Effects of ipsilateral and contralateral distracters. *The Journal of the Acoustical Society of America*. 2005; 118(5):3164–3176. [PubMed: 16334898]
- Wilson RH, McArdle RA, Smith SL. An evaluation of the BKB-SIN, HINT, QuickSIN, and WIN materials on listeners with normal hearing and listeners with hearing loss. *Journal of Speech, Language, and Hearing Research : JSLHR*. 2007; 50(4):844–856.10.1044/1092-4388(2007/059)

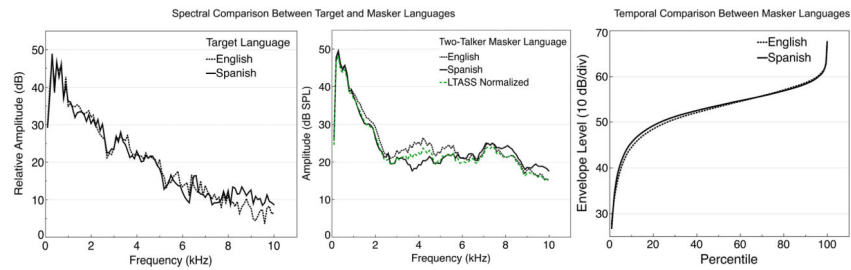


Figure 1.

[color online]. Panel A: spectral comparisons between the average LTASS of the 30 disyllabic words used for testing for the two target languages. Panel B and C: spectral and temporal comparisons, respectively, of the two-talker Spanish masker and the twotalker English masker. Panel B illustrates LTASS comparisons of the two two-talker maskers, as well as the spectrum of the LTASS-normalized competing speech and SSN maskers. Panel C illustrates the cumulative distribution of filtered envelope values for the two two-talker maskers, indicating masker envelope minima that were available to the listener. A larger proportion of relatively low envelope values indicate a greater opportunity for dip listening.



Figure 2.
[color online]. Example of four-alternative forced choice picture pointing with hand-drawn illustrations of lion/leon, woman/mujer, garden/jardin, and chicken/pollo target words.

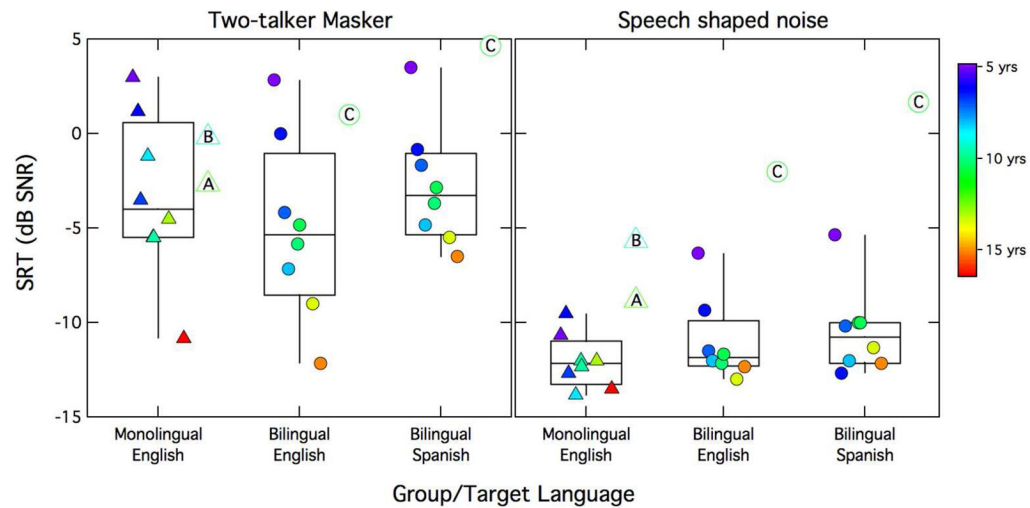


Figure 3.

[color online]. SRTs for bilingual and monolingual children in the presence of two-talker maskers and SSN. The height of the box indicates the interquartile range of performance scores, while the intermediate horizontal line indicates the median. The whiskers are calculated using the following two formulae: upper whisker = 3rd quartile + 1.5*(interquartile range), lower whisker = 1st quartile – 1.5*(interquartile range). Individual data points are shown for the 16 children, color-coded to indicate their age at testing, and organized along the abscissa by age. Pilot data from two monolingual English children and one bilingual English/Spanish child with bilateral sensorineural hearing loss (Listeners A, B and C described within Discussion section) are also shown.

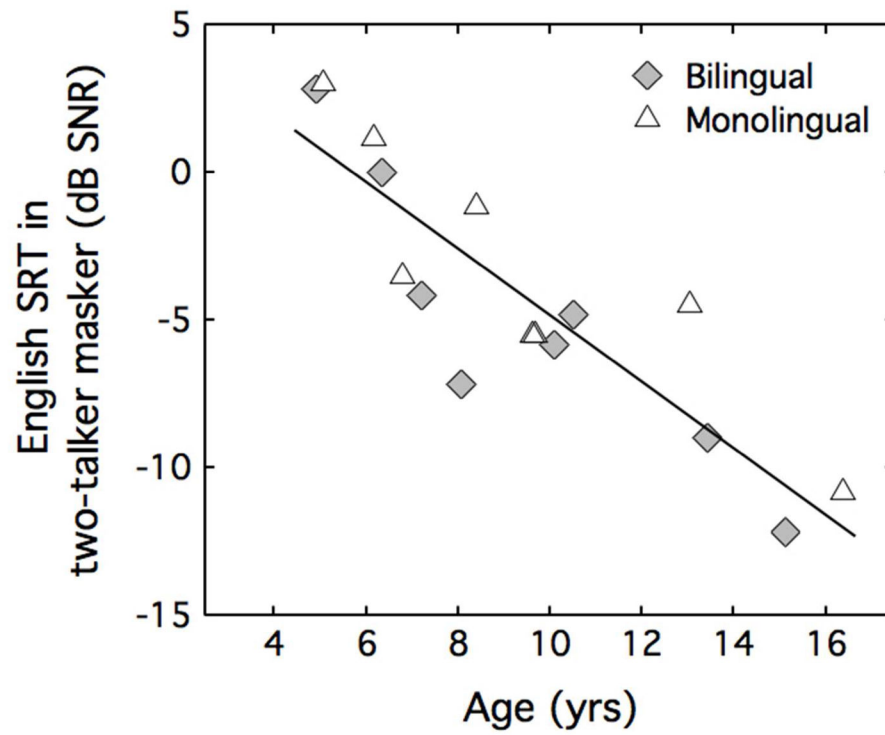


Figure 4. Correlation between English SRT scores in the presence of the English two-talker masker for both groups of listeners (bilingual and monolingual children) and age at testing. SRTs consistently decreased as listeners got older ($r = -0.88$).

Table 1

The 30 disyllabic English and Spanish words used for testing and corresponding lexical information for each token, including lexical frequency, number of phonological neighbors, familiarity ratings, imageability ratings, cross linguistic lexical normative data, and age of acquisition data (English only). A ‘-’ indicates that the word was not available within the database searched.

| | Lexical frequency (log10) | | Number of Phonological Neighbors | | Familiarity | | Imageability | | Kuperman et al. Age of Acquisition (age) in years | |
|----------|---------------------------|---------|----------------------------------|---------|-------------|---------|--------------|---------|---|---------|
| | English | Spanish | English | Spanish | English | Spanish | English | Spanish | English | Spanish |
| paper | 2.24 | 2.27 | 12 | 2 | 635 | 666 | 590 | 634 | 4.00 | 4.00 |
| chicken | 1.5 | 1.11 | 1 | 18 | 544 | 626 | 619 | 651 | 3.26 | 3.26 |
| water | 2.64 | 2.47 | 13 | 2 | 641 | 680 | 632 | 637 | 2.37 | 2.37 |
| table | 2.31 | 2.24 | 10 | 17 | 599 | 658 | 582 | 677 | 4.39 | 4.39 |
| children | 2.82 | 2.29 | 0 | 5 | 608 | 648 | 597 | 639 | 4.10 | 4.10 |
| garden | 2.05 | 1.8 | 3 | 0 | 567 | 621 | 635 | 586 | 5.33 | 5.33 |
| baby | 2.27 | 1.25 | 4 | 7 | 597 | 604 | 608 | 681 | 3.84 | 3.84 |
| flower | 1.46 | 1.53 | 3 | 3 | 566 | 674 | 618 | 625 | 3.11 | 3.11 |
| oven | 1.28 | 0.91 | 4 | 4 | 577 | 652 | 599 | 643 | 5.67 | 5.67 |
| hanger | 0.25 | 0.74 | 2 | 10 | - | 537 | - | 490 | 6.78 | 6.78 |
| tiger | 0.99 | 0.72 | 5 | 1 | 513 | 554 | 606 | 633 | 4.00 | 4.00 |
| monkey | 1 | 1.28 | 6 | 21 | 531 | 585 | 588 | 656 | 4.21 | 4.21 |
| doctor | 2.13 | 1.96 | 2 | 2 | 573 | 603 | 600 | 544 | 4.60 | 4.60 |
| pencil | 1.22 | 0.9 | 3 | 1 | 598 | 616 | 607 | 592 | 4.06 | 4.06 |
| candy | 0.86 | 1.6 | 10 | 0 | 447 | - | 575 | - | 4.00 | 4.00 |
| monster | 1.19 | 1.34 | 2 | 0 | - | 606 | - | 575 | 4.58 | 4.58 |
| balloon | 0.63 | 1.05 | 4 | 1 | 520 | 560 | 583 | 622 | 4.37 | 4.37 |
| sweater | 1.08 | 0.35 | 4 | 0 | - | - | - | - | 5.22 | 5.22 |
| necklace | 0.52 | 0.86 | 2 | 7 | 536 | 582 | 606 | 655 | 5.00 | 5.00 |
| dolphin | 0.37 | 0.74 | 1 | 1 | - | - | 626 | - | 6.05 | 6.05 |
| button | 1.25 | 1.16 | 6 | 4 | 573 | 649 | 580 | 649 | 4.78 | 4.78 |
| zebra | 0.36 | 0.28 | 1 | 2 | - | - | 648 | - | 4.79 | 4.79 |
| lemon | 1.15 | 0.88 | 2 | 3 | 518 | - | 632 | - | 4.74 | 4.74 |
| turkey | 1.16 | 0.85 | 5 | 20 | - | 474 | 408 | 479 | 3.95 | 3.95 |

| | Lexical frequency (log ₁₀) | | | Number of Phonological Neighbors | | | Familiarity | | Imageability | | Kupermanetal Age of Acquisition (age) in years | |
|-------------|--|---------------|---------------|----------------------------------|----------------|----------------|----------------|----------------|---------------|---------|--|---------|
| | English | Spanish | English | Spanish | English | Spanish | English | Spanish | English | Spanish | English | Spanish |
| dragon | 0.93 | 0.66 | 1 | 3 | - | 417 | 568 | 586 | 5.58 | | | |
| feather | 1.05 | 1.23 | 10 | 1 | - | 576 | - | 590 | 4.67 | | | |
| lion | 1.25 | 1.49 | 7 | 7 | 511 | 485 | 626 | 672 | 4.42 | | | |
| elbow | 1.22 | 0.93 | 2 | 19 | 564 | 563 | 602 | 427 | 4.78 | | | |
| ruler | 0.94 | 1.44 | 5 | 2 | 571 | 636 | 543 | 550 | 5.94 | | | |
| woman | 2.53 | 2.69 | 2 | 0 | 623 | 673 | 626 | 661 | 4.95 | | | |
| mean | 1.36 | 1.30 | 4.40 | 5.43 | 564 | 598 | 596 | 606 | 4.6 | | | |
| std | (0.71) | (0.63) | (3.48) | (6.66) | (46.64) | (66.99) | (45.23) | (65.57) | (0.91) | | | |

Table 2

Speaking rate (syllables/second) for the English/Spanish bilingual talker used to produce the 30 disyllabic target words in both languages. Overall duration of speech produced, number of words/second, and number of syllables/second for the two English/Spanish bilingual talkers used to create the two-talker English and two-talker Spanish maskers.

| | English | | | Spanish | | |
|-------------|------------------|---------------|---------------|----------------|---------------|---------------|
| | Duration (sec) | words/sec | syllables/sec | Duration (sec) | words/sec | syllables/sec |
| Masker | | | | | | |
| Talker 1 | 181 | 3.47 | 4.18 | 178 | 2.64 | 5.00 |
| Talker 2 | 169 | 4.05 | 4.89 | 203 | 2.81 | 4.58 |
| mean | 175.00 | 3.76 | 4.54 | 190.50 | 2.73 | 4.79 |
| std | (8.49) | (0.41) | (0.5) | (17.68) | (0.12) | (0.3) |
| | syllables/sec | | | | | |
| Target | Disyllabic Word | English | Spanish | | | |
| | paper/papel | 0.24 | 0.24 | | | |
| | chicken/pollo | 0.26 | 0.21 | | | |
| | water/agua | 0.24 | 0.25 | | | |
| | table/mesa | 0.26 | 0.31 | | | |
| | children/niños | 0.32 | 0.40 | | | |
| | garden/jardín | 0.26 | 0.33 | | | |
| | baby/bebé | 0.24 | 0.22 | | | |
| | flower/ flores | 0.29 | 0.39 | | | |
| | oven/horno | 0.21 | 0.22 | | | |
| | hanger/gancho | 0.26 | 0.29 | | | |
| | tiger/tigre | 0.26 | 0.21 | | | |
| | monkey/mono | 0.27 | 0.23 | | | |
| | doctor/doctor | 0.31 | 0.27 | | | |
| | pencil/lápiz | 0.27 | 0.37 | | | |
| | candy/dulce | 0.29 | 0.33 | | | |
| | monster/monstruo | 0.33 | 0.33 | | | |
| | balloon/globo | 0.28 | 0.23 | | | |
| | sweater/suéter | 0.32 | 0.32 | | | |

| | English | | | Spanish | | |
|-----------------|----------------|---------------|---------------|----------------|-----------|---------------|
| | Duration (sec) | words/sec | syllables/sec | Duration (sec) | words/sec | syllables/sec |
| necklace/collar | 0.34 | 0.26 | | | | |
| dolphin/delfin | 0.27 | 0.36 | | | | |
| button/botón | 0.22 | 0.31 | | | | |
| zebra/cebra | 0.33 | 0.30 | | | | |
| lemon/limón | 0.26 | 0.27 | | | | |
| turkey/pavo | 0.25 | 0.17 | | | | |
| dragon/dragón | 0.29 | 0.26 | | | | |
| feather/pluma | 0.21 | 0.19 | | | | |
| lion/león | 0.32 | 0.25 | | | | |
| elbow/codo | 0.27 | 0.20 | | | | |
| ruler/regla | 0.30 | 0.25 | | | | |
| woman/mujer | 0.28 | 0.23 | | | | |
| mean | 0.28 | 0.27 | | | | |
| std | (0.03) | (0.06) | | | | |

Table 3

Demographics of the eight bilingual children.

| Listener ID | Age at Testing (yrs) | Sex | Age Exposed to English (mos) | Pre-K Language | | Parental Level of Education | | Speaks | | Understands | | Hears Spanish From | | Hears English From | | | | |
|-------------|----------------------|-----|------------------------------|-------------------|---------------|-----------------------------|---------|---------|---------|-------------|---|--------------------|---|--------------------|-----|---|---|-----|
| | | | | Mother | Father | English | Spanish | English | Spanish | M | F | S | O | M | F | S | O | |
| 1 | 4.9 | F | 0 | English & Spanish | Post-Graduate | Post-Graduate | 5 | 5 | 5 | 5 | 5 | 3 | 2 | 3 | 3 | 3 | | |
| 2 | 6.3 | F | 2 | English & Spanish | Elementary | Elementary | 5 | 4 | 4 | 4 | 4 | 4 | 4 | 2 | 1 | 2 | 4 | |
| 3 | 7.2 | F | 12 | English & Spanish | Secondary | DNR | 3 | 5 | 5 | 5 | 5 | 4 | 4 | 2 | 1 | 4 | 1 | 2 |
| 4 | 8.1 | F | 0 | English | Post-Graduate | Secondary | 5 | 5 | 5 | 5 | 5 | 2 | 4 | 2 | 4 | 1 | 4 | 4 |
| 5 | 10.1 | M | 0 | N/A | Post-Graduate | Post-Graduate | 5 | 3 | 5 | 4 | 4 | 3 | 3 | 2 | 2 | 2 | 3 | 4 |
| 6 | 10.5 | F | 48 | English & Spanish | Elementary | Elementary | 5 | 5 | 5 | 4 | 4 | 4 | 4 | 2 | 3 | 1 | 2 | 4 |
| 7 | 13.4 | M | 60 | N/A | University | University | 5 | 5 | 5 | 5 | 5 | 3 | 3 | DNR | DNR | 2 | 2 | DNR |
| 8 | 15.1 | M | 0 | DNR | Post-Graduate | Post-Graduate | 5 | 4 | 5 | 4 | 4 | 2 | 3 | 1 | 1 | 3 | 2 | 4 |

Age at testing, sex (M = male; F = female), age at first exposure to English, and preschool language are reported. Parental report with respect to *Speaks English/Spanish*: 5 = speaks all of the time, 4 = speaks frequently, 3 = can have a simple conversation, 2 = says a couple of words or phrases, 1 = does not speak any. With respect to *Understands English/Spanish*: 5 = understands everything they are told, 4 = understands the majority of what they are told, 3 = understands basic directions, 2 = understands a couple of words or phrases, 1 = does not understand anything. With respect to *Hears English/Spanish from*: M = mother, F = Father, S = siblings/other family members, O = other; 1 = never, 2 = sometimes, 3 = most of the time, 4 = all of the time. N/A = not applicable; DNR = did not respond.

Table 4

The SNR associated with each listener's SRT for each masker condition (twotalker and spectrally shaped noise). The linguistic group and age at testing is also reported, as well as pilot data from three children with hearing loss (reported in the Discussion section).

| Group | Listener | Age (yrs) | English | | Spanish | |
|-------------|----------|-----------|----------|--------|----------|--------|
| | | | 2-talker | SSN | 2-talker | SSN |
| Bilingual | 1 | 4.9 | 2.83 | -6.33 | 3.50 | -5.33 |
| | 2 | 6.3 | 0.00 | -9.33 | -0.83 | -12.67 |
| | 3 | 7.2 | -4.17 | -11.50 | -1.67 | -10.17 |
| | 4 | 8.1 | -7.17 | -12.00 | -4.83 | -12.00 |
| | 5 | 10.1 | -5.83 | -12.17 | -3.67 | -10.00 |
| | 6 | 10.5 | -4.83 | -11.67 | -2.83 | -10.00 |
| | 7 | 13.4 | -9.00 | -13.00 | -5.50 | -11.33 |
| | 8 | 15.1 | -12.17 | -12.33 | -6.50 | -12.17 |
| | mean | 9.45 | -5.04 | -11.04 | -2.79 | -10.46 |
| | std | (3.52) | (4.78) | (2.18) | (3.18) | (2.32) |
| Monolingual | 9 | 5.1 | 3.00 | -10.67 | | |
| | 10 | 6.2 | 1.67 | -9.50 | | |
| | 11 | 6.8 | -3.50 | -12.67 | | |
| | 12 | 8.4 | -1.17 | -13.83 | | |
| | 13 | 9.6 | -5.50 | -12.00 | | |
| | 14 | 9.7 | -5.50 | -12.33 | | |
| | 15 | 13.0 | -4.50 | -12.00 | | |

| Group | English | | | Spanish | | |
|------------------|----------|-----------|--------------|--------------|--------------|------|
| | Listener | Age (yrs) | 2-talker SSN | 2-talker SSN | 2-talker SSN | SSN |
| | 16 | 16.4 | -10.83 | -13.50 | | |
| | mean | 9.40 | -3.29 | -12.06 | | |
| | std | (3.75) | (4.42) | (1.42) | | |
| Hearing Impaired | A | 9.3 | -2.66 | -8.83 | | |
| | B | 11.9 | -0.17 | -5.66 | | |
| | C | 10.3 | 1.00 | -2.00 | 4.67 | 1.67 |

Table 5

Parameter estimates for the regression model analyzing data from monolingual English children between the SSN and two-talker masker conditions.

| Effect | Estimate | Std. Error | Prob t |
|-------------|----------|------------|--------|
| Intercept | -7.708 | 0.949 | <.0001 |
| Masker Type | -4.35 | 0.636 | .0002 |

Table 6

Parameter estimates for the regression model analyzing data for monolingual English and bilingual English/Spanish children between the SSN and two-talker masker conditions.

| Effect | Estimate | Std. Error | Prob t |
|--------------------------------|-----------------|-------------------|---------------|
| Intercept | -7.875 | 0.766 | <.0001 |
| Linguistic Group | -0.167 | 0.766 | 0.831 |
| Masker Type | -3.677 | 0.414 | <.0001 |
| Linguistic Group * Masker Type | 0.677 | 0.414 | 0.124 |

Table 7

Parameter estimates for the regression model analyzing data from bilingual English/Spanish children between the SSN and two-talker masker conditions for both English and Spanish languages.

| Effect | Estimate | Std. Error | Prob t |
|---------------------------------|-----------------|-------------------|---------------|
| Intercept | -7.333 | 1.031 | 0.0002 |
| Stimulus Language | -0.708 | 0.277 | 0.0380 |
| Masker Type | -3.417 | 0.422 | <.0001 |
| Stimulus Language * Masker Type | 0.417 | 0.166 | 0.0403 |

Table 8

Pilot data were collected for Listeners A, B, and C, three children with sensorineural hearing loss. Their listener ID, age at testing, linguistic status (monolingual or bilingual), configuration and severity of hearing loss, age of initial hearing loss diagnosis, and age fitted with amplification are reported.

| Listener ID | Age (yrs) | Bilingual | Shape and Severity of Hearing Loss | Age of Dx (yrs) | Age fit with HA (yrs) |
|-------------|-----------|-----------|---|-----------------|-----------------------|
| A | 9.4 | N | mild through 1 kHz, precipitously sloping to moderate-severe at 1.5 kHz through 8 kHz, bilaterally | 5.5 | 5.6 |
| B | 11.9 | N | mild through 3 kHz bilaterally, rising to normal hearing through 8 kHz in left ear, and through 6 kHz in right ear dropping to mild at 8 kHz in right ear | 3.5 | 4 |
| C | 10.3 | Y | mild-to-moderate at 250 Hz sloping to moderate-severe at 1 kHz through 8 kHz, bilaterally | 0.25 | 0.33 |