AJA

Research Article

Spanish Pediatric Picture Identification Test

Lisa Lucks Mendel,[a] Monique A. Pousson,[a] Johnnie K. Bass,[b] Jordan Alyse Coffelt,[c] Melanie Morris,[d] and Kati A. Lane[e]

Purpose: The purpose of this study was to construct and validate a recorded word recognition test for monolingual Spanish-speaking children utilizing a picture board and a picture-pointing task.
Design: The Spanish Pediatric Picture Identification Test was developed and validated in this study. Test construction steps included (a) producing new digital recordings of word lists created by Comstock and Martin (1984) using a bilingual Spanish–English female, (b) obtaining list equivalency, (c) creating digitally illustrated pictures representing the word lists, (d) validating the pictures using monolingual Spanish-speaking and bilingual Spanish–English children, and (e) re-establishing list equivalency and obtaining performance–intensity functions using a picture-pointing task with monolingual Spanish-speaking children and bilingual Spanish–English adults.
Results: Normative data for three Spanish word recognition lists were established. Performance–intensity functions at sensation levels from 0 to 40 dB SL in 8-dB steps were obtained, establishing list equivalency for Lists 1, 2, and 3.
Conclusions: The Spanish Pediatric Picture Identification Test was developed and validated as a picture-pointing task for word recognition with monolingual Spanish-speaking children. The two validated channel recordings include an English translation for ease of testing by clinicians lacking Spanish language skills. Future validation will be conducted with bilingual Spanish–English children with normal hearing and with hearing loss.

In audiologic evaluations, assessment of speech understanding ability is crucial to making conclusions regarding a patient's auditory status. There are many types of audiologic tests that provide assessments of speech understanding, with the most common tests being the speech recognition threshold (SRT) and word recognition testing. While the SRT provides information regarding the lowest level one can repeat spondaic words, suprathreshold word recognition tests provide an estimate of speech understanding based on the number of words identified correctly at a level well above the patient's hearing threshold. Word recognition testing serves a variety of purposes, including, but not limited to, describing the extent of a hearing loss and its effect on speech understanding, assisting in differential diagnosis of various disorders, determining the need for amplification, and verifying benefit from hearing instruments and/or audiologic rehabilitation (Gelfand, 2016; McArdle & Hnath-Chisolm, 2009).

According to the 2017 U.S. Census Bureau (n.d.), there were over 37 million people over the age of 5 years who spoke Spanish at home. With this large population of Spanish speakers in the United States, there is a demand for validated Spanish speech perception materials to be used with both monolingual Spanish-speaking and bilingual Spanish–English individuals during audiologic evaluations. Audiologic test materials should be appropriate for the age, education, and linguistic background of the listener; therefore, an appropriate audiologic assessment consists of procedures and test materials that are developmentally and culturally appropriate and free of any cultural bias (American Speech-Language-Hearing Association, 2002; Davis, 1978; Desjardins et al., 2019; Gaeta & John, 2015). Characteristics of stimuli used in such assessments are likely to influence performance on word recognition tasks. 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.

The focus of this article is the description of the development of a picture identification word recognition test

[a]School of Communication Sciences and Disorders, The University of Memphis, TN
[b]St. Jude Children's Research Hospital, Memphis, TN
[c]Memphis Hearing Aid, Germantown, TN
[d]Mobile Audiology Program, Georgia Department of Education, Atlanta
[e]Huntsville Hospital for Women & Children, AL

Correspondence to Lisa Lucks Mendel: lmendel@memphis.edu

Editor-in-Chief: Sumitrajit Dhar
Editor: Kathy R. Vander Werff

Received September 9, 2019
Revision received November 13, 2019
Accepted March 7, 2020
https://doi.org/10.1044/2020_AJA-19-00049

Disclosure: The authors have declared that no competing interests existed at the time of publication.

using Spanish stimuli. The data presented here are primarily from monolingual Spanish speakers, most of whom live outside the United States and speak only Spanish. Ultimately, we hope this test will be validated on bilingual Spanish–English listeners as well, given the growing population of bilingual Spanish–English individuals in the United States. Shi (2014) reported that there are few bilingual audiologists to work with non–English-speaking clients, and there are no specific guidelines available to properly and optimally evaluate the speech perception performance of a bilingual Spanish–English individual. Grosjean (1997, 1998) suggested that bilingual speakers cannot be treated in the same way as monolingual speakers because of a myriad of issues related to language history, language skills, language stability, language proficiency, and so forth. In addition, Weiss and Dempsey (2008) suggested that Spanish, as a native language, is not necessarily the dominant language and that the best way to test bilinguals is by testing in both English and Spanish in order to control for variables such as age of acquisition, use of language in different environments, and so forth. However, Grosjean argues that testing bilinguals individually in each language is not effective because these individuals often use both languages simultaneously.

Unfortunately, there are only a few commercially available standardized and validated word recognition tests for pediatric monolingual Spanish-speaking or bilingual Spanish–English individuals. With the growing Hispanic population in the United States, the field of audiology needs valid and reliable word recognition tests that use the Spanish language.

In 1984, Comstock and Martin developed, recorded, and validated four lists of 25 Spanish words to assess word recognition scores of Spanish-speaking children using a picture-pointing task. These lists were developed using words taken from the vocabulary of native Spanish-speaking preschool children. A performance–intensity (PI) function for each list was obtained from 15 native Spanish-speaking adults to determine list equivalency. Once equivalency for the lists was verified, PI functions were obtained from 22 native Spanish-speaking children aged 3–8 years. On the basis of these two experiments, it was determined that the lists were effective for testing children if the vocabulary of the child was considered. Unfortunately, after repeated attempts to contact the authors to find these materials, these recordings could not be located and are no longer available for use by the general audiology community.

Given the increase of the Hispanic population in the United States and the limited number of commercially available materials for Spanish-speaking children, we rerecorded the stimuli used by Comstock and Martin (1984) in order to validate these items in a new test called the Spanish Pediatric Picture Identification Test (SPPIT). This study describes the first step in the validation of this recorded word recognition test for monolingual Spanish-speaking children that can be administered by clinicians who do not speak Spanish. Follow-up studies will validate the SPPIT with bilingual Spanish–English children.

Development of the SPPIT

In this study, we produced and validated our own recordings of the word lists created by Comstock and Martin (1984). Phase 1 of this study involved producing digital recordings of these stimuli and administering the lists to bilingual Spanish–English adults with normal hearing to determine list equivalency. In Phase 2, digitally illustrated pictures that represented the stimuli in the word lists were created and validated using monolingual Spanish-speaking and bilingual Spanish–English children. In Phase 3, list equivalency was re-established, PI functions were obtained, and the final word lists were validated using a picture-pointing task with monolingual Spanish-speaking and bilingual Spanish–English children.

The participants in each phase of this study were either monolingual Spanish speakers or bilingual Spanish–English speakers. The bilingual Spanish–English participants (Phases 1 and 2) were native Spanish speakers, where Spanish was the first language they learned and the primary language used in and out of the home. They had some English skills but were not considered fluent in English. The monolingual Spanish-speaking participants (Phase 3) spoke Spanish as their only language in all situations. More specific information about the participants in Phase 3 can be found in Table 1.

Prior to data collection for each phase of the study, all participants (and/or a parent) signed an informed consent form approved by The University of Memphis Institutional Review Board for participation in this study; participants tested at St. Jude Children's Research Hospital (SJCRH) also signed an informed consent form approved by the SJCRH Institutional Review Board. Basic ethical considerations were taken for the protection of the research participants throughout the project.

Phase 1

The objectives of Phase 1 were (a) to produce digital recordings of the stimuli and (b) to establish list equivalency.

Phase 1: Method

Digital Recordings

Digital recordings of the word lists from Comstock and Martin (1984) were produced using a Marantz Professional Solid State Recorder (PMD660). The original word lists are presented in Appendix A. A 50-year-old bilingual Spanish–English female from Mexico, fluent in both Spanish and English, produced the stimuli for the recordings. A tie-clip electret condenser lapel microphone was placed 3 in. from the talker's mouth, and she was instructed to speak each stimulus word at a natural speaking rate and using natural intonation. The carrier phrase "*Di la palabra* (Say the word)" was recorded and inserted prior to each stimulus item to introduce each target word and help signal the listener. Each stimulus item was recorded 4 times, and the second or third presentation of each word was used in the

Table 1. Demographic information for the participants in Phase 3.

Participant number	Country of Spanish language	Birth country	Current place of residence	Preferred language—participant	Preferred language—family	Other fluent languages of family
1	Ecuador	USA	Florida	Spanish	Spanish	English (parents)
2	Puerto Rico	Puerto Rico	Puerto Rico	Spanish	Spanish	Limited English
3	Ecuador	Ecuador	Ecuador	Spanish	Spanish	Limited English
4	Puerto Rico	Puerto Rico	Puerto Rico	Spanish	Spanish	Limited English
5	Dominican Republic	Dominican Republic	Dominican Republic	Spanish	Spanish	None listed
6	Puerto Rico	Puerto Rico	Puerto Rico	Spanish	Spanish	None listed
7	Puerto Rico	Puerto Rico	Puerto Rico	Spanish	Spanish	None listed
8	Puerto Rico	Puerto Rico	Florida	Spanish	Spanish	English (family)
9	Guatemala	USA	Illinois	Spanish	Spanish	None listed
10	Honduras	Honduras	Honduras	Spanish	Spanish	None listed
11	Ecuador	Ecuador	Ecuador	Spanish	Spanish	None listed
12	Peru	Peru	Peru	Spanish	Spanish	None listed
13	Guatemala	Guatemala	Georgia	Spanish	Spanish	None listed
14	Mexico	USA	Mississippi	Spanish	Spanish	None listed
15	Puerto Rico	Puerto Rico	Puerto Rico	Spanish	Spanish	English (father)
16	Mexico	Mexico	USA	Spanish	Spanish	None listed
17	Puerto Rico	Nicaragua	Nicaragua	Spanish	Spanish	None listed
18	Puerto Rico	Puerto Rico	Puerto Rico	Spanish	Spanish	None listed
19	Nicaragua	Nicaragua	Nicaragua	Spanish	Spanish	None listed
20	Puerto Rico	Puerto Rico	USA	Spanish	Spanish	None listed
21	Honduras	Honduras	Honduras	Spanish	Spanish	English (participant)

final stimulus recording to avoid any variation in intonation during production. Adobe Audition (Version 2.0) was used in postproduction to ensure that all stimuli were leveled to the same root-mean-square level, a 1000-Hz calibration tone was created, and a 2-s interstimulus interval was inserted between the stimulus words.

For the English translation, a 25-year-old native English-speaking female spoke the translation of each of the stimulus items using the same recording parameters. A two-channel recording was created such that the Spanish target word was presented to the listener via one channel and the English translation was presented to the examiner in the other channel. Thus, in one channel, the carrier phrase is produced in Spanish, followed by the Spanish target word. In the other channel, the English translation of the stimulus item is presented immediately after the Spanish word is spoken.

Assessing List Equivalency

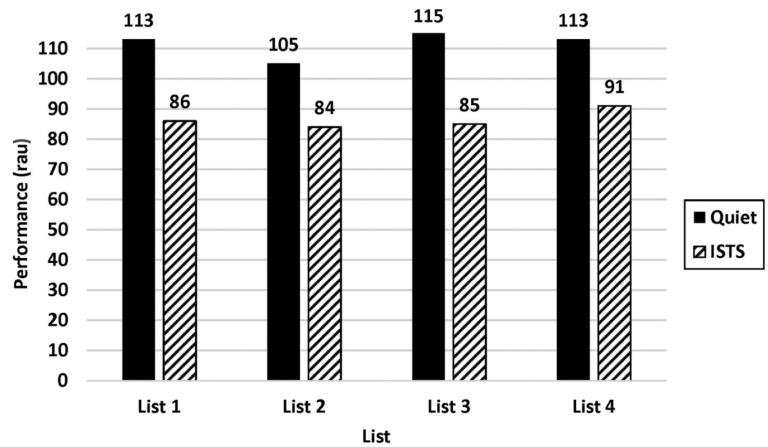
Participants. Fifteen bilingual Spanish–English adults (nine women, six men), aged 22–45 years ($M = 33.2$), were enrolled in the list equivalence portion of this study. To qualify, participants needed a pure-tone average (PTA) of 20 dB HL or better. Five subjects did not qualify due to hearing thresholds falling outside a normal range, leaving a total sample of 10 eligible participants. All participants were of Hispanic heritage from a variety of countries (Puerto Rico, Mexico, Peru, Guatemala, Ecuador, El Salvador, and Honduras). All participants were native Spanish speakers, where Spanish was the first language they learned and the primary language used in and out of the home. There were no known cognitive, speech, or language disabilities. Air-conduction thresholds were obtained at octave frequencies between 250 and 6000 Hz, and all participants had a PTA of 20 dB HL or better and normal middle ear function as evidenced by a Type A tympanogram. The better hearing ear was the ear with the better PTA and the only ear used in the testing.

Procedure. After otoscopy and tympanometry were performed and hearing thresholds were obtained in the better ear, an SRT was obtained using the Trisyllabic Words for SRT in Spanish for Adults recording (Zubick et al., 1983). Then, the four recorded word lists were presented at 40 dB SL in quiet and in noise using the International Speech Test Signal (Holube et al., 2010) at a 0-dB signal-to-noise ratio. The order of list presentation was randomized, and no participant heard each stimulus item more than once. Percent correct performance on each list in each condition was determined. All percent correct scores were converted into rationalized arcsine units (raus) prior to data analysis (Studebaker, 1985). The rau, like the arcsine transform, stabilizes the error variance and helps control for ceiling effects. To ensure accuracy in scoring, the recorded talk-back responses from the participants were rescored by the examiner. Intrajudge scoring reliability was conducted on 30% of the data, using the following formula: [agreements / (agreements + disagreements)] × 100%. Intrajudge scoring reliability was 98.7%.

Phase 1: Results

Figure 1 displays the mean performance in raus as a function of list in quiet and in noise. A two-way repeated-measures analysis of variance was conducted comparing performance between lists and conditions, revealing a significant main effect for condition. All scores obtained in

Figure 1. Mean percent correct word recognition scores transformed into rationalized arcsine units (raus) for all lists in quiet and in noise (ISTS). ISTS = International Speech Test Signal.



noise were significantly poorer than those obtained in quiet, $F(1, 20) = 306.521$, $p < .001$. In addition, a significant interaction effect was found between list and condition, $F(1, 20) = 13.102$, $p < .001$. Post hoc Tukey all-pairwise comparisons revealed that List 2 had significantly poorer scores compared to Lists 1, 3, and 4 in quiet. However, within the noise condition, there were no significant differences obtained between lists. In order to be sure that all lists were equivalent, stimulus items from List 2 were redistributed to the other three lists to even out the number of errors across lists. List equivalence was then re-established in Phase 3.

Phase 2

The objective of Phase 2 was to create digitally illustrated pictures for each stimulus item and validate them on Spanish-speaking children. Each picture board page consists of four pictures, of which three serve as a stimulus item in one of the three lists and one serves as a distractor image.

Phase 2: Method

Pictorial Representation

Pictorial representations of each stimulus word were created using Adobe Illustrator. Although there were only 100 stimuli, 113 pictures were drawn digitally and evaluated because some stimulus items could be represented in different ways to reflect their meaning. Difficult-to-depict stimuli were represented in two different ways to see which representation yielded better performance on the picture identification task. All pictorial representations were placed on flash cards and presented in random order to 14 monolingual Spanish-speaking children (seven boys, seven girls) aged 2–12 years ($M = 3.9$). For all children, Spanish was their only language. The children were shown each picture on a single flash card and were asked "*Qué es (What is)?*" to have them verbally name the picture. The children verbally identified the picture if they recognized the image. If the children were unable to verbally identify the item, they were shown a picture board of four different items and asked to point to the appropriate picture in a closed-set task.

Phase 2: Results

Picture identification results in Table 2 were determined by tallying the number of correct identifications of each target picture and scoring them as percent correct per stimulus item. Many pictures were easily identified by most, if not all, participants. However, several stimulus items received poor scores for identification due to several reasons. In some instances, the picture was not easily identifiable by the subjects. In other cases, the English translation we used to depict the stimulus item was not accurate for Spanish speakers from different regions. For example, "*mecha*" could be translated as "dynamite" in some regions; however, it was translated as "match" by Comstock and Martin (1984). In addition, we used an incorrect translation of the word "*masa*" as "flower," which actually refers to "flour" used in baking. Finally, there were several words that appeared to be outside the expressive/receptive vocabulary of our target population, such as "*callo*," which means "corn blister."

Considering the various types of errors we observed from the original lists from Comstock and Martin (1984), we felt that some of the confusing items should be eliminated as target stimuli for the test. Items in Table 2 with the lowest percent correct score in a group of four are in bold. We also considered the results from Phase 1, which indicated more errors occurred on List 2, in determining which items should be eliminated. Therefore, the number of usable word lists for the picture-pointing response was reduced from four to three for our final recordings. The word lists were then rebalanced to include stimuli that received a high percentage of picture recognition for Lists 1, 2, and 3, whereas List 4 served as nontarget items that had either poor performance on the picture identification task or a large number of errors from Phase 1. Appendix B represents

Table 2. Phase 2 picture identification results.

List 1		List 2	
mala (sick female)	100%	*pala* (shovel)	71%
boca (mouth)	100%	*troca* (truck)	100%
lloro (crying)	79%	***pozo* (hole)**	**43%**
ocho (eight)	100%	*ojo* (eye)	100%
cama (bed)	100%	*caja* (box)	93%
hueso (bone)	93%	*queso* (cheese)	93%
casa (house)	100%	*pasa* (raisin)	79%
saco (coat)	71%	***callo* (corn blister)**	**29%**
mano (hand)	100%	*sapo* (frog)	93%
pelo (hair)	71%	***pecho* (chest)**	**64%**
lodo (mud)	71%	*rojo* (red)	100%
***cuna* (crib)**	**64%**	*sopa* (soup)	100%
rata (rat)	100%	*rama* (branch)	57%
***misa* (mass)**	**64%**	*niña* (girl)	93%
plato (plate)	93%	***pato* (duck)**	**86%**
comer (eat)	100%	*corer* (to run)	100%
***mata* (plant)**	**36%**	*bata* (bathrobe)	36%
boda (wedding)	79%	*bota* (boot)	93%
pera (pear)	71%	*baña* (to bathe)	100%
baño (bathtub)	93%	***vaso* (glass)**	**79%**
reza (to pray)	71%	*vela* (candle)	50%
silla (chair)	93%	*tía* (aunt)	100%
moto (motorcycle)	93%	*foco* (light bulb)	86%
cena (supper)	79%	*fecha* (date)	64%
seco (to dry)	86%	*dedo* (finger)	79%
List 3		**List 4**	
sala (living room)	93%	***ala* (bird's wing)**	**64%**
***foca* (seal)**	**71%**	*coca* (Coke)	100%
yoyo	57%	*torro* (bull)	93%
oso (bear)	79%	***oro* (gold)**	**36%**
***capa* (cape)**	**71%**	*cara* (face)	79%
***peso* (dollar)**	**50%**	*beso* (kiss)	100%
***masa* (flower)**	**57%**	*taza* (cup)	71%
gallo (rooster)	79%	*taco*	93%
***palo* (stick)**	**93%**	*carro* (car)	100%
burro (donkey)	86%	*perro* (dog)	79%
pollo (chicken)	93%	***codo* (elbow)**	**57%**
luna (moon)	100%	*soda*	93%
papa (potato)	86%	***faja* (belt)**	**57%**
piña (pineapple)	86%	*risa* (laughing)	79%
uno (one)	100%	*gato* (cat)	100%
llover (to rain)	100%	***coser* (sew)**	**43%**
pata (foot)	100%	*vaca* (cow)	93%
***borra* (erase)**	**36%**	*bote* (tin can)	64%
uña (fingernail)	57%	***garra* (rag)**	**50%**
brazo (arm)	100%	*mono* (monkey)	93%
***tela* (screen)**	**43%**	*mesa* (table)	100%
***liga* (rubber band)**	**64%**	*chiva* (goat)	71%
coco (coconut)	79%	***roto* (torn)**	**71%**
***mecha* (match)**	**21%**	*cejo* (eyebrow)	64%
huevo (egg)	100%	***velo* (veil)**	**57%**

Note. Scores reflect the percentage of children who correctly identified the picture associated with that item. Items in bold became nontarget items and were moved to List 4 and validated in Phase 3.

the final version of the word lists. The stimulus items in bold in List 4 were not used as a test list but rather as distractor items for the picture-pointing task.

Final Stimulus Set

Twenty-five picture boards were created with four illustrations on each, representing the 25 rows of words found in Appendix B. Each picture board contains a target word from Lists 1, 2, and 3 as well as a distractor item from List 4. The position of each picture was randomly determined so that the target stimulus was not in the same position on each picture board. Stimulus items were presented in a consonant–vowel–consonant–vowel context such that the vowel used in each item was the same across lists. For example, for Item 25, the vowel /e/ was used for each word across lists: List 1, *dedo* (finger); List 2, *seco* (to dry); List 3, *huevo* (egg); List 4, *velo* (veil). Examples of several picture boards are shown in Figure 2.

Phase 3

The goal of Phase 3 was to validate the final picture board by (a) re-establishing list equivalency, (b) obtaining PI functions on the final stimulus set, and (c) validating the final word lists and picture boards using a picture-pointing task with Spanish-speaking children.

Phase 3: Method

Adobe Audition (Version 2.0) was used to reorder the recorded stimuli into tracks that represented the newly arranged lists established from Phase 2 and presented in Appendix B. In addition, the carrier phrase was changed to "*Enséñame* (Show me)" to be used in the final picture-pointing task, and it was rerecorded by the same female Spanish speaker.

Participants. Ten bilingual Spanish–English adults (eight women, two men), aged 21–55 years (M = 34), and 21 monolingual Spanish-speaking children (nine girls, 12 boys), aged 4–14 years (M = 7.72), were enrolled in the validation portion of this study. All participants were native Spanish speakers and of Hispanic heritage from a variety of countries (Mexico, Guatemala, El Salvador, Colombia, Ecuador, Puerto Rico, Dominican Republic, Honduras, Peru, and Nicaragua). For the monolingual listeners, Spanish was their only language, and for the bilingual Spanish–English listeners, Spanish was the first language they learned and the primary language used in and out of the home (see Table 1). There were no known cognitive, speech, or language disabilities. All participants had a PTA of 20 dB HL or better and normal middle ear function as evidenced by a Type A tympanogram in the test ear on the day of testing. The test ear was the better ear for each participant. Air-conduction thresholds were obtained at octave frequencies between 500 and 4000 Hz.

Instrumentation. Validation data were collected in two locations in Memphis, Tennessee: (a) Speech Perception Assessment Laboratory (SPAL) at The University of Memphis and (b) SJCRH. The 10 adults were tested in the SPAL, and 21 children were tested at SJCRH.

The adults evaluated in the SPAL were tested in a sound-treated booth meeting American National Standards Institute (ANSI) S.31-1999 (R2013). Air-conduction thresholds were obtained with a Grason-Stadler GSI 61 using Telephonics TDH-50 supra-aural earphones. Tympanograms

Figure 2. Sample picture boards. (a) Item 2: *boca* (mouth)/*troca* (truck)/*coca* (Coke)/*foca* (seal). (b) Item 3: *lloro* (crying)/*yoyo*/*toro* (bull)/*pozo* (hole). (c) Item 15: *plato* (plate)/*pato* (duck)/*uno* (one)/*gato* (cat). (d) Item 16: *coser* (sew)/*corer* (to run)/*comer* (to eat)/*llover* (to rain). (e) Item 22: *liga* (rubber band)/*silla* (chair)/*chiva* (goat)/*tia* (aunt). (f) Item 25: *dedo* (finger)/*huevo* (egg)/*velo* (veil)/*seco* (to dry).

were obtained using a Madsen MI 34 middle ear analyzer. The stimuli were presented using a Sony CD player (SCD-CE595) routed through the GSI 61. The 21 participants at SJCRH were tested in a sound-treated booth meeting ANSI S.31-1999 (R2013). Air-conduction thresholds were obtained using a GSI AudioStar Pro audiometer (Grason-Stadler, Inc.) with Eartone 3A insert earphones, unless the child would not tolerate insert phones. In that case, Sennheiser HDA 200 supra-aural headphones were used. Tympanograms were obtained using GSI TympStar and GSI TympStar Pro tympanometers (Grason-Stadler, Inc.). A Sony CD player (SCD-CE595) routed through the GSI AudioStar Pro was used to administer the speech stimuli.

Procedure—adults. After otoscopy and tympanometry were performed and the hearing thresholds obtained, SRTs were measured in the better ear using the Auditec of St. Louis recording of the Trisyllabic Words for SRT in Spanish for Adults (Zubick et al., 1983). The SPPIT word list recordings were presented at 0, 8, 16, 24, 32, and 40 dB SL above the SRT. The lists and levels were counterbalanced such that each participant listened randomly to Lists 1, 2, and 3 presented at two different sensation levels (SLs) each. Because each list was heard by each participant at two levels, the first presentation for each list was at a very low SL and the second at a higher SL to minimize learning and practice effects. The adults did not use the picture-pointing response; they repeated the word they heard for each list.

Procedure—children. SRTs were measured in the better ear after otoscopy and tympanometry were performed and after hearing thresholds were obtained. The Spanish Pediatric Speech Recognition Threshold Test (Mendel et al., 2019) was used to obtain SRTs for the children. The Spanish Pediatric Speech Recognition Threshold is a picture-pointing task designed to be used with Spanish-speaking children, which was validated using this participant sample (Mendel et al., 2019). Once an SRT was obtained, the SPPIT was presented to the better hearing ear at a minimum of two of a possible four levels (8, 16, 32, and 40 dB SL) above the child's SRT. The presentation of the lists at the different SLs was counterbalanced such that each child heard each of the three lists at two different SLs. Some of the participants, however, were unable to complete all lists in the protocol, such that, in some cases, a child may have only heard each list at one SL. As with the adults, the first presentation of each list was at a very low SL and the second at a higher SL to minimize learning and practice effects. The participants were instructed to point to the picture that represented the word they heard using the SPPIT picture board, which consisted of a "two-row by two-column" four-item grid, with one image on each board representing the target stimulus and the remaining three images representing distractors. The presentation order of the lists was counterbalanced across participants.

Phase 3: Results

PI Functions

Figure 3 displays PI functions from the adult responses to the three lists of stimuli presented using an open-set response mode at SLs from 0 to 40 dB SL in 8-dB steps. A two-way general linear model was conducted to assess possible differences between the lists and SLs. There was no significant interaction between lists and levels, $F(10, 42) = 0.720$, or a main effect for list, $F(2, 42) = 0.640$, $p = .533$. However, a significant main effect for level was observed, $F(5, 42) = 52.46$, $p < .001$. The differences among SLs were assessed using a post hoc Bonferroni-adjusted p value, and significant differences were found between 0 dB SL compared to all other levels (8, 16, 24, 32, and 40 dB SL), $p < .01$. Furthermore, significant differences were measured between 8 dB SL and higher levels of 16, 24, 32, and 40 dB SL, $p < .01$. No additional differences were noted above 16 dB SL compared to higher levels (24, 32, and 40 dB SL).

Once list equivalency was established for the audio recordings of the stimuli, the children listened to the stimuli and responded using the picture-pointing task and the SPPIT picture boards. Figure 4 displays the PI functions for the children at 8, 16, 32, and 40 dB SL, and Table 3 displays the percent correct performance of the children on the three lists including standard deviations. A linear mixed-model analysis was conducted to assess possible differences between the three lists and SLs (8, 16, and 32 dB). We did not include 40 dB SL in the analysis because the scores plateaued at that level. Both lists and SLs were considered random factors and included as fixed effects. Results indicated no significant interaction between lists and SLs, $F(4, 62.71) = 0.958$, $p = .437$, but significant main effects were measured for list, $F(2, 47.29) = 6.46$, $p = .003$, and SL, $F(2, 44.35) = 57.85$, $p < .001$. A post hoc analysis of the main effects using a Bonferroni adjustment for multiple comparisons revealed significantly higher SPPIT scores for List 3 compared to List 1 ($p = .009$) and for List 3 compared to List 2 ($p = .004$), but only at 8 dB SL. This difference is not considered to be clinically significant given that this test should be presented at a higher SL.

Item Analysis

A two-way analysis of variance was conducted comparing the number of errors at each presentation level across lists. There were no significant differences between

Figure 3. Adult performance–intensity (PI) functions. Error bars reflect ± 1 *SD*.

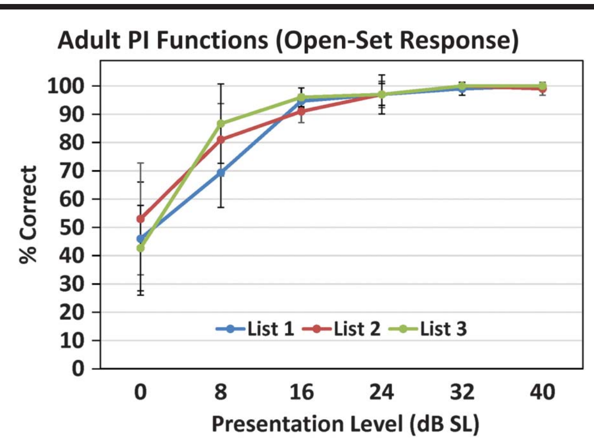


Figure 4. Children's performance–intensity (PI) functions. Error bars reflect ± 1 *SD*.

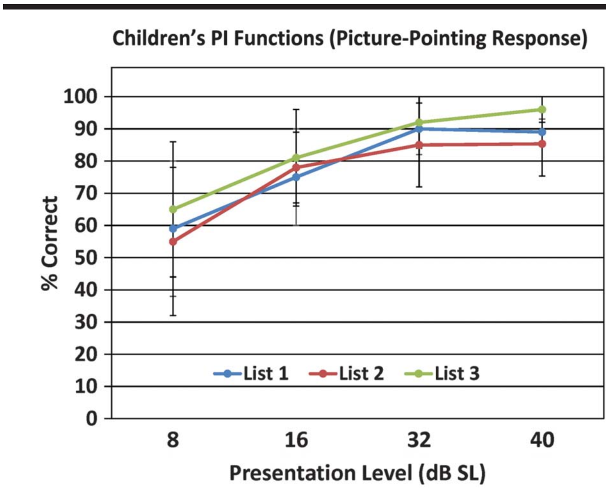


Table 3. Mean percent correct performance and standard deviations for the children for each list at the four presentation levels.

	Presentation levels			
List	8 dB SL	16 dB SL	32 dB SL	40 dB SL
List 1				
M	59%	75%	90%	89%
SD	17.03	14.94	9.13	7.20
List 2				
M	55%	78%	85%	85%
SD	22.69	11.09	13.12	9.65
List 3				
M	65%	81%	92%	96%
SD	20.88	15.02	7.80	4.40

lists at any level, $F(1, 114) = 0.114$, $p = .995$. In List 1, the highest number of errors across all presentation levels occurred on the word *boca* (mouth), which had a total of 29 errors in identification, followed by *troca* (truck) with a total of 23 errors in List 2 and by *pata* (foot) with a total of 21 errors in List 3. Each list had about the same number of items with errors across all test levels. For example, List 1 had five items with 20 or more errors, List 2 had five items with 20 or more errors, and List 3 had one item with 20 or more errors.

Discussion

The goal of this study was to develop and validate the SPPIT as a picture-pointing Spanish word recognition test to be used with Spanish-speaking children. The results presented here provide normative data for monolingual Spanish-speaking children. Future studies will provide data for bilingual Spanish–English children with varying degrees of hearing loss. Several steps were followed in the standardization of this test, including producing digital recordings of these stimuli, creating and validating digitally illustrated pictures that represented the stimuli in the word lists, and establishing list equivalency.

Stimulus Creation

The initial attempt at establishing list equivalence with adults (Phase 1) yielded significant differences in mean performance on one list and a large number of errors on specific test items. These differences would not have been observed if testing had been conducted only in a quiet condition. By adding noise to make the task more difficult, significant differences in mean performance on the original List 2 became apparent. By rearranging the stimuli across the lists and re-establishing list equivalence, the final version of the SPPIT was created, resulting in three test lists of stimuli and one list containing distractor items.

In most cases, the pictorial representations of the stimuli were accurate and identifiable (Phase 2) by the children. Caution had to be taken to ensure that the English translation used to depict each item was accurate, not only in regard to a direct translation from English to Spanish but also to be sure the vocabulary level was appropriate for children. It is still possible that some Spanish-speaking children will have difficulty with some stimulus items due to regional and dialectical differences. However, by moving most of the potentially confusing items to the distractor list (List 4), we believe this difficulty should be minimized.

Establishing List Equivalence

The PI functions shown in Figure 3 for the adult listeners who responded to each stimulus item without pictures (open-set response) show quite a bit of variance in performance for the low presentation levels, but eventually, the standard deviations become minimal and scores become maximal at the higher presentation levels. Figure 4, which displays the PI functions for the children who used the picture-pointing response, also shows a broad range of scores at low presentation levels, which narrows as the presentation level increases. However, the scores for the children do not reach 100%, and the standard deviations do not reduce as much as they did for the adults. These findings are consistent with those obtained by Weisleder and Hodgson (1989), who evaluated a Spanish language test for adults and found large standard deviations at low presentation levels. In addition, maximal scores did not always reach 100% at the highest presentation levels. Variability in the scores obtained by the children in this study could also have been due to a low number of scores per SL. In order to reduce the effect of learning and practice, each child who participated in this study only heard the SPPIT at a minimum of two of possible five SLs. Thus, the number of scores contributing to the average per presentation level was reduced and may have limited the power; this could have contributed to the variability as well.

The slopes of the PI functions in Figure 4 for this study were relatively flat compared to slopes of other word recognition tests. The percentages of increase in recognition ability per decibel increment for our lists were 1.21%/dB (List 1), 1.24%/dB (List 2), and 1.02%/dB (List 3). These values are lower than those reported for traditional auditory word recognition tests in English. Given that there are few monosyllabic words in Spanish and that the SPPIT contains bisyllabic words, it is reasonable to compare our results with English monosyllabic tests. Tillman and Carhart (1966) reported 5.6%/dB for the NU-6, whereas Wilson et al. (1976) and Beattie et al. (1977) reported 3.6%/dB and 4.2%/dB, respectively; for the W-22, Beattie et al. reported 4.6%/dB.

We could not find information regarding slopes of PI functions for any other Spanish picture-pointing test for a direct comparison; however, we compared our slopes to other picture-pointing tests in English. Chermak et al. (1988) reported the slope of the PI function for the WIPI at 5.65%/dB, whereas Elliott and Katz (1980) reported the slope for the NU-CHIPS ranging from as low as 0.67%/dB to as high as 3%/dB. Thus, some similarity is seen in the lower slope of the PI functions for the picture-pointing tests.

Regional varieties of the Spanish language can be quite different from one another, specifically with regard to pronunciation and vocabulary (Comstock & Martin, 1984). Thus, the outcomes from this study are limited to a population familiar with Latin American Spanish given the ethnic backgrounds of the talker on the recordings and the children tested. We recognize that, even within Latin America, there are many different dialect areas that could have had an influence on our results. A separate recording of these stimulus items is currently being validated using a male talker with a European Spanish dialect in order to provide audiologists with additional dialectical options for their Spanish-speaking patients.

Limitations

There are several limitations to the data presented here that are worthy of mention. First, grouping individuals with different dialects and from different regions could have influenced the results of this study. Shi (2014) reports that dialectal differences can have an effect on one's speech perception ability. It is possible that the pictures portrayed in the SPPIT could be identified with different words than those in our stimulus lists based on the regional dialect of the child. It should be noted that most of the speech perception tests currently used in English typically do not account for dialectal or regional differences.

Second, although we did provide information regarding the participants' language use, the results from this study would have greater external validity if we had also provided more details regarding the participants' linguistic characteristics and their language proficiency. If a patient or family reported Spanish as the preferred language with no known cognitive or language disorders, we made the assumption that the patient was proficient in his or her preferred language of Spanish.

Third, the normative data presented here are limited to monolingual Spanish-speaking children with hearing within a normal range or with minimal hearing loss. Future research should investigate listeners with different degrees and configurations of hearing loss to determine their performance on the SPPIT. Future research should also validate this test on bilingual Spanish–English children to determine expected speech recognition scores for that population. It is important to validate this test using different populations of Spanish-speaking children because bilingualism can change across the life span. As children are exposed to a second language, they may no longer be truly monolingual. Thus, the normative data presented here for monolingual children using the SPPIT may be different than scores for bilingual listeners. Desjardins et al. (2019) showed that testing speech recognition in both languages of a bilingual listener may help with counseling and recommending appropriate treatments. Clinicians who use speech perception assessments in both languages will reduce the possibility of misdiagnosing a child's speech recognition ability due to lexical errors.

In addition to the SPPIT, there are a few other speech recognition tests that might be appropriate for the bilingual (Spanish/English) child and the monolingual (Spanish) child. For example, Calandruccio et al. (2014) developed a picture identification task using disyllabic words that could be used with monolingual English-speaking, monolingual Spanish-speaking, and bilingual Spanish–English children in background noise, competing speech, or both. This test was designed such that the examiner would determine which language is most appropriate for testing on a case-by-case basis. Because bilingual Spanish–English children are often acquiring both languages at the same time, it is difficult to determine which language is most appropriate for testing.

In conclusion, the SPPIT picture board can provide ease and accuracy of scoring speech perception ability for the monolingual English-speaking clinician. Instead of judging the accuracy of an oral response, the audiologist can simply judge if the appropriate word/picture was identified. In addition, the picture-pointing task is also useful for nonverbal or shy children. This test could also be used for Spanish-speaking adults with limited or no English proficiency. Future research will validate this test on bilingual listeners to broaden its use for both monolingual and bilingual Spanish–English children.

Acknowledgments

The Spanish Pediatric Picture Identification Test was developed under Grant H325100322 from the U.S. Department of Education, awarded to Linda Jarmulowicz (Principal Investigator), Teresa Wolf (Co-Principal Investigator), and Jennifer P. Taylor (Project Associate). Portions of this article were presented at the annual convention of the American Speech-Language-Hearing Association in November 2014 in Orlando, FL; at AudiologyNOW! in March 2015 in San Antonio, TX; and at the American Academy of Audiology Conference in 2018 in Nashville, TN. The Spanish Pediatric Picture Identification Test is available from Auditec, Inc. Appreciation is expressed to Chhayakant Patro and Sungmin Lee for their assistance in producing the recordings and to Skylar McSorley, Skye Jones, and Stephanie White for assisting with data collection at both locations.

References

American Speech-Language-Hearing Association. (2002). *Characteristics and needs of children with hearing loss and/or APD.* https://www.asha.org/policy/gl2002-00005/#sec1.4

Beattie, R. C., Edgerton, B. J., & Svihovec, D. V. (1977). A comparison of the Auditec of St. Louis cassette recordings of NU-6 and CID W-22 on a normal-hearing population. *Journal of Speech and Hearing Disorders, 42*(1), 60–64. https://doi.org/10.1044/jshd.4201.60

Calandruccio, L., Gomez, B., Buss, E., & Leibold, L. J. (2014). Development and preliminary evaluation of a pediatric Spanish–English speech perception task. *American Journal of Audiology, 23*(2), 158–172. https://doi.org/10.1044/2014_AJA-13-0055

Chermak, G. D., Wagner, D. P., & Bended, R. B. (1988). Interlist equivalence of the Word Intelligibility by Picture Identification Test administered in broad-band noise. *International*

Journal of Audiology, 27(6), 324–333. https://doi.org/10.3109/00206098809081603

Comstock, C. L., & Martin, F. N. (1984). A children's Spanish word discrimination test for non–Spanish-speaking clinicians. *Ear and Hearing, 5*(3), 166–170. https://doi.org/10.1097/00003446-198405000-00008

Davis, H. (1978). Audiometry: Pure tone and simple speech tests. In H. Davis & S. R. Silverman (Eds.), *Hearing and deafness* (pp. 183–221). Holt, Rinehart & Winston.

Desjardins, J. L., Barraza, E. G., & Orozco, J. A. (2019). Age-related changes in speech recognition performance in Spanish–English bilinguals' first and second languages. *Journal of Speech, Language, and Hearing Research, 62*(7), 2553–2563. https://doi.org/10.1044/2019_JSLHR-H-18-0435

Elliott, L. L., & Katz, D. R. (1980). *Northwestern University Children's Perception of Speech (NU-CHIPS)*. Auditec.

Gaeta, L., & John, A. B. (2015). Considerations in speech recognition testing of bilingual and Spanish-speaking patients. Part II: Young children. *Journal of Educational, Pediatric & (Re)Habilitative Audiology, 1,* 1–11.

Gelfand, S. (2016). *Essentials of audiology*. Thieme. https://doi.org/10.1055/b-006-161125

Grosjean, F. (1997). The bilingual individual. *Interpreting, 2*(1–2), 163–187. https://doi.org/10.1075/intp.2.1-2.07gro

Grosjean, F. (1998). Studying bilinguals: Methodological and conceptual issues. *Bilingualism: Language and Cognition, 1*(2), 131–149. https://doi.org/10.1017/S136672899800025X

Holube, I., Fredelake, S., Vlaming, M., & Kollmeier, B. (2010). Development and analysis of an International Speech Test Signal (ISTS). *International Journal of Audiology, 49*(12), 891–903. https://doi.org/10.3109/14992027.2010.506889

McArdle, R., & Hnath-Chisolm, T. (2009). Speech audiometry. In J. Katz (Ed.), *Clinical handbook of audiology* (pp. 64–79). Lippincott Williams & Wilkins.

Mendel, L. L., Pousson, M., Bass, J. K., Lunsford, R. E., & McNiece, C. (2019). Spanish Pediatric Speech Recognition Threshold test. *American Journal of Audiology, 28*(3), 597–604. https://doi.org/10.1044/2019_AJA-18-0132

Shi, L.-F. (2014). Speech audiometry and Spanish–English bilinguals: Challenges in clinical practice. *American Journal of Audiology, 23*(3), 243–259. https://doi.org/10.1044/2014_AJA-14-0022

Studebaker, G. A. (1985). A "rationalized" arcsine transform. *Journal of Speech and Hearing Research, 28*(3), 455–462. https://doi.org/10.1044/jshr.2803.455

Tillman, T. W., & Carhart, R. (1966). *An expanded test for speech discrimination utilizing CNC monosyllabic words: Northwestern University Test No. 6 (SAM-TR-66-55)*. Defense Technical Information Center. https://doi.org/10.21236/AD0639638

U.S. Census Bureau. (n.d.). *QuickFacts*. Retrieved 2019, from https://www.census.gov/quickfacts/fact/table/US/PST045217

Weisleder, P., & Hodgson, W. R. (1989). Evaluation of four Spanish word-recognition-ability lists. *Ear and Hearing, 10*(6), 387–392. https://doi.org/10.1097/00003446-198912000-00012

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*(1), 5–17. https://doi.org/10.3766/jaaa.19.1.2

Wilson, R. H., Coley, K. E., Haenel, J. L., & Browning, K. M. (1976). Northwestern University Auditory Test No. 6: Normative and comparative intelligibility functions. *Journal of the American Auditory Society, 5,* 221–228.

Zubick, H. H., Irizarry, L. M., Rosen, L., Feudo, P., Kelly, J. H., & Strome, M. (1983). Development of speech-audiometric materials for native Spanish-speaking adults. *International Journal of Audiology, 22*(1), 88–102. https://doi.org/10.3109/00206098309072772

Appendix A

Word Lists From Comstock and Martin (1984) Digitally Recorded for the SPPIT

List 1	List 2	List 3	List 4
mala (sick female)	*pala* (shovel)	*sala* (living room)	*ala* (bird's wing)
boca (mouth)	*troca* (truck)	*foca* (seal)	*coca* (coke)
lloro (crying)	*pozo* (hole)	*yoyo*	*torro* (bull)
ocho (eight)	*ojo* (eye)	*oso* (bear)	*oro* (gold)
cama (bed)	*caja* (box)	*capa* (cape)	*cara* (face)
hueso (bone)	*queso* (cheese)	*peso* (dollar)	*beso* (Kiss)
casa (house)	*pasa* (raisin)	*masa* (flower)	*taza* (cup)
saco (coat)	*callo* (corn blister)	*gallo* (rooster)	*taco*
mano (hand)	*sapo* (frog)	*palo* (stick)	*carro* (car)
pelo (hair)	*pecho* (chest)	*burro* (donkey)	*perro* (dog)
lodo (mud)	*rojo* (red)	*pollo* (chicken)	*codo* (elbow)
cuna (crib)	*sopa* (soup)	*luna* (moon)	*soda*
rata (rat)	*rama* (branch)	*papa* (potato)	*faja* (belt)
misa (mass)	*niña* (girl)	*piña* (pineapple)	*risa* (laughing)
plato (plate)	*pato* (duck)	*uno* (one)	*gato* (cat)
comer (eat)	*corer* (to run)	*llover* (to rain)	*coser* (sew)
mata (plant)	*bata* (bathrobe)	*pata* (foot)	*vaca* (cow)
boda (wedding)	*bota* (boot)	*borra* (erase)	*bote* (tin can)
pera (pear)	*baña* (to bathe)	*uña* (fingernail)	*garra* (rag)
baño (bathtub)	*vaso* (glass)	*brazo* (arm)	*paña* (cloth)
reza (to pray)	*vela* (candle)	*tela* (screen)	*mesa* (table)
silla (chair)	*tia* (aunt)	*liga* (rubber band)	*chiva* (goat)
moto (motorcycle)	*foco* (light bulb)	*coco* (coconut)	*roto* (torn)
cena (supper)	*fecha* (date)	*mecha* (match)	*cejo* (eyebrow)
seco (to dry)	*dedo* (finger)	*huevo* (egg)	*velo* (veil)

Appendix B

Redistributed Stimulus Items Resulting in the Final Test Lists

List 1	List 2	List 3	List 4
mala (sick female)	*sala* (living room)	*pala* (shovel)	***ala* (bird's wing)**
boca (mouth)	*troca* (truck)	*coca* (Coke)	***foca* (seal)**
toro (bull)	*lloro* (crying)	*yoyo*	***pozo* (hole)**
ojo (eye)	*oso* (bear)	*ocho* (eight)	***oro* (gold)**
caja (box)	*cara* (face)	*cama* (bed)	***capa* (cape)**
beso (kiss)	*hueso* (bone)	*queso* (cheese)	***peso* (dollar)**
pasa (raisin)	*taza* (cup)	*casa* (house)	***masa* (flower)**
saco (coat)	*taco*	*gallo* (rooster)	***callo* (corn blíster)**
carro (car)	*sapo* (frog)	*mano* (hand)	***palo* (stick)**
perro (dog)	*burro* (donkey)	*pelo* (hair)	***pecho* (chest)**
rojo (red)	*pollo* (chicken)	*lodo* (mud)	***codo* (elbow)**
soda	*luna* (moon)	*sopa* (soup)	***cuna* (crib)**
papa (potato)	*rama* (branch)	*rata* (rat)	***faja* (belt)**
niña (girl)	*piña* (pineapple)	*risa* (laughing)	***misa* (mass)**
plato (plate)	*uno* (one)	*gato* (cat)	***pato* (duck)**
correr (to run)	*comer* (to eat)	*llover* (to rain)	***coser* (sew)**
vaca (cow)	*bata* (bathrobe)	*pata* (foot)	***mata* (plant)**
bote (tin can)	*bota* (boot)	*boda* (wedding)	***borra* (erase)**
pera (pear)	*uña* (fingernail)	*baña* (to bathe)	***garra* (rag)**
brazo (arm)	*baño* (bathtub)	*mono* (monkey	***vaso* (glass)**
mesa (table)	*reza* (to pray)	*vela* (candle)	***tela* (screen)**
chiva (goat)	*tía* (aunt)	*silla* (chair)	***liga* (rubber band)**
foco (light bulb)	*coco* (coconut)	*moto* (motorcycle)	***roto* (torn)**
cena (supper)	*ceja* (eyebrow)	*fecha* (date)	***mecha* (match)**
dedo (finger)	*seco* (to dry)	*huevo* (egg)	***velo* (veil)**

Note. Lists 1, 2, and 3 contain stimulus items, and List 4 contains distractors.