



Published in final edited form as:

J Am Acad Audiol. 2012 June ; 23(6): 464–475. doi:10.3766/jaaa.23.6.8.

Assessing spoken word recognition in children who are deaf or hard of hearing: A translational approach

Karen Iler Kirk¹, Lindsay Prusick¹, Brian French², Chad Gotch², Laurie S. Eisenberg³, and Nancy Young⁴

¹The University of Iowa, Iowa City, IA

²Washington State University, Pullman, WA

³House Research Institute, Los Angeles, CA

⁴Children's Memorial Hospital, Chicago, IL

Abstract

Under natural conditions, listeners use both auditory and visual speech cues to extract meaning from speech signals containing many sources of variability. However, traditional clinical tests of spoken word recognition routinely employ isolated words or sentences produced by a single talker in an auditory-only presentation format. The more central cognitive processes used during multimodal integration, perceptual normalization and lexical discrimination that may contribute to individual variation in spoken word recognition performance are not assessed in conventional tests of this kind. In this paper, we review our past and current research activities aimed at developing a series of new assessment tools designed to evaluate spoken word recognition in children who are deaf or hard of hearing. These measures are theoretically motivated by a current model of spoken word recognition and also incorporate “real-world” stimulus variability in the form multiple talkers and presentation formats. The goal of this research is to enhance our ability to estimate real-world listening ability and to predict benefit from sensory aid use in children with varying degrees of hearing loss.

Spoken word recognition tests have been part of the audiological assessment of individuals who are deaf or hard of hearing for the last 60 years. Such tests have been used to assess the effects of hearing loss on spoken word recognition and speech perception, (Hudgins, Hawkins, Karling, & Stevens, 1947; Skinner, et al., 2002), to determine cochlear implant candidacy (Gifford, Dorman, Spahr, & Bacon, 2007; Kirk, 2000a), measure cochlear implant outcomes (Dorman, Gifford, Spahr, & McKarns, 2008; Eisenberg, Kirk, Martinez, Ying, & Miyamoto, 2004; Holt, Kirk, Eisenberg, Martinez, & Campbell, 2005) and to guide the development of an individual's aural rehabilitation programs (Eisenberg, 2007; Mackersie, 2002).

In research settings, listeners who are deaf or hard of hearing may be administered spoken word recognition tests to evaluate the effectiveness of signal processing strategies (Dorman, Loizou, Kemp, & Kirk, 2000; Firszt, Holden, Reeder, & Skinner, 2009; Gifford, Olund, & Dejong, 2011; Gifford, Shallop, & Peterson, 2008; Skinner, et al., 2002), to better understand perceptual processes that support spoken word recognition (Holt, Beer, Kronenberger, Pisoni, & Lalonde, 2012; Kirk, Pisoni, & Osberger, 1995; Krull, Choi, Kirk, Prusick, & French, 2010; Loebach, Bent, & Pisoni, 2008; Loebach & Pisoni, 2008) or to

identify factors that contribute to individual variability (Desjardin, Ambrose, Martinez, & Eisenberg, 2009; Geers, 2006; Pisoni, Kronenberger, Roman, & Geers, 2011). Obviously, no one test can achieve all of these aims. Thus, the choice of spoken word recognition test(s) is dictated by the information one wishes to obtain.

Our test development work originally was motivated by the observation that children with cochlear implants performed poorly on traditional tests of open-set word recognition despite the fact that parents reported their children could understand some speech through listening alone. We hypothesized that traditional tests, such as the Phonetically Balanced Kindergarten Word List (PBK) (Haskins, 1949), included vocabulary that was unfamiliar to children who are deaf or hard of hearing due to the constraints imposed by phonetic balancing of test items. Furthermore, we noted some shortcomings with traditional spoken word recognition tests. First, most tests used stimuli produced by one talker in carefully articulated speech. Such tests, in which acoustic variability is highly constrained, may not accurately reflect spoken word recognition abilities under more natural listening situations. Increasing stimulus variability by introducing multiple talkers or varying speaking rate reduces spoken word recognition performance in listeners with normal hearing (Mullenix, Pisoni, & Martin, 1989; Nygaard, Sommers, & Pisoni, 1992; Sommers, Nygaard, & Pisoni, 1994) and in listeners who are deaf or hard of hearing (Kaiser, Kirk, Lachs, & Pisoni, 2003; Kirk, 2000b; Kirk, Pisoni, & Miyamoto, 1997). Secondly, although most traditional clinical tests yield descriptive information about spoken word recognition, they reveal little about the underlying perceptual and cognitive processes employed by listeners who are deaf or hard of hearing. Finally, most traditional tests utilized an auditory-only presentation format that might not adequately characterize the performance of listeners who are deaf or hard of hearing. For example, although some adults and children with sensory aids demonstrate substantial auditory-only word recognition, others obtain high levels of speech understanding only when auditory *and* visual speech cues are available (Bergeson, Pisoni, & Davis, 2003; Hay-McCutcheon, Pisoni, & Kirk, 2005; Kaiser, et al., 2003). Furthermore, the ability to combine and integrate auditory and visual speech information has been found to be an important predictor of speech perception benefit with a sensory aid (Bergeson & Pisoni, 2004; Bergeson, Pisoni, & Davis, 2005; Lachs, Pisoni, & Kirk, 2001b) and thus has important implications for understanding the underlying representation and processing of speech in listeners who use these devices. We saw the need to take a new, translational approach to test development – one that builds upon a body of basic and clinical research concerning spoken word recognition by listeners with normal hearing or hearing loss.

Over the last 15 years, we have created a series of word and sentence recognition tests. These measures are theoretically motivated by a model of spoken word recognition and also incorporate “real-world” stimulus variability in the form multiple talkers and presentation formats. The goal of this research is to enhance our ability to estimate real-world listening ability and to predict benefit from sensory aid use in children with varying degrees of hearing loss. Below we review our past and current test development activities.

The Lexical Neighborhood Tests

Traditional tests of spoken word recognition in children yield descriptive information regarding spoken word recognition abilities but reveal little about the underlying perceptual processes that support spoken word recognition (Meyer & Pisoni, 1999). To address this problem, we developed new measures to assess spoken word recognition in children with cochlear implants. Stimuli were selected according to two criteria. First, the test words had to be familiar to young children with relatively limited vocabularies. Stimuli were drawn from the Child Language Data Exchange System database (MacWhinney & Snow, 1985) which contains transcripts of young children’s verbal exchanges with a caregiver or another

child. All stimulus words were drawn from productions by typically-developing children between the ages of 3–5 years, and thus represent early-acquired vocabulary. The second criterion was that the new measures should be grounded in a model of spoken word recognition and lexical access. Test development was theoretically motivated by the assumptions underlying the Neighborhood Activation Model (Luce & Pisoni, 1998). This theory posits that both word frequency (how often words occur in the language) and phonemic similarity of words in the lexicon influence spoken word recognition performance. One measure of lexical similarity is the number of “neighbors” or words in the lexicon that differ by one phoneme from the target word (Greenberg & Jenkins, 1964; Landauer & Streeter, 1973). Based on computational analyses, word lists were constructed to allow systematic examination of the effects of word frequency, lexical density and word length. The resulting Lexical Neighborhood Test consists of two lists of 50 monosyllabic words, whereas the Multisyllabic Lexical Neighborhood Test (MLNT) consists of two lists of 24 two-to-three syllable words. On each list, half of the words are lexically easy (i.e., occur often and have few phonemically similar neighbors with which to compete for lexical selection) and half are lexically hard (i.e., occur infrequently and come from dense lexical neighborhoods).

In the first of a series of studies, Kirk et al. (1995) used these new measures to examine the effect of lexical characteristics on spoken word recognition performance by children with cochlear implants and to compare their performance on the LNT and MLNT with their performance on the PBK. Participants were 27 children with profound deafness who had used a cochlear implant for at least one year. Test stimuli were presented A-only via live voice at approximately 70 dB SPL; children responded by repeating the word they heard using spoken and/or signed English. The percentage of words correctly identified was significantly higher for lexically easy words than for hard words. Furthermore, spoken word recognition performance was consistently higher on the lexically controlled lists than on the PBK. These early results suggested that pediatric cochlear implant users are sensitive to the acoustic-phonetic similarity among words, that they organize words into similarity neighborhoods in long-term memory, and that they use this structural information in recognizing isolated words in an open-set response format. The results also suggested that the PBK underestimated the participants’ spoken word recognition abilities, perhaps because of the vocabulary constraints inherent in creating phonetically balanced word lists.

In an effort to determine the minimum participant age at which testing with these materials should be attempted, we tested spoken word recognition in 3- and 4-year old children with normal hearing using the PBK, LNT and MLNT (Kluck, Pisoni, & Kirk, 1997). Stimulus presentation and response collection procedures were the same as described above. Both the 3- and 4-year old groups of children completed the tasks with scores that were close to ceiling. High scores and lack of variability in this population precluded establishing test-retest reliability for children with normal hearing. However, performance by these children provided a “benchmark” for assessing the spoken word recognition abilities of children who are deaf or hard of hearing; children with normal hearing who are 3 years of age or older can complete these tasks.

To evaluate the influence of vocabulary knowledge on spoken word recognition test performance, we compared pediatric cochlear implant recipients’ familiarity with words on the PBK, LNT and MLNT using parent ratings on a seven-point scale (Kirk, Sehgal, & Hay-McCutcheon, 2000). Results showed a significant difference in children’s familiarity with test vocabulary across the three tests; words on the LNT were rated as most familiar followed by the MLNT and the PBK, respectively. This suggests that poor performance on the PBK may result, in part, because children with profound deafness are unfamiliar with the test items. There were no significant differences in familiarity between the lexically easy

and hard words within the LNT and the MLNT. Additionally, word familiarity was significantly related to chronological age for the MLNT and PBK, but not the LNT. These results provide further support for the appropriateness of the LNT as a spoken word recognition test for children of widely varying ages.

Studies conducted in our lab and elsewhere demonstrated that performance on the LNT and MLNT is strongly correlated with other traditional measures of spoken word recognition and spoken language processing in children with cochlear implants (Geers, Brenner, & Davidson, 2003; Pisoni, Svirsky, Kirk, & Miyamoto, 1997). We examined the relationship among measures of spoken language processing in pediatric cochlear implants users with exceptionally good speech perception abilities (Pisoni, et al., 1997). The results revealed that performance on the LNT was strongly correlated with open-set sentence recognition, receptive vocabulary, receptive and expressive language abilities, speech intelligibility, nonword repetition, and working memory span. This pattern of results suggests that the LNT and MLNT are measuring the same underlying construct, i.e., phonological coding processes that are used to encode, store, retrieve and manipulate spoken words.

Multiple Talker Versions of the LNT and MLNT

We next developed recorded versions of the LNT and MLNT that incorporated stimulus variability (Kirk, Eisenberg, Martinez, & Hay-McCutcheon, 1999). Three male and three female talkers were recorded producing the stimuli. The intelligibility of the target words was determined by a group of 60 college students with normal hearing (10 listeners per talker). Listeners heard the tokens from each talker presented under headphones in quiet at 70 dB SPL and wrote down the word they heard. The mean intelligibility of the six talkers ranged from 92–100% words correct. The intelligibility results were used to create equivalent audio-recorded LNT and MLNT lists in single-talker and multiple-talker conditions. Tokens from one male were used for the single-talker version because his mean word intelligibility score was closest to grand mean for the six talkers. The multiple-talker version contained tokens produced by the remaining two males and three females.

Test-retest reliability and inter-list equivalency of the audio-recorded versions of the LNT and MLNT were evaluated in 16 pediatric cochlear implant users with profound, prelingual hearing loss (Kirk, et al., 1999). Stimuli were presented at approximately 70 dB SPL via loudspeaker at 0 degrees azimuth. Children responded by repeating the word they heard in spoken and/or signed English. Each child was tested twice, with the time between sessions ranging from 3 hours to 15 days, depending on availability. The LNT and MLNT had high test-retest reliability (r values ranged from .83 to .95 across tests and lists) and no significant inter-list differences were noted. Spoken word recognition scores were higher for lexically easy words than for lexically hard words, replicating the earlier findings obtained with live voice stimuli.

Stimulus Variability Effects on Spoken Word Recognition by Children with Cochlear Implants

Kirk et al. (1998) used the recorded versions of the LNT and MLNT to examine the effects of lexical competition, talker variability, and word length on A-only spoken word recognition in 20 children with prelingual, profound deafness who used a CI. Each participant was administered one LNT and one MLNT word list in the single-talker condition and the remaining LNT list and MLNT list in the multiple-talker condition. Lexically easy words were recognized with greater accuracy than lexically hard words regardless of talker condition or word length. Multisyllabic words were identified better than monosyllabic words, presumably because longer words contain more linguistic redundancy and have few lexical neighbors with which to compete for selection. The children in this

study also were sensitive to talker characteristics, with differences in performance noted between the single-talker and multiple-talker lists.

Utility of the LNT and MLNT

The LNT and MLNT provide reliable measures of auditory-only spoken word recognition abilities of children with profound hearing loss who use a CI. Their utility is demonstrated by the fact that since their introduction in the mid 1990's, the LNT and MLNT have been used to help determine CI candidacy and postimplant benefit in FDA pediatric clinical trials of new CI systems. These tests also have been used by a number of different researchers to examine factors that influence cochlear implant outcomes (Eisenberg, et al., 2006; Geers, et al., 2003; Kirk, Pisoni, & Miyamoto, 2000).

An Auditory Lexical Sentence Test for Children

Eisenberg et al. (2002) created lexically controlled sentences from the subset of words used to develop the LNT and MLNT. Sentence lists were generated using the definitions and procedures of Kirk et al. (1995) in accordance with the Neighborhood Activation Model (Luce & Pisoni, 1998). Three key words were used in constructing each of the 5- to 7-word sentences. These sentences were combined into two lists, each with five practice and 20 test sentences that were syntactically correct but semantically neutral. One list contained sentences with lexically easy key words; the other list contained key words that were lexically hard. Eisenberg et al. (2002) used these sentences to investigate the effects of word frequency and lexical density on the children's' recognition of isolated words and the same words produced in a sentence context. The lexically controlled sentences and the isolated key words comprising the sentences were audio-recorded separately as two different tests and used in three experiments. In Experiment 1, 48 children with normal hearing between the ages of 5–12 years repeated the isolated words and sentences at one of six different levels to generate performance-intensity functions. In Experiment 2, 12 normal hearing children aged 5–14 years repeated the words and sentences under spectrally degraded conditions intended to simulate CI speech processing. In Experiment 3, 12 children with CIs aged 5–14 years repeated the unprocessed stimuli. Children also completed a test of vocabulary recognition. Figure 1 illustrates the results of all three experiments. Across three experiments, sentences containing lexically easy key words were recognized with greater accuracy than sentences containing lexically hard key words. Sentence scores were significantly higher than word scores for the children with normal hearing and nine high-performing children with CIs. Three low-performing children with CIs showed the opposite pattern for isolated word and sentence stimuli. A statistically significant relationship was observed between chronological age and sentence scores for children with normal hearing who heard degraded speech. For children with CIs, the relationship between language abilities and spoken word recognition was strong and significant. This result demonstrates the influence of linguistic knowledge on phonological processing of words.

Summary of Auditory Test Development

The results of our studies concerning lexical effects on auditory spoken word and sentence recognition by children with CIs suggest that the ability to encode novel sound patterns is a fundamental prerequisite for building a grammar from the spoken language input a child receives. These new spoken word recognition measures have been found to be important predictors of spoken language acquisition (Pisoni, et al., 1997). Furthermore, listeners with hearing aids and/or cochlear implants are sensitive to differences among talkers, despite receiving a degraded auditory signal. Furthermore, their spoken word recognition is influenced by talker characteristics. Perceptually robust speech perception tests can be used to assess several aspects of spoken word recognition in the clinic and laboratory that appear

to generalize and transfer to conditions in natural listening situations, where the listener is faced with many sources of variability.

Multimodal Spoken Word Recognition and Speech Perception

Under natural conditions, listeners use both auditory and visual speech cues to extract meaning from speech signals containing many sources of variability introduced by different talkers, dialects, speaking rates, and background noise. Sumbly and Pollack (1954) were the first researchers to document that the addition of visual speech cues to the auditory speech signal yielded substantial improvements in speech perception for listeners with normal hearing. Visual speech cues are especially helpful under adverse listening conditions such as noise. Similar results have been obtained by other investigators for listeners with normal hearing (Demorest, Bernstein, & DeHaven, 1996; Massaro & Cohen, 1995; Sommers, Tye-Murray, & Spehar, 2005; Summerfield, 1987) and for listeners who are deaf or hard of hearing (Erber, 1971, 1972; Moody-Antonio, et al., 2005; Walden, Prosek, & Worthington, 1975). The cognitive processes by which individuals combine and integrate auditory and visual speech information with lexical and syntactic knowledge have become an important area of research in the field of speech perception. It has been hypothesized that visual cues enhance speech understanding because they provide segmental and suprasegmental information that is complementary (i.e., place of articulation) to the acoustic speech cues, and because they reduce the attentional demands placed on the auditory signal (Grant & Seitz, 2000; Summerfield, 1987). Multimodal spoken word recognition therefore appears to be more than the simple addition of auditory and visual information (Bernstein, Demorest, & Tucker, 2000). A well-known example of the robustness of multimodal spoken word recognition is the “McGurk” effect (McGurk & MacDonald, 1976). When presented with an auditory stimulus /ba/ and a visual stimulus /ga/, many listeners report perceiving /da/. Thus, information from separate sensory modalities can be combined to produce percepts that differ predictably from either the auditory or visual percept alone. However, substantial individual variability is noted in the ability to integrate the two types of speech cues (Demorest & Bernstein, 1992). One factor that appears to influence auditory-visual, or multimodal integration is degree of hearing loss. Individuals with lesser degrees of hearing loss obtain better auditory-only (A) and auditory-plus-visual (AV) performance compared to individuals with greater degrees of hearing loss (Erber, 1972; Seewald, Ross, Giolas, & Yonovitz, 1985; Tillberg, Ronneberg, Svard, & Ahlner, 1996). The nature of a listener’s early linguistic experience also influences multimodal speech perception. Individuals with an earlier onset of deafness demonstrate better V-only spoken word recognition compared to listeners with a late onset of deafness (Bergeson, et al., 2005; Tillberg, et al., 1996). Grant et al. (Grant & Seitz, 1998) proposed that variability in multimodal spoken word recognition depends not only on how lexical access is influenced by an individuals’ access to auditory and visual speech cues, but also on the processes by which the two types of cues are integrated, and by the impact of top-down contextual constraints and memory processes. They concluded that approximately 20–30% of the variance in multimodal spoken word recognition could be accounted for by differences in integration efficiency [but see Massaro and Cohen (2000) for an alternative point of view].

Multimodal Spoken Word Recognition in Children with Cochlear Implants

The ability to integrate and make use of auditory and visual speech cues is an important predictor of speech perception benefit with a sensory aid (Bergeson, et al., 2003, 2005; Lachs, Pisoni, & Kirk, 2001a) and thus has important implications for understanding the underlying representation of speech in listeners who use these devices. Adults who are deaf or hard of hearing use information from the auditory and visual modalities to access common, multimodal lexical representations in memory. In contrast, children who are deaf

or hard of hearing must rely on the auditory signal provided by a sensory aid to develop speech perception, speech production and language skills.

Lachs et al. (2001b) examined how prelingually deafened children combine visual information available in the talker's face with auditory speech cues provided by their cochlear implant to enhance spoken language comprehension. Twenty-seven children with cochlear implants identified spoken sentences from the Common Phrases Test (Robbins, Renshaw, & Osberger, 1995) presented via live voice under A and AV presentation formats. Five additional measures of spoken word recognition were administered to assess A-only speech perception skills. Speech intelligibility also was measured to assess the speech *production* abilities of these children. A measure of visual enhancement, R_a , was used to assess the gain in performance provided by the AV presentation format relative to the maximum possible performance obtainable in the A format. R_a is calculated using Equation 1

$$R_a = (AV - A) / (100 - A) \quad \text{Eq. (1)}$$

where A and AV represent the percent correct scores obtained in the respective presentation formats, normalized to the amount by which the speech intelligibility could have possibly improved above scores in the A format. Another measure of audiovisual gain R_v was computed using scores in the V-only and AV conditions. R_v is calculated using Equation 2,

$$R_v = (AV - V) / (100 - V) \quad \text{Eq. (2)}$$

where V and AV represent the percent correct scores in the respective presentation formats. This alternative measure of gain may be more appropriate for individuals with greater degrees of hearing loss because they may rely more on visual input than on auditory input for speech perception (Grant, Walden, & Seitz, 1998). The results demonstrated that children who are deaf or hard of hearing who were best at recognizing spoken words through listening alone also were better at combining the complementary sensory information about speech articulation available under AV conditions. We found that children who are deaf or hard of hearing who received more benefit from AV stimulus presentation also produced more intelligible speech.

The Multimodal Lexical Sentence Test for Children (MLST-C)

Our results suggest a close link between speech perception and production and a common underlying linguistic basis for auditory-visual enhancement effects. Despite the importance of audiovisual speech integration in daily listening situations, few pediatric tests of multimodal spoken word recognition are available for use in clinic and research settings. Over the last five years we have conducted translational research to develop such a multimodal test of spoken word recognition for children who are deaf or hard of hearing. Our new test utilizes stimuli with controlled lexical characteristics and incorporates stimulus variability in the form of multiple talkers. Below we describe test development and present preliminary data obtained from children who are deaf or hard of hearing.

Sentence Creation and Recording

The MLST-C consists of 100 semantically-neutral sentences of five-to-seven words in length. Each sentence contains three key words which are controlled for the lexical characteristics of word frequency and lexical density. Fifty sentences were those created by Eisenberg et al (2002). An additional 50 sentences were created in order to allow for orthogonal key word combinations of word frequency and lexical density (Krull, et al.,

2010). Within a sentence, all key words represent the same lexical category. Key words are drawn from four different lexical categories: 1) High Frequency-Sparse, 2) High-Frequency-Dense, 3) Low-Frequency-Sparse, and 4) Low-Frequency-Dense. Category 1 contains key words that are easy to identify in a spoken word recognition task, whereas Category 4 contains key words that are hard to identify. We have used words from these two lexical categories in our previous tests of spoken word recognition (Eisenberg, et al., 2002; Holt, Kirk, & Hay-McCutcheon, 2011; Kirk, et al., 1995). Table 1 illustrates the four lexical categories and provides an example sentence for each.

Five male and five female talkers were audiovisually recorded in a high definition format producing the sentences in a conversational speaking style. Fixed distances were maintained between each talker and the camera, microphone, lighting and background. All talkers wore a black t-shirt and the video shot composition was adjusted to accommodate each talker's physiological characteristics. Stimulus editing was carried out to ensure that each sentence began and ended with the talker's articulators in a resting state. Next, the root mean square amplitude of each sentence was equated to the mean computed across all talkers and sentences.

Determining Sentence Intelligibility

Kirk et al. (2011) examined the intelligibility of each sentence as a function of talker, presentation format and key word lexical category. Adult participants with normal hearing (N=189) were tested to determine sentence intelligibility. Participants were divided into three groups per talker, and each participant group was administered sentences produced by a single talker in only one of the three possible presentation formats (A, V or AV). The auditory speech signal was presented at 60 dB mixed with speech-shaped noise at a -2 dB signal-to-noise ratio. Responses were scored as the percent of key words correctly identified in each sentence. The data were subjected to a repeated measures analysis of variance with word frequency and lexical density as within-subject factors, and presentation format and talker as between-subject factors. There were significant main effects of lexical density, presentation format and talker. There also were numerous significant interactions, including an interaction between word frequency and lexical density. Scores in the three presentation formats differed significantly from one another, with very poor performance in the V format and very good performance in the AV format. Across talkers, key word accuracy ranged from 5–15% in the V format, from 68–95% correct in the A format, and from 87–98% correct in the AV format. Words from sparse lexical neighborhoods were identified with significantly greater accuracy than words from dense neighborhoods but there was no effect of word frequency. Presumably, all of the key words in this pediatric sentence test were highly familiar to the adult participants with normal hearing.

Generating Equivalent Lists

The sentence intelligibility data were submitted to psychometric analyses in order to generate multiple lists of sentences that were equivalent in difficulty (Gotch, et al., 2011). Sentences were inspected across talkers, within lexical category, and within presentation format to control for level of difficulty as judged by the sentence mean. A given sentence from a particular talker was eliminated if it deviated too far from the mean intelligibility of that sentence across all talkers. Sentences also were eliminated if they had negative item-total correlations for any two of the three presentation formats. List generation also met the following constraints: 1) balanced in terms of talker gender; 2) at least two male and two female talkers per list; 3) equal number of sentences from each of the four lexical categories; 4) similar mean intelligibility scores across lists *within* each presentation format. Thus, mean intelligibility scores for the lists differ *across* presentation formats ($V < A < AV$), but

they are similar when all of the lists are presented in any one of the three presentation formats. This process resulted in 21 lists with eight sentences and 24 key words per list.

Performance of Children who are deaf or hard of hearing on the MLST-C

Kirk et al. (2011) conducted a preliminary investigation to determine whether the MLST-C sentence lists were equivalent, valid and reliable in the population for whom they were developed, children who are deaf or hard of hearing.

Participants

Thirty-one children have been tested to date. All participants met the following inclusion criteria: 1) aged 4–12 years with a minimum receptive vocabulary age of 3 years, 2) spoken English as the primary language, 3) normal or corrected-to-normal vision, 4) no additional handicapping conditions, 5) a bilateral symmetrical sensorineural hearing loss with pure tone average thresholds >30dB HL, and 6) normal tympanometry bilaterally. All participants used hearing aid(s), cochlear implant(s) or a combination of the two sensory aids. Table 2 presents the participant characteristics. All participants were paid for their time.

Methods

Testing was carried out at The University of Iowa, the House Research Institute in Los Angeles, and Children's Memorial Hospital in Chicago. Participants were tested over two or three sessions, depending on fatigue. Participants first completed vision screening, an audiological assessment and a receptive vocabulary test to insure that they meet inclusion criteria. Speech perception testing was carried out in a sound-attenuating booth. Participants were administered the 21 sentence lists twice. Time between repeated administrations of the sentence lists ranged from several hours to a week or more. Each participant was administered one third (seven) of the lists per presentation format (V, A or AV). Sentence assignment to presentation format was counterbalanced across participants. Auditory speech stimuli were presented in quiet at 60 dB SPL in the A and AV presentation formats. Participants respond by repeating the sentence they heard and were encouraged to guess if they were unsure. Participant responses also were audio recorded for subsequent scoring verification. At the time the sentence lists were administered for a second time, participants also were given the HINT-C, an auditory sentence recognition test.

Results

Figure 2 presents the average percent of key words correctly identified by each sensory aid group as a function of presentation format. A repeated measures analysis of variance was conducted. Because we observed floor effects in the V condition, only scores from the A and AV presentation formats were included in this analysis. The children demonstrated significantly poorer performance in the A condition than in the AV condition; the percent of key words correctly identified was 81% vs. 91% respectively. The average performance for children with hearing aids was 93% correct compared with 79% correct for children with cochlear implants. This translates to a score difference of 1.5 sentences, which is not significant.

Figure 3 illustrates the the percent of key words correctly identified as a function of list number within a given presentation format. As mentioned above, each participant was administered 1/3 of the sentences in the V, A and AV presentation formats. The order and sentence assignments were counterbalanced across participants. Because we have not yet completed testing, some sentence blocks have been administered to a larger group of participants compared others. In Figure 3, white bars represent lists that have been administered to slightly more than 20 participants and grey bars represent lists that have

been administered to fewer than 10 participants, typically between three and seven. In order to evaluate list equivalency, one of our steps has been to focus on effect sizes to evaluate differences among the various lists within a given presentation format. Therefore, each list was compared with the other 20 lists administered in a given presentation format. For example, List 1 was compared with Lists 2–20. A list was considered potentially nonequivalent if it differed by more than a half standard deviation from 60% (or 12) of the other lists with which it was compared within a given presentation format. Potentially problematic, or non-equivalent lists are indicated with an asterisk in Figure 3. None of the lists in the V presentation format was problematic. Note that the V presentation format was much more difficult than either the A or AV presentation formats. There are six potentially problematic lists in the A condition and three in the AV condition. With one exception (List 15 in the AV format) these problematic lists are from sets that have been administered to relatively few participants to date, typically six or seven participants. That said, given that the goal is to create several test forms comprised of two-to- three lists, there appears to be a sufficient number of lists to reach this goal while minimizing repetition and maintaining test quality.

Table 3 presents the correlations between performance on the first and second administration of the MLST-C. All correlations were consistently high ($r = .85$, $p < .01$) indicating very good test-retest reliability (i.e., stability) of the MLST-C scores. Figure 4 illustrates the relationship between performance in the A presentation format of the MLST-C and performance on the HINT-C. There was a significant correlation between performance on the two measures ($r = 0.80$, $p < .01$) providing validity evidence to support use of the scores from the MLST-C.

Conclusion

Our research has yielded a set of theoretically motivated tests of spoken word recognition that incorporate “real-world” stimulus variability in the form of multiple talkers and presentation formats. By tapping into the more central cognitive processes used during multimodal integration, perceptual normalization and lexical discrimination, the MLST-C may reveal information about the way in which children encode, store and retrieve items from their mental lexicons. Such tests may also further our understanding of individual variations in spoken word recognition performance by children who are deaf or hard of hearing.

The development of our new multimodal sentence test is based on strong psychometric principles following standard design principles at each step of the process. Preliminary data suggest that when our test forms (consisting of two or three equivalent lists per form) are built for clinical and research purposes, the scores will be supported for such use by the reliability and validity data being gathered. This work will help to ensure more accurate assessment of spoken word recognition in children who are deaf or hard of hearing. The test contains multiple lists that are equivalent in all three presentation formats, making them useful in auditory processing of spoken language as well as auditory-visual enhancement. Current work is ongoing to gather additional reliability, validity, and list equivalency data in a larger group of children who are deaf or hard of hearing who use sensory aids. We believe these new tests will lead to better assessment paradigms that can aid in the selection of sensory aids and inform the development of aural rehabilitation programs.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

David Pisoni has been a mentor and colleague throughout much of this work. His contributions were invaluable. Rachael Frush Holt, Vidya Krull and Sangsook Choi conducted test development research at Indiana University School of Medicine and Purdue University that is cited in this paper. We thank Jennifer Karpik, Amy Martinez, Diane Hammes Ganguly, Nicholas Giuliani, Amanda Silberer, Nathaniel Wisecup, Lisa Weber and Susan Stentz for assistance with stimulus preparation, editing and/or data collection. Most importantly, we thank our pediatric participants and their families for their important contributions to our research. This work was supported by grants R01 D00064 and R01 DC00875.

This research was supported by NIDCD grants awarded to R. Miyamoto and K. I. Kirk (R01 DC00064) and K. I. Kirk (R01 DC00875)

References

- Bergeson, TR.; Pisoni, DB. Audiovisual speech perception in deaf adults and children following cochlear implantation. In: Calvert, GA.; Spence, C.; Stein, BE., editors. *The Handbook of Multisensory Perception*. Cambridge, MA: MIT Press; 2004. p. 749-771.
- Bergeson TR, Pisoni DB, Davis RA. Longitudinal study of audiovisual speech perception by children with hearing loss who have cochlear implants. *The Volta Review*. 2003; 103:347-370. [PubMed: 21743753]
- Bergeson TR, Pisoni DB, Davis RA. Development of audiovisual comprehension skills in prelingually deaf children with cochlear implants. *Ear and Hearing*. 2005; 26:149-164. [PubMed: 15809542]
- Bernstein LE, Demorest ME, Tucker PE. Speech perception without hearing. *Perception & Psychophysics*. 2000; 62:233-252. [PubMed: 10723205]
- Demorest ME, Bernstein LE. Sources of variability in speechreading sentences: A generalizability analysis. *Journal of Speech and Hearing Research*. 1992; 35:876-891. [PubMed: 1405543]
- Demorest ME, Bernstein LE, DeHaven G. Generalizability of speechreading performance on nonsense syllables, words, and sentences: Subjects with normal hearing. *Journal of Speech and Hearing Research*. 1996; 39:697-713. [PubMed: 8844551]
- Desjardin JL, Ambrose SE, Martinez AS, Eisenberg LS. Relationships between speech perception abilities and spoken language skills in young children with hearing loss. *Int J Audiol*. 2009; 48(5): 248-259. [PubMed: 19842800]
- Dorman MF, Gifford RH, Spahr AJ, McKarns SA. The benefits of combining acoustic and electric stimulation for the recognition of speech, voice and melodies. *Audiol Neurootol*. 2008; 13(2):105-112. [PubMed: 18057874]
- Dorman MF, Loizou PC, Kemp LL, Kirk KI. Word recognition by children listening to speech processed into a small number of channels: data from normal-hearing children and children with cochlear implants. *Ear Hear*. 2000; 21(6):590-596. [PubMed: 11132785]
- Eisenberg LS. Current state of knowledge: speech recognition and production in children with hearing impairment. *Ear Hear*. 2007; 28(6):766-772. [PubMed: 17982364]
- Eisenberg LS, Johnson KC, Martinez AS, Cokely CG, Tobey EA, Quittner AL, et al. Speech recognition at 1-year follow-up in the childhood development after cochlear implantation study: Methods and preliminary findings. *Audiol Neurootol*. 2006; 11(4):259-268. [PubMed: 16699259]
- Eisenberg LS, Kirk KI, Martinez AS, Ying EA, Miyamoto RT. Communication abilities of children with aided residual hearing: comparison with cochlear implant users. *Arch Otolaryngol Head Neck Surg*. 2004; 130(5):563-569. [PubMed: 15148177]
- Eisenberg LS, Martinez AS, Holowecky SR, Pogorelsky S. Recognition of lexically controlled words and sentences by children with normal hearing and children with cochlear implants. *Ear Hear*. 2002; 23(5):450-462. [PubMed: 12411778]
- Erber NP. Auditory and audiovisual reception of words in low-frequency noise by children with normal hearing and by children with impaired hearing. *Journal of Speech and Hearing Research*. 1971; 14:496-512. [PubMed: 5163883]
- Erber NP. Auditory, visual, and auditory-visual recognition of consonants by children with normal and impaired hearing. *Journal of Speech and Hearing Disorders*. 1972; 15:413-422.

- Firszt JB, Holden LK, Reeder RM, Skinner MW. Speech recognition in cochlear implant recipients: comparison of standard HiRes and HiRes 120 sound processing. *Otol Neurotol.* 2009; 30(2):146–152. [PubMed: 19106769]
- Geers AE. Factors influencing spoken language outcomes in children following early cochlear implantation. *Adv Otorhinolaryngol.* 2006; 64:50–65. [PubMed: 16891836]
- Geers AE, Brenner C, Davidson L. Factors associated with development of speech perception skills in children implanted by age five. *Ear and Hearing.* 2003; 24:24S–35S. [PubMed: 12612478]
- Gifford RH, Dorman MF, Spahr AJ, Bacon SP. Auditory function and speech understanding in listeners who qualify for EAS surgery. *Ear Hear.* 2007; 28(2 Suppl):114S–118S. [PubMed: 17496661]
- Gifford RH, Olund AP, Dejong M. Improving speech perception in noise for children with cochlear implants. *J Am Acad Audiol.* 2011; 22(9):623–632. [PubMed: 22192607]
- Gifford RH, Shallop JK, Peterson AM. Speech recognition materials and ceiling effects: considerations for cochlear implant programs. *Audiol Neurootol.* 2008; 13(3):193–205. [PubMed: 18212519]
- Gotch, CM.; French, BF.; Kirk, KI.; Prusick, L.; Eisenberg, LS.; Martinez, A., et al. Deriving equivalent forms of a multimodal lexical sentence test. Paper presented at the 13th Symposium on Cochlear Implants in Children; 2011.
- Grant KW, Seitz PF. Measures of auditory-visual integration in nonsense syllables and sentences. *The Journal of the Acoustical Society of America.* 1998; 104(4):2438–2450. [PubMed: 10491705]
- Grant KW, Seitz PF. The use of visible speech cues for improving auditory detection of specific words. *Journal of the Acoustical Society of America.* 2000; 108:1197–1208. [PubMed: 11008820]
- Grant KW, Walden BE, Seitz PF. Auditory-visual speech recognition by hearing-impaired subjects: Consonant recognition, sentence recognition, and auditory-visual integration. *The Journal of the Acoustical Society of America.* 1998; 103(5):2677–2705. [PubMed: 9604361]
- Greenberg JH, Jenkins JJ. Studies in the psychological correlates of the sound system of American English. *Word.* 1964; 20:157–177.
- Haskins, H. A phonetically balanced test of speech discrimination for children. Northwestern University; Evanston, IL: 1949.
- Hay-McCutcheon M, Pisoni DB, Kirk KI. Audiovisual speech perception in elderly cochlear implant recipients. *Laryngoscope.* 2005; 115:1887–1894. [PubMed: 16222216]
- Holt RF, Beer J, Kronenberger WG, Pisoni DB, Lalonde K. Contribution of Family Environment to Pediatric Cochlear Implant Users' Speech and Language Outcomes: Some Preliminary Findings. *J Speech Lang Hear Res.* 2012
- Holt RF, Kirk KI, Eisenberg LS, Martinez AS, Campbell W. Spoken word recognition development in children with residual hearing using cochlear implants and hearing aids in opposite ears. *Ear Hear.* 2005; 26(4 Suppl):82S–91S. [PubMed: 16082270]
- Holt RF, Kirk KI, Hay-McCutcheon M. Assessing multimodal spoken word-in-sentence recognition in children with normal hearing and children with cochlear implants. *J Speech Lang Hear Res.* 2011; 54(2):632–657. [PubMed: 20689028]
- Hudgins C, Hawkins J, Karling J, Stevens S. The development of recorded auditory tests for measuring hearing loss of speech. *Laryngoscope.* 1947; 57:57–89. [PubMed: 20287775]
- Kaiser AR, Kirk KI, Lachs L, Pisoni DB. Talker and lexical effects on audiovisual word recognition by adults with cochlear implants. *Journal of Speech, Language, and Hearing Research.* 2003; 46:390–404.
- Kirk, KI. Challenges in the clinical investigation of cochlear implant outcomes. In: Niparko, JK.; Kirk, KI.; Mellon, NK.; McConkey Robbins, A.; Tucci, DL.; Wilson, BS., editors. *Cochlear Implants: Principles and Practices.* Philadelphia: Lippincott Williams & Williams; 2000a. p. 225-259.
- Kirk KI. Cochlear implants: New developments and results. *Current Opinion in Otolaryngology & Head and Neck Surgery.* 2000b; 8:415–420.
- Kirk, KI.; Eisenberg, LS.; French, BF.; Prusick, L.; Martinez, A.; Ganguly, DH., et al. Development of the Multimodal Lexical Sentence Test for Children. Paper presented at the 13th Symposium on Cochlear Implants in Children; 2011.

- Kirk, KI.; Eisenberg, LS.; French, BF.; Prusick, L.; Martinez, AS.; Ganguly, DH., et al. The Multimodal Lexical Sentence Test for Children: Performance of children with hearing loss. Paper presented at the 13th Symposium on Cochlear Implants in Children; 2011.
- Kirk KI, Eisenberg LS, Martinez AS, Hay-McCutcheon M. Lexical neighborhood test: test-retest reliability and interlist equivalency. *The Journal of the American Academy of Audiology*. 1999; 10:113–123.
- Kirk KI, Hay-McCutcheon M, Sehgal ST, Miyamoto RT. Speech perception in children with cochlear implants: Effects of lexical difficulty, talker variability, and word length. *Annals of Otolaryngology, Rhinology, and Laryngology*. 1998; 109(12 Part 2):79–81.
- Kirk KI, Pisoni DB, Miyamoto RC. Effects of stimulus variability on speech perception in listeners with hearing impairment. *Journal of Speech, Language, and Hearing Research*. 1997; 40:1395–1405.
- Kirk, KI.; Pisoni, DB.; Miyamoto, RT. Lexical discrimination by children with cochlear implants: effects of age at implantation and communication mode. In: Waltzman, SB.; Cohen, NL., editors. *Cochlear Implants*. New York: Thieme; 2000.
- Kirk KI, Pisoni DB, Osberger MJ. Lexical effects on spoken word recognition by pediatric cochlear implant users. *Ear Hear*. 1995; 16(5):470–481. [PubMed: 8654902]
- Kirk KI, Sehgal ST, Hay-McCutcheon M. Comparison of children's familiarity with tokens on the PBK, LNT, and MLNT. *Ann Otol Rhinol Laryngol Suppl*. 2000; 185:63–64. [PubMed: 11141008]
- Gluck, M.; Pisoni, DB.; Kirk, KI. Research in Spoken Language Processing No. 21. 1997. Performance of normal-hearing children on open-set speech perception tests; p. 349-366.
- Krull V, Choi S, Kirk KI, Prusick L, French BF. Lexical effects on spoken-word recognition in children with normal hearing. *Ear Hear*. 2010; 31(1):102–114. [PubMed: 19701087]
- Lachs L, Pisoni DB, Kirk KI. Use of audiovisual information in speech perception by prelingually deaf children with cochlear implants: A first report. *Ear and Hearing*. 2001a; 22:236–251. [PubMed: 11409859]
- Lachs L, Pisoni DB, Kirk KI. Use of audiovisual information in speech perception by prelingually deaf children with cochlear implants: a first report. *Ear Hear*. 2001b; 22(3):236–251. [PubMed: 11409859]
- Landauer TK, Streeter LA. Structural differences between common and rare words: Failure of equivalence assumptions for theories of word recognition. *Journal of Verbal Learning and Verbal Behavior*. 1973; 12:119–131.
- Loebach JL, Bent T, Pisoni DB. Multiple routes to the perceptual learning of speech. *J Acoust Soc Am*. 2008; 124(1):552–561. [PubMed: 18646998]
- Loebach JL, Pisoni DB. Perceptual learning of spectrally degraded speech and environmental sounds. *J Acoust Soc Am*. 2008; 123(2):1126–1139. [PubMed: 18247913]
- Luce PA, Pisoni DB. Recognizing spoken words: the neighborhood activation model. *Ear Hear*. 1998; 19(1):1–36. [PubMed: 9504270]
- Mackersie CL. Tests of speech perception abilities. *Current Opinions in Otolaryngology & Head and Neck Surgery*. 2002; 10:392–397.
- MacWhinney B, Snow C. The child language data exchange system. *Journal of Child Language*. 1985; 12:271–296. [PubMed: 4019604]
- Massaro DW, Cohen MM. Perceiving talking faces. *Current Directions in Psychological Science*. 1995; 4:104–109.
- Massaro DW, Cohen MM. Tests of auditory-visual integration efficiency within the framework of the fuzzy logical model of perception. *Journal of the Acoustical Society of America*. 2000; 108:784–789. [PubMed: 10955645]
- McGurk H, MacDonald JW. Hearing lips and seeing voices. *Nature*. 1976; 264:746–748. [PubMed: 1012311]
- Meyer TA, Pisoni DB. Some computational analyses of the PBK test: Effects of frequency and lexical density on spoken word recognition. *Ear and Hearing*. 1999; 20(4):363–371. [PubMed: 10466571]
- Moody-Antonio S, Takayanagi S, Masuda A, Auer ET Jr, Fisher L, Bernstein LE. Improved speech perception in adult congenitally deafened cochlear implant recipients. *Otol Neurotol*. 2005; 26(4):649–654. [PubMed: 16015162]

- Mullenix JW, Pisoni DB, Martin CS. Some effects of talker variability on spoken word recognition. *Journal of the Acoustical Society of America*. 1989; 85:365–378. [PubMed: 2921419]
- Nygaard, LC.; Sommers, MS.; Pisoni, DB. Effects of speaking rate and talker variability on the representation of spoken words in memory. *International Conference on Spoken Language Processing*; 1992.
- Pisoni DB, Kronenberger WG, Roman AS, Geers AE. Measures of digit span and verbal rehearsal speed in deaf children after more than 10 years of cochlear implantation. *Ear Hear*. 2011; 32(1 Suppl):60S–74S. [PubMed: 21832890]
- Pisoni, DB.; Svirsky, MA.; Kirk, KI.; Miyamoto, RT. Looking at the “Stars”: A first report on the intercorrelations among measures of speech perception, intelligibility and language development in pediatric cochlear implant users. Bloomington, IN: Indiana University; 1997.
- Robbins, AM.; Renshaw, JJ.; Osberger, MJ. Common Phrases Test. Indianapolis, IN: Indiana University School of Medicine; 1995.
- Seewald RC, Ross M, Giolas TG, Yonovitz A. Primary modality for speech perception in children with normal and impaired hearing. *Journal of Speech and Hearing Research*. 1985; 28:36–46. [PubMed: 3981996]
- Skinner MW, Holden LK, Whitford LA, Plant KL, Psarros C, Holden TA. Speech recognition with the nucleus 24 SPEAK, ACE, and CIS speech coding strategies in newly implanted adults. *Ear Hear*. 2002; 23(3):207–223. [PubMed: 12072613]
- Sommers MS, Nygaard LC, Pisoni DB. Stimulus variability and spoken word recognition. I. Effects of variability in speaking rate and overall amplitude. *J Acoust Soc Am*. 1994; 96(3):1314–1324. [PubMed: 7962998]
- Sommers MS, Tye-Murray N, Spehar B. Auditory-visual speech perception and auditory-visual enhancement in normal-hearing younger and older adults. *Ear and Hearing*. 2005; 26:263–275. [PubMed: 15937408]
- Sumby WH, Pollack I. Visual contribution of speech intelligibility in noise. *Journal of the Acoustical Society of America*. 1954; 26:212–215.
- Summerfield, Q. Some preliminaries to a comprehensive account of audio-visual speech perception. In: Dodd, B.; Campbell, R., editors. *Hearing by Eye: The Psychology of Lipreading*. Hillsdale, NJ: Lawrence Erlbaum Associates; 1987. p. 3-51.
- Tillberg I, Ronneberg J, Svard I, Ahlner B. Audio-visual speechreading in a group of hearing aid users: The effects of onset age, handicap age and degree of hearing loss. *Scandinavian Audiology*. 1996; 25:267–272. [PubMed: 8976000]
- Walden BE, Prosek RB, Worthington DW. Auditory and audiovisual feature transmission in hearing-impaired adults. *Journal of Speech and Hearing Research*. 1975; 18:272–280.

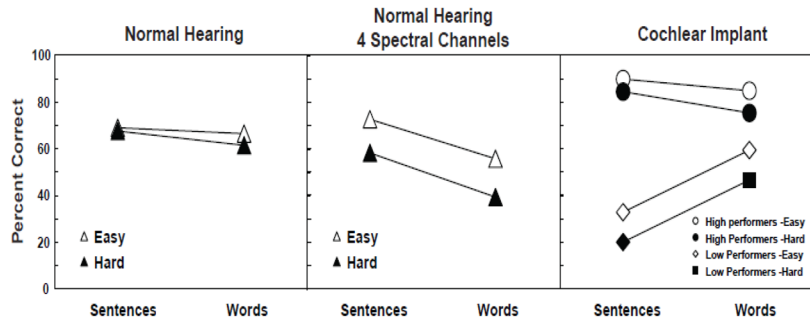


Figure 1. Average percent of lexically easy and hard words plotted as a function of sentence and word contexts.

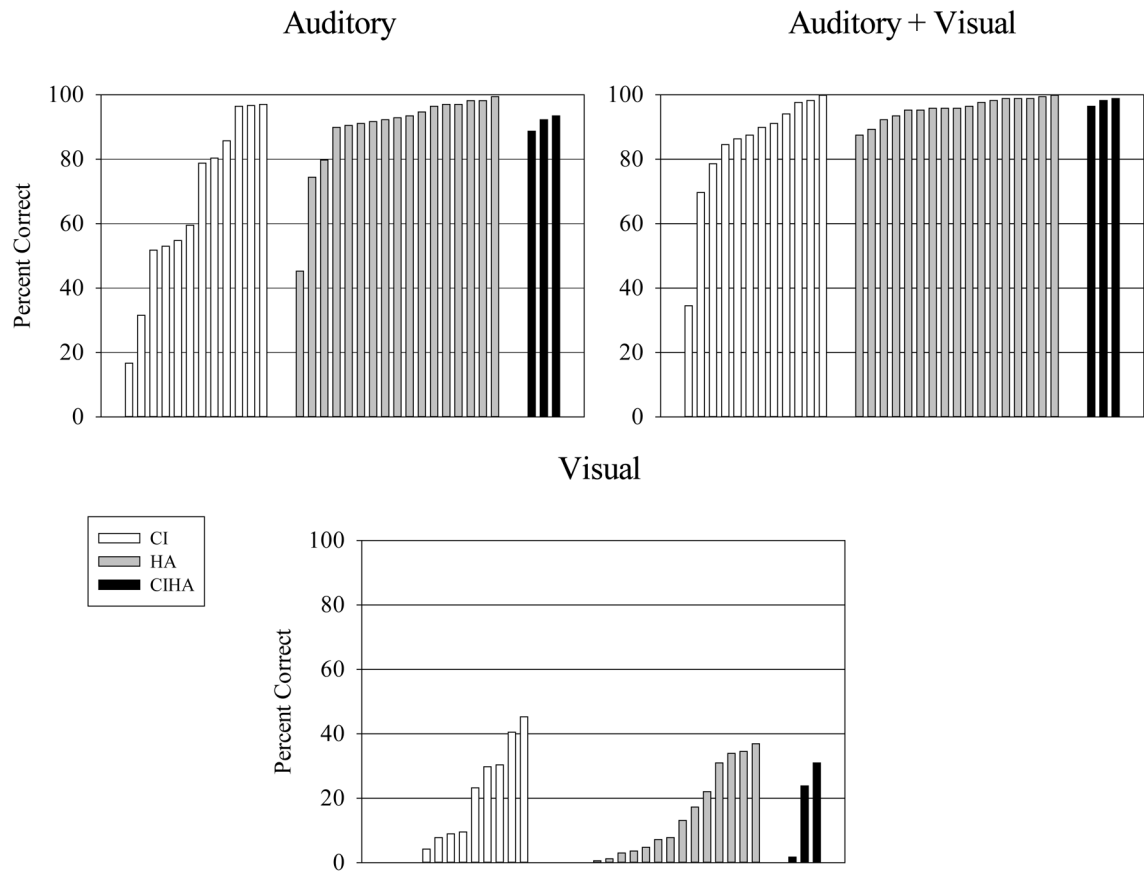


Figure 2. Percent of key words correctly identified by children with hearing loss as a function of presentation format and type of sensory aid.

Auditory

Auditory + Visual

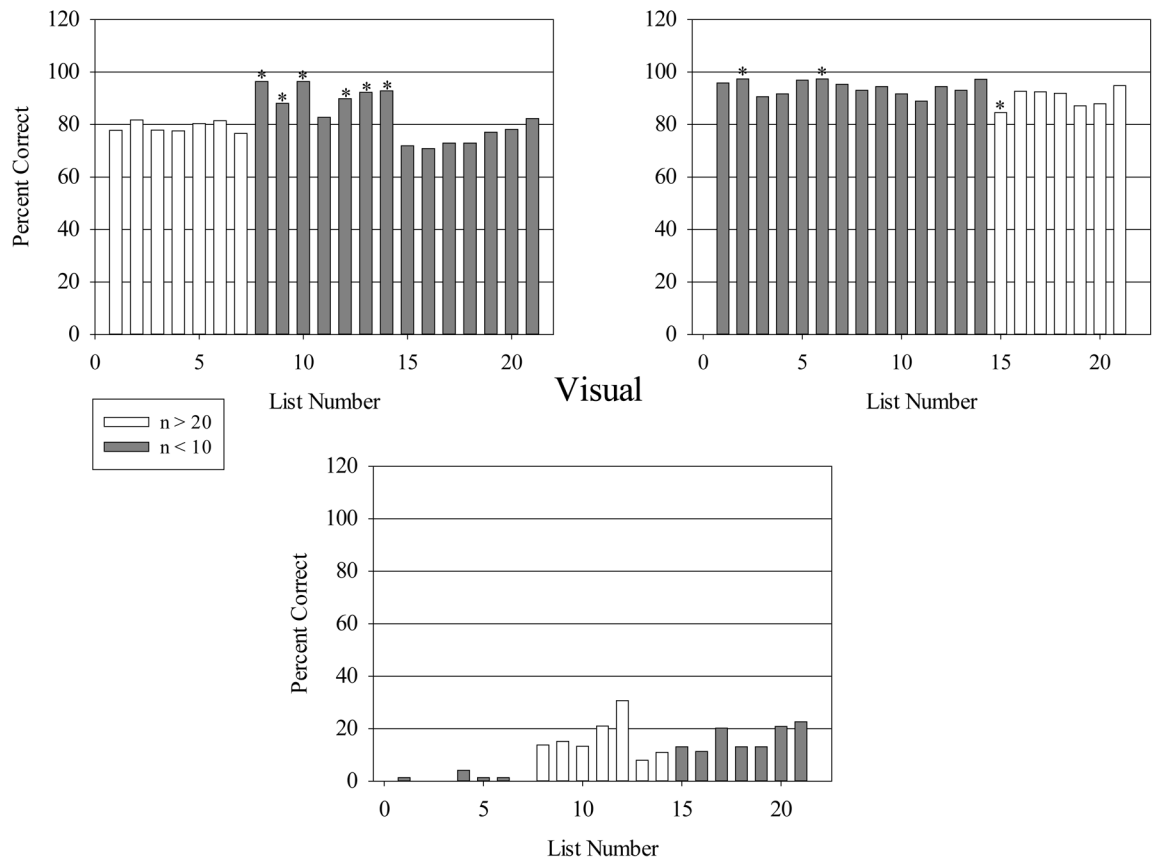


Figure 3. Mean percent of key words correctly identified by children who are deaf or hard of hearing as a function of MLST-C list number and presentation format.

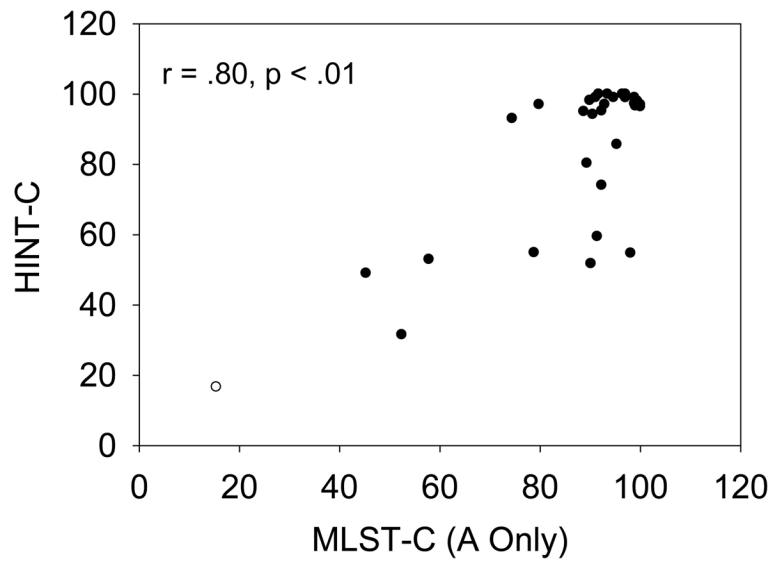


Figure 4. The relationship between the performance of children who are deaf or hard of hearing on the MLST-C in the auditory presentation format and performance on the HINT-C.

Table 1

Example sentences from each lexical category with key words in italics

Lexical Category	Example Sentence
High Frequency – Sparse (Easy)	I <i>wonder</i> who <i>brought</i> the <i>food</i> .
High Frequency – Dense	We <i>made</i> the <i>boats white</i> for Joe.
Low Frequency – Sparse	<i>Adam</i> put the <i>banjo</i> in the <i>garage</i> .
Low Frequency – Dense (Hard)	<i>Start walking</i> to your <i>seat</i> .

Table 2

Participant characteristics

	CI Users (N=11)	HA Users (N=17)	CI + HA Users (N=3)
Mean Age at Test	10.10 yrs	9.05 yrs	11.44 yrs
Mean Age at Onset	0.28 mos	1.25 mos	0 mos
Mean Age Fit with Current Sensory Aid	3.20 yrs	3.18 yrs	3.54 yrs
Hearing Loss Classification	Profound	Mild - 5 Moderate - 9 Severe - 2	CI: Profound HA: Severe - 2 Profound - 1
Mean Receptive Vocabulary Age	9.66 yrs	9.75 yrs	21.93 yrs (n=1)
Gender	8 males 3 females	8 males 9 females	2 males 1 female

Table 3

Correlations between performance on the MLST-C and the HINT-C

MLST-C Presentation Format	Correlation Coefficient
Visual-Only	$r = .085^*$
Auditory-Only	$r = .093^*$
Auditory + Visual	$r = .097^*$

 $p < .01$