J Am Acad Audiol 10: 113–123 (1999)

Lexical Neighborhood Test: Test–Retest Reliability and Interlist Equivalency

Karen Iler Kirk*
Laurie S. Eisenberg†
Amy S. Martinez†
Marcia Hay-McCutcheon*

Abstract

This study examined the reliability and interlist equivalency of two new recorded spoken word recognition measures, the Lexical Neighborhood Test (LNT) and the Multisyllabic Lexical Neighborhood Test (MLNT), and evaluated the effects of lexical difficulty on spoken word recognition by children with hearing loss. Participants were 16 children with prelingual, profound deafness who used a cochlear implant. Test–retest reliability was high and no significant interlist differences were observed for both measures. In addition, we found that lexically "easy" words (i.e., those that occur often and have few phonemically words with which they are similar) were recognized correctly more often than lexically "hard" words (i.e., those with opposite characteristics). The results demonstrate that the LNT and MLNT provide reliable information about the spoken word recognition abilities of children with profound hearing loss who use cochlear implants. In addition, these new measures also provide more detailed information about the way in which these children organize and access spoken words from long-term lexical memory.

Key Words: Children, cochlear implants, hearing loss, spoken word recognition

Abbreviations: CHILDES = Child Language Data Exchange System; LNT = Lexical Neighborhood Test; MLNT = Multisyllabic Lexical Neighborhood Test; NAM = Neighborhood Activation Model; OC = oral communication; PB-K = Phonetically Balanced Kindergarten Word Lists; PTA = average of the pure-tone threshold hearing levels at 500, 1000, and 2000 Hz; SPL = speech presentation level; TC = total communication

Historically, several criteria have been considered essential in the selection of test items for measures of spoken word recognition by listeners with hearing impairment, including word familiarity, homogeneity of audibility, and phonetic balancing (Hudgins et al, 1947; Hirsh et al, 1952; Bilger, 1984). Phonetic balancing was included because it was assumed that all speech sounds must be included in order to test hearing. Subsequent research demonstrated that phonetic balancing was not needed to achieve equivalent word lists (Tobias, 1964; Carhart, 1965; Hood and Poole, 1980) and that other nonauditory factors, such as subject age or language level, also influence spoken word recognition (Smith and Hodgson, 1970; Hodgson, 1985; Jerger, 1984). Nonetheless, the Phonetically Balanced Kindergarten word lists (PB-K), which were developed by Haskins in 1949, remain one of the most commonly used open-set measures to assess spoken word recognition in children with hearing loss who use cochlear implants. Recorded versions of the PB-K are available commercially, allowing comparison of results obtained at different clinical and research centers. The PB-K also has been used to evaluate potential cochlear implant candidates as well as to measure postimplant performance. It still remains today one of the most important outcome measures for assessing success with a cochlear implant.

*Department of Otolaryngology—HNS, Indiana University School of Medicine, Indianapolis, Indiana; †Children's Auditory Research and Evaluation Center, House Ear Institute, Los Angeles, California.

Reprint requests: Karen Iler Kirk, Department of Otolaryngology—HNS, Indiana University School of Medicine, 702 Barnhill Drive, Indianapolis, IN 46202

Early studies found that children with cochlear implants generally performed relatively poorly on the PB-K (Staller et al, 1991; Fryauf-Bertschy et al, 1992, Miyamoto et al, 1993). For example, Staller et al (1991) reported that the mean PB-K score for 28 children with approximately 2 years of cochlear implant experience was only 11 percent words correct. More recently, we have seen a wider range of performance on the PB-K as children have been implanted at younger ages and as cochlear implant signal processing strategies have improved (Shea et al, 1994; Fyrauf-Bertschy et al, 1997; Waltzman et al, 1997; Sehgal et al, in press). However, overall performance on the PB-K remains relatively low. For example, Fryauf-Bertschy et al (1997) reported PB-K results for 34 prelingually deafened children who had used their Nucleus 22-channel cochlear implant for 24 months. Although one child did achieve a score of 35 percent words correct, the majority of children correctly identified less than 20 percent of the words.

One possible reason why children with cochlear implants perform poorly on the PB-K is that when test item selection is constrained by phonetic balancing, the resulting lists may contain words that are unfamiliar to young deaf children who typically have very limited vocabularies. In fact, Fryauf-Bertschy et al (1997) found it necessary to use a reduced set of PB-K test items with younger children who could not demonstrate knowledge of the vocabulary on the full set of 50 items.[1] Children should be able to repeat unfamiliar words if their sensory aid provides adequate auditory information for complete phoneme identification. If not, then children will most likely select a phonemically similar word within their current active vocabulary.

Development of New Open-Set Word Recognition Tests

Because of these concerns regarding the PB-K, two new measures, the Lexical Neighborhood Test (LNT) and the Multisyllabic Lexical Neighborhood Test (MLNT) (Kirk et al, 1995), were developed at Indiana University School of Medicine to assess the open-set speech perception performance of children with hearing loss who used a cochlear implant or other sensory aid. These measures were developed according to two criteria. First, the specific words for these lists were selected to be familiar to young children with relatively limited vocabularies. Test items on the LNT and MLNT were selected from the Child Language Data Exchange System or CHILDES database (MacWhinney and Snow, 1985), which contains data from published studies concerning normal language development. This database consists of transcripts of verbal exchanges between a child or children and a caregiver or another child in the environment. All tokens on the LNT and MLNT were selected from productions by children between the ages of 3 and 5 years, and therefore are assumed to represent early-acquired vocabulary.

The second criterion was that the new measures should be based on what is known about word recognition and lexical access using a current model of spoken word recognition. This latter criterion was included so that test results would not only provide descriptive information about the benefits of using a sensory aid, but would also yield additional information about the way in which children with hearing loss organize and access spoken words from their long-term memory. The LNT and MLNT are theoretically motivated tests and are based on the assumptions of the Neighborhood Activation Model (NAM) (Luce, 1986; Luce and Pisoni, 1998). The NAM proposes that words are organized into "similarity neighborhoods" based on their frequency of occurrence (i.e., how often words occur in the language) and the density (i.e., acoustic-phonetic similarity) of words within the lexical neighborhood. One measure of lexical density is the number of words, or lexical neighbors, that can be generated from a target word by adding, deleting, or substituting one phoneme at a time (Greenberg and Jenkins, 1964; Landauer and Streeter, 1973). Words with many similar lexical neighbors come from "dense" neighborhoods, whereas those with few neighbors come from "sparse" neighborhoods. Figure 1 illustrates two lexical neighborhoods, one sparse and one dense. In adult listeners with normal hearing, higher frequency words that occur often and come from sparse neighborhoods are lexically "easy," that is, they are recognized with greater accuracy than lexically "hard" words (those that occur infrequently and come from "dense" neighborhoods) (Luce, 1986; Luce et al, 1990; Luce and Pisoni, 1998).

The NAM provides a two-stage account of how the structure and organization of sound

[1]As pointed out by an anonymous reviewer, the use of a reduced set of PB-K items jeopardizes the phonetic balancing of the stimulus items.

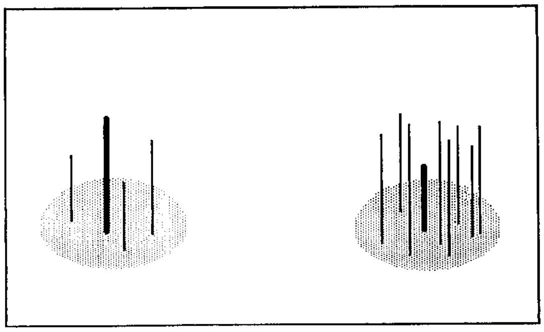

Figure 1 Two examples of a lexical neighborhood. The thick bar represents the target word, the narrow bars represent phonetically similar words, or neighbors, and the height of the bars represents word frequency. Lexically easy target words are high in word frequency and occur in sparse neighborhoods; lexically hard target words have the opposite characteristics.

patterns in words contributes to the perception of spoken words (Luce and Pisoni, 1998). According to NAM, a stimulus input activates a set of similar acoustic-phonetic representations in memory. These acoustic-phonetic representations are assumed to be activated in a multidimensional acoustic-phonetic space with activation levels proportional to the degree of similarity to the stimulus word. Over the course of perceptual processing, the stimulus pattern corresponding to the input receives successively higher levels of activation, while the activation levels of acoustically similar patterns become attenuated. This initial stage of activation is followed by a process of "lexical selection" among a large number of potential candidates that are consistent with the acoustic-phonetic input. Word frequency is assumed to act as a biasing factor by multiplicatively adjusting the activation levels of the acoustic-phonetic representations. In lexical selection, the activation levels are then summed and the probabilities of choosing each acoustic-phonetic representation are computed based on the overall activation level. According to this model, word recognition occurs when a given acoustic-phonetic representation is chosen based on the computed probabilities (for further details, see Luce and Pisoni, 1998).

Development of the LNT and MLNT Word Lists

The LNT consists of two 50-item lists of monosyllabic words and the MLNT consists of two 24-item lists of words, each with two or three syllables. Within each list of the LNT and MLNT, half of the items are lexically easy and half are lexically hard. Lexical difficulty was determined using Logan's (1992) analysis of the lexical properties of words in the CHILDES database. Among the monosyllabic words in the corpus, word frequency ranged from 1 to 519 occurrences with a median of 4 occurrences; monosyllabic lexical neighborhood density ranged from 0 to 19 neighbors with a median of 4 neighbors per target word. Lexically easy words on the LNT were above the median for word frequency and below the median for lexical density. The reverse pattern was used to select the lexically easy words. For the multisyllabic words in the CHILDES database, word frequency ranged from 1 to 100 occurrences with a median of 2 occurrences and lexical density ranged from two to seven words with a median of zero neighbors. On the MLNT, lexically easy words were those that had word frequencies greater than two and neighborhood densities of zero; lexically hard MLNT words were those with word frequencies of less than two occurrences and neighborhood densities greater than 0 neighbors.

Comparison of Performance on the PB-K, LNT, and MLNT

Kirk et al (1995) examined performance on the PB-K, LNT, and MLNT in a group of 28 pediatric cochlear implant users. All tests were administered via live voice. The results demonstrated that word recognition performance was significantly higher for the LNT word lists than for the PB-K, suggesting that the PB-K underestimates spoken word recognition in these children. Performance also was significantly better for the multisyllabic stimuli on the MLNT than for the monosyllabic stimuli on the LNT, which indicates that these children were able to use linguistic redundancy and context (i.e., word length) to aid them in spoken word recognition. Finally, the results revealed that children with cochlear implants identified lexically easy words with greater accuracy than lexically hard words. This latter finding suggests that children with cochlear implants are sensitive to the acoustic-phonetic similarities among spoken words, that they organize spoken words into similarity neighborhoods in long-term memory, and that they use this structural information and context in recognizing isolated words.

The present paper reports on the development of recorded versions of the LNT and MLNT word lists. There were two purposes for the present investigation. The first was to establish test–retest reliability and interlist equivalency of these recorded materials. The second purpose was to determine whether the effects of lexical difficulty on spoken word recognition previously seen with live voice stimulus presentation could be replicated using recorded versions of the LNT and MLNT.

METHODS

Participants

Participants were recruited from among the population of children with the Nucleus 22-channel cochlear implant followed at either Indiana University School of Medicine or the House Ear Institute. Selection criteria included (1) ≥ 5 years of age and (2) ≥ 6 months of cochlear implant experience. Sixteen children participated in this study. Table 1 presents the demographic characteristics of the children who were included in this study. Five of the children had a congenital hearing loss; for the remaining participants, age at onset ranged from 0.75 years to 6.0 years of age. One child developed a profound hearing loss as a result of meningitis; for the remaining 15 children, the etiology was unknown. Thirteen of these children used oral communication (OC) and the remaining three children used total communication (TC) (i.e., the simultaneous use of spoken and signed English). All participants used the Spectra processor with the SPEAK strategy. The number of active electrodes ranged from 13 to 20 across children. The children were paid $10.00 per hour for their participation.

Stimuli

The LNT and MLNT word lists were used to assess open-set spoken word recognition in this investigation. To create recorded lists, the 100 LNT and the 48 MLNT stimulus words were recorded by a male talker with a General American dialect. During recording, the talker was seated inside a sound-attenuated, single-walled booth (Industrial Acoustics Company Audiometric Testing Room, Model 402). He was asked to produce each word in a conversational manner as it appeared on a computer screen in front of him. The talker's productions were transduced using a high-quality Shure microphone (Model SM98). The recordings were digitally sampled on line at a rate of 22.05 kHz with 16-bit amplitude quantization using a Tucker-Davis Technologies System II with an A/D converter (DD1) and low-pass filter of 10.4 kHz (anti-aliasing filter, FT5). The silence preceding and following each target word was removed using a commercial waveform editor (CSpeechSP, version 4.1 Beta). Each word was then stored as an individual digital sound file. Next, the RMS amplitude of the words was equated across all tokens within a test (i.e., the LNT or the MLNT) using a signal-processing software package (Luce and Carrell, 1981).

Table 1 Participant Characteristics

Participants	*Etiology*	*Age at Profound Loss (Yr)*	*Unaided PTA* (dB HL)*	*Age CI Fit (Yr)*	*Processor*	*Strategy*	*Number of Active Electrodes*	*Age at Testing (Yr)*
1	Unknown	Congenital	112	2.6	Spectra	SPEAK	20	5.6
2	Unknown	6.0	93	6.3	Spectra	SPEAK	20	7.4
3	Unknown	Congenital	110	4.7	Spectra	SPEAK	20	6.8
4	Unknown	Congenital	118	2.6	Spectra	SPEAK	19	6.0
5	Unknown	Congenital	118	4.9	Spectra	SPEAK	17	14.6
6	Unknown	1.6	101	7.2	Spectra	SPEAK	20	8.2
7	Unknown	0.8	120	4.1	Spectra	SPEAK	18	8.3
8	Unknown	Congenital	107	3.6	Spectra	SPEAK	20	5.6
9	Unknown	0.4	107	3.9	Spectra	SPEAK	19	5.8
10	Unknown	1.8	107	5.8	Spectra	SPEAK	20	9.0
11	Unknown	0.9	103	3.4	Spectra	SPEAK	20	6.8
12	Meningitis	2.1	118	5.4	Spectra	SPEAK	14	9.4
13	Unknown	1.1	111	2.1	Spectra	SPEAK	20	6.5
14	Unknown	0.7	92	5.5	Spectra	SPEAK	20	7.1
15	Unknown	0.8	120	2.3	Spectra	SPEAK	20	4.8
16	Unknown	1.3	92	4.3	Spectra	SPEAK	20	5.5

*Unaided pure-tone average calculated by substituting 120 dB HL for no response at a given frequency.

The intelligibility of the target words was determined by a group of 10 college students with normal hearing. The students were presented with the target words at 70 dB SPL in quiet under headphones and asked to write down each word they heard. The mean percent of words correctly identified by these listeners was 95 percent for the LNT words lists and 99 percent for the MLNT word lists.

Procedures

Word recognition testing was carried out at two testing sites (Indiana University School of Medicine and the House Ear Institute) by one of several audiologists or speech-language pathologists experienced in testing children with a hearing loss. Children were tested twice; the time between the two test sessions ranged from 3 hours to 15 days. All children were tested in a quiet room by an examiner who sat to one side of the child. Stimuli were presented at 70 dB SPL via a loudspeaker placed at 0° azimuth from the child. Each subject was administered the two 50-word lists of the LNT and the two 24-word lists of the MLNT. Test order and list order within each test were the same at both sessions for a given child, but were counterbalanced across children. Stimulus presentation and response collection were under computer control. Children responded by repeating the word they heard and their responses were transcribed immediately by the examiner. Once a child had responded, the experimenter presented the next stimulus item. Responses were scored as the percent of words and phonemes correctly identified. Children who used TC (simultaneous spoken and signed English) were asked to speak and sign their responses; children who used OC were asked to repeat the word and to write or explain their answer (e.g., "What is that?" or "What do you do with that?") if their response was unintelligible. If a child was not able to clarify an unintelligible response, that response was phonetically transcribed by the examiner and scored for phonemes correct, but no word credit was given.

Data Analysis

The children's mean word and phoneme scores on the LNT and MLNT were arcsine transformed to stabilize the variance. Four factorial repeated measures analyses of variance (ANOVA) were then computed to assess statistically the effects of lexical difficulty, list, and session. In addition, Pearson product-moment correlations between individual participant's scores at sessions 1 and 2 were computed for both word and phoneme scores to determine test–retest reliability.

RESULTS

Table 2 presents the correlations between word scores obtained during sessions 1 and 2 for the LNT and MLNT measures. Table 3 presents the same correlational data for the phoneme scores obtained at sessions 1 and 2. The correlations were computed separately for each list of the LNT and MLNT. For both the LNT and MLNT, the correlations between scores obtained at the two sessions were very high regardless of list number administered or the lexical difficulty of the items ($r \geq .83$). This was true for both word and phoneme recognition. Furthermore, all correlations were highly significant ($p < .0001$). These correlational results demonstrate that test–retest reliability is quite high for the LNT and MLNT.

Mean results for the experimental test conditions are displayed in Figure 2 (LNT word and phoneme scores) and Figure 3 (MLNT word and phoneme scores). Error bars represent ±2 standard errors of the mean. The data were further stratified according to lexical difficulty (easy vs hard), test list (1 vs 2), and test session (first vs second).

LNT Scores

Figure 2 illustrates the main effect of lexical difficulty on LNT word recognition performance. Mean word scores were significantly higher for the lexically easy words than for the

Table 2 Correlations between Percentage of Words Correctly Identified on the LNT and MLNT at Sessions 1 and 2 as a Function of List Number and Lexical Difficulty.

Test	*List and Lexical Difficulty*	*Correlation between Sessions 1 and 2 (r)*
LNT	List 1 easy words	.89
	List 1 hard words	.85
	List 2 easy words	.91
	List 2 hard words	.87
MLNT	List 1 easy words	.84
	List 1 hard words	.83
	List 2 easy words	.92
	List 2 hard words	.90

All correlations are significant ($p < .0001$).

Table 3 Correlations between Percentage of Phonemes Correctly Identified on the LNT and MLNT at Sessions 1 and 2 as a Function of List Number and Lexical Difficulty.

Test	*List and Lexical Difficulty*	*Correlation between Sessions 1 and 2 (r)*
LNT	List 1 easy phonemes	.92
	List 1 hard phonemes	.91
	List 2 easy phonemes	.94
	List 2 hard phonemes	.90
MLNT	List 1 easy phonemes	.88
	List 1 hard phonemes	.90
	List 2 easy phonemes	.95
	List 2 hard phonemes	.93

All correlations are significant ($p < .0001$).

hard words (F [1, 15] = 21.09, $p < .001$). Across lists and sessions, the mean percent of words correctly identified was 48 percent for the lexically easy words and 42 percent for the lexically hard words. There were no significant differences in word recognition performance as a function of list or test session, nor were any of the interactions significant.

In contrast to the word scores, analysis of the LNT phoneme scores revealed no statistically significant differences between easy and hard words. Although differences between test lists were not significant, LNT phoneme recognition scores did differ significantly between sessions 1 and 2 (F [1, 15] = 7.69, $p < .02$). Session 2 mean phoneme scores always exceeded session 1 mean scores regardless of test list or lexical difficulty of the test items. The mean percent of phonemes correctly identified was 59 percent for session 1 and 63 percent for session 2. The increased scores seen with repeated administration indicate practice effects for phoneme recognition. That is, the children appeared to learn something about phoneme recognition with repeated administrations of these tests in a short period of time. Finally, there were no significant interactions.

MLNT Scores

Figure 3 illustrates the percent of words and phonemes correctly identified on the MLNT. On average, the participants correctly identified 60 percent of the lexically easy words and 45 percent of the lexically hard words on the MLNT. The main effect of lexical difficulty was significant (F [1, 15] = 23.94, $p < .001$). There were no main effects for test lists and sessions but there was a significant interaction between lexical difficulty and test list (F [1, 15] = 10.33, $p = .006$). Figure 3 illustrates that the difference in performance between lexically easy and hard MLNT stimulus words was greater for list 2 than for list 1.

Analysis of MLNT phoneme scores revealed a significant main effect of lexical difficulty. On average, 67 percent of the phonemes in lexically easy words were identified correctly as compared to only 60 percent of the phonemes in the lexically hard words. There were no main effects for list or test session; moreover, no interactions were found.

DISCUSSION

Summary of Results

For both the LNT and MLNT, significant effects were observed for lexical difficulty on word recognition. Lexically easy words were recognized with greater accuracy than lexically

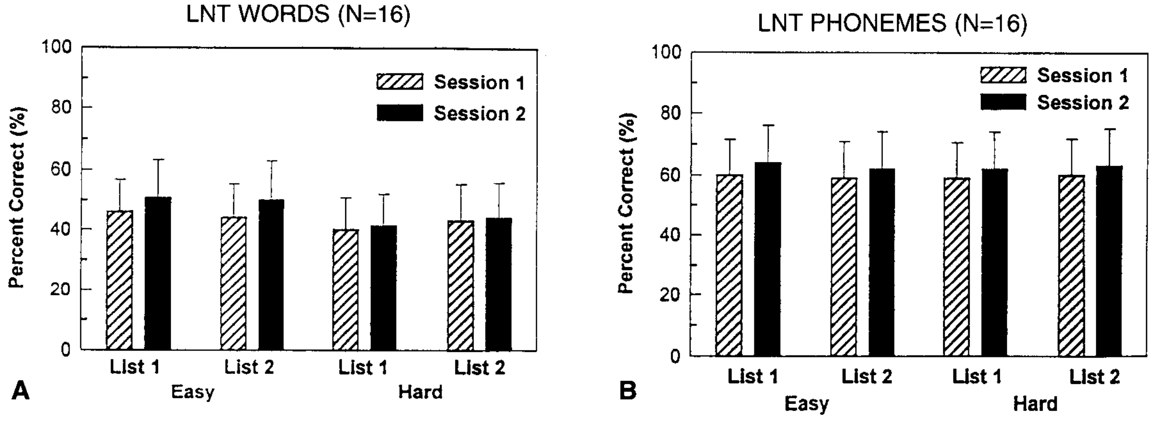

Figure 2 The percent of words and phonemes correctly identified on the LNT as a function of lexical difficulty, list number, and test session.

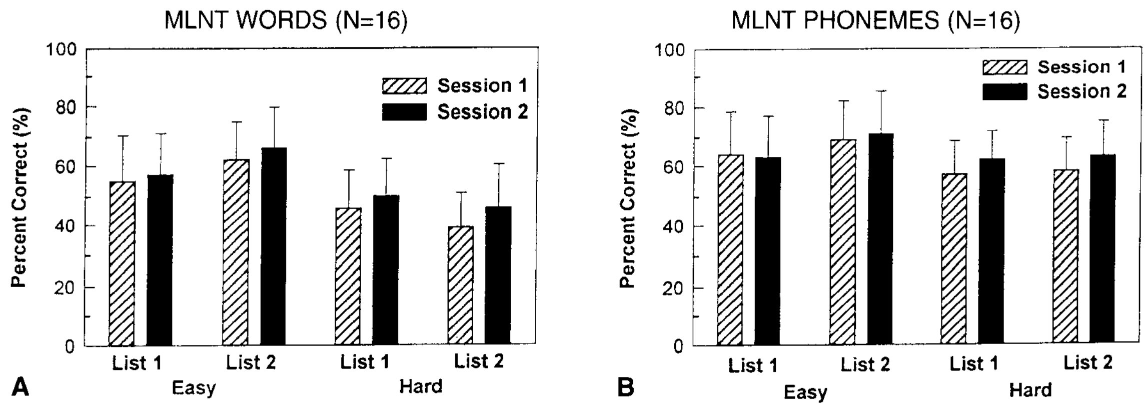

Figure 3 The percent of words and phonemes correctly identified on the MLNT as a function of lexical difficulty, list number, and test session.

hard words. These findings, obtained with recorded versions of the LNT and MLNT word lists, replicate the earlier results of Kirk et al (1995) using live-voice stimulus presentation. Interestingly, the children in the current study had higher word and phoneme recognition scores on the LNT and MLNT than did children in the previous study, who were tested via live voice. The improved performance noted here might relate to the processor strategy employed by the participants. The majority of children tested in the previous study used an earlier generation of speech processing strategy, the Multipeak or MPEAK strategy, whereas the children in the present study used the SPEAK strategy. The newer generation strategy, SPEAK, has been shown to yield higher word recognition scores in both adults and children with cochlear implants (Skinner et al, 1994; Sehgal et al, in press). Despite the difference in overall performance levels, the present study replicates the earlier results concerning lexical effects on spoken word recognition by children with cochlear implants. The fact that lexical effects were noted with the newly recorded LNT and MLNT stimuli suggests that these effects are robust across presentation formats, individual talkers, and specific stimulus tokens. Furthermore, the present results support the notion that children with cochlear implants are able to perceive the acoustic-phonetic similarities among words and that they organize words into similarity neighborhoods in long-term memory.

With respect to phoneme scores, there was a discrepancy between the LNT and the MLNT measures. For the LNT, phoneme recognition did not differ between the lexically easy and hard words. However, for the MLNT, phoneme recognition was significantly higher for lexically easy than for lexically hard words. In the previous investigation of Kirk et al (1995), no lexical effects on phoneme recognition were evident for either the monosyllabic stimuli on the LNT or the multisyllabic words on the MLNT. These previous findings suggested that children with cochlear implants recognize words in the context of other words rather than on a phoneme-by-phoneme basis. In the present investigation, the results for the LNT are consistent with previous research, but the results for the MLNT are not. There are several possible explanations for this discrepancy. First, in the previous investigation, the MLNT consisted of one 30-item list, with half of the items lexically easy and half lexically hard. When the recorded measures were developed, we added additional words selected from the CHILDES database to create two MLNT lists of 24 items. A second MLNT list was created so that children would not be administered the same list at every postimplant testing interval. It may be that the change in stimuli yielded different results. Second, it appears that the children in the present study were responding to the multisyllabic stimuli on the MLNT in an all-or-none fashion. Recall that multisyllabic words come from sparser neighborhoods than monosyllabic words and therefore have fewer lexical neighbors than monosyllabic words (Pisoni et al, 1985). For example, on the MLNT, the lexically easy words had no lexical neighbors and the lexically hard words had an average of about two lexical neighbors per target word. Thus, there are few phonemically similar words with which the MLNT target words can be confused. It is likely that the children either identified a given multisyllabic word correctly, thereby getting all of the phonemes correct (which happened significantly more often for lexically easy words), or their error response was not close to the target and they missed all of the phonemes.

That is, they were unlikely to select a lexical neighbor as an error response for the target. Although these results differ from the previous investigation, they still suggest that children do not identify words primarily on a phoneme-by-phoneme basis. Rather, spoken word recognition is influenced by competition from other phonetically similar words in the lexical neighborhood and even phoneme recognition is affected by lexical context.

A consistent finding among the measures obtained from this study was the equivalence among the two lists within each test. Word and phoneme recognition on the LNT and MLNT did not differ as a function of the list administered. This suggests that alternate lists can be used at different test intervals to minimize learning effects without significantly influencing performance. There was a statistically significant interaction between list and lexical difficulty for one of the four ANOVAs computed on the data. Specifically, on MLNT word scores, greater differences were evident between lexically easy and hard words on list 2 than on list 1. Figure 3 illustrates that scores for the lexically easy words on the MLNT were slightly higher for list 2 than for list 1, whereas scores for the lexically hard words were lower for list 2 than for list 1. A review of the lexical statistics for the two lists of the MLNT revealed that the lexical characteristics of the easy words on lists 1 and 2 were very similar. The mean word frequency was seven and six occurrences in the corpus for the easy words on MLNT lists 1 and 2, respectively; the average lexical density and neighborhood frequency (the mean word frequency of all of a target word's lexical neighbors) of the lexically easy words was zero for both lists. However, differences in the lexical properties of the lexically hard words were noted between MLNT lists 1 and 2. The lexically hard words on lists 1 and 2 were similar in mean word frequency (1.9 and 1.3 occurrences, respectively) and in mean lexical density (1.8 and 1.5 neighbors, respectively), but they differed in terms of mean neighborhood frequency (3.1 for list 1 vs 8.8 for list 2). Recall that mean neighborhood frequency refers to how often the lexical neighbors occur in the language. The higher the value, the more competition these lexical neighbors generate for lexical selection. The differences in mean neighborhood frequency between MLNT lists 1 and 2 suggest that it is more difficult to identify lexically hard words on list 2 than on list 1, and this may have contributed to the interaction evident for these data.

For three of the four measures (LNT words, MLNT words, and phonemes), performance did not differ between test sessions. Furthermore, the strong correlations found between sessions ($r \geq .83$) provide evidence demonstrating that these measures are highly reliable from one administration to another. For one dependent measure, LNT phoneme scores, word recognition was significantly higher for session 2 than for session 1, suggesting that learning occurred over repeated administrations within a relatively brief time period. If it is necessary to test children several times within a brief interval of time, it would be best to administer alternate lists in order to minimize learning effects that may be independent of the child's spoken word recognition skills. Alternatively, the clinician may prefer to use only word scores where such learning was not evident.

Clinical Utility of the LNT and MLNT

The open-set spoken word recognition skills of the children in the present study spanned a wide range of abilities. Individual scores (averaged across lists and levels of lexical difficulty) ranged from 4 percent to 82 percent on the LNT and from 6 percent to 84 percent on the MLNT; the median scores were 47 percent for the LNT and 56 percent for the MLNT. In this regard, children selected to participate in this study were representative of the population of children with cochlear implants. Thus, the present findings suggest that the LNT and MLNT should yield reliable estimates of spoken word recognition in children with cochlear implants regardless of their overall performance levels. This is an important consideration in measuring progress over time with a cochlear implant or other sensory aid. Another important aspect of these new measures is that they provide more than descriptive information about speech perception abilities. The use of theoretically motivated measures of spoken word recognition may help us learn more about the underlying factors that influence spoken word recognition in children with hearing impairment; their use may also provide important diagnostic information about the way these children process spoken language. For example, if children do not show the expected lexical effects on spoken word recognition, this would suggest that they may be organizing and accessing words from memory in a different way than listeners with normal hearing. Furthermore, the magnitude of lexical effects on spoken word recognition may yield important information about how well children can make fine acoustic-phonetic distinctions among phonetically similar words. Relatively gross acoustic cues available from coarse coding of the speech signal may be sufficient to identify many lexically easy words in sparse neighborhoods, but finer acoustic-phonetic encoding is required to access lexically hard words from dense neighborhoods. Thus, a large decrease in the identification of lexically hard words compared to easy words might suggest that a child is unable to encode the fine acoustic-phonetic cues available in the speech signal.

The validity of the LNT and MLNT was not specifically addressed in the present investigation. However, several recent studies have shown that performance on the LNT and MLNT is significantly correlated with other measures of spoken word recognition and spoken language processing in children with cochlear implants. Pisoni et al (1997) examined the relationship between various measures of spoken language processing in a group of pediatric cochlear implants users with exceptionally good speech perception abilities. Children who scored in the 80th percentile or higher on the PB-K at 2 years postimplant were included in this investigation. Intercorrelations were computed among a number of measures 1 year postimplant to explore whether a common factor or factors underlying the development of spoken language abilities could be identified. The results revealed that spoken word recognition performance on the LNT and the MLNT was significantly correlated with open-set sentence recognition in the auditory-only modality as measured by the Common Phrases test (Robbins et al, 1995) ($r > .80$), with receptive vocabulary recognition as measured by the Peabody Picture Vocabulary Test (Dunn, 1965) ($r > .61$), and with both expressive and receptive language abilities as measured by the Reynell Language Development Scales (Reynell and Huntley, 1985) ($r > .65$). Performance on the LNT and MLNT was also significantly correlated with the children's speech intelligibility ($r > .71$). In a separate study, Kirk (1996) found that word recognition performance on the LNT was significantly correlated with receptive language abilities on the Reynell Language Developmental Scales for children who used OC but not for those who used TC. Finally, Pisoni and Geers (1998) found significant correlations between the spoken word recognition of children with cochlear implants, as measured by the LNT, and their ability to recall lists of digits presented through listening alone ($r > .58$).

The investigations cited above suggest that the LNT and MLNT are measuring something about the encoding, storage, retrieval, and manipulation of spoken language. The strong correlations between measures of lexical access and other language processing measures suggest that the development of the lexicon serves as the interface between the initial sensory input and the representation of sound patterns of words in lexical memory (Pisoni et al, 1997). Furthermore, the data of Pisoni and Geers (1998) demonstrate the role of working memory and rehearsal in lexical access. According to Baddeley et al (1998), one aspect of working memory, the phonological loop, serves to store unfamiliar sound patterns while more permanent memories are constructed. They suggest that the primary role of this phonological loop is in the acquisition of new vocabulary. Children must be able to retain, process, and encode unfamiliar sound patterns before they become part of their lexicon. Thus, the ability to map sounds onto meaning is a necessary prerequisite to building a grammar from the ambient spoken language in a child's environment (Pisoni et al, 1997). As such, it appears that measures of spoken word recognition such as the LNT and MLNT may be one important predictor of a child's ability to acquire spoken language.

Acknowledgment. This research was funded in part by the National Institute on Deafness and Other Communication Disorders grants K08 DC00126-02, K08 DC00083-05, R01 DC0064-08, R01 DC00111-22, and T32 DC00012-20. We thank David Ertmer, Julie Schulte, and Ann Kalberer for their help in preparing the recorded stimuli. We also thank Susan Sehgal, Jennifer Boyett, and Annmarie Henson for their assistance with data collection, and Terri Kerr for help with data analysis. David B. Pisoni provided valuable comments on an earlier version of this manuscript.

REFERENCES

Baddeley A, Gathercole S, Papagno C. (1998). The phonological loop as a language learning device. *Psychol Rev* 105:158–173.

Bilger RC. (1984). Speech recognition test development. *ASHA Reports* 14:2–7.

Carhart R. (1965). Problems in the measurement of speech discrimination. *Arch Otolaryngol* 82:253–260.

Dunn LM. (1965). *Peabody Picture Vocabulary Test.* Circle Pines, MN: American Guidance Service.

Fryauf-Bertschy H, Tyler RS, Kelsay DM, Gantz BJ. (1992). Performance over time of congenitally deaf and postlingually deafened children using a multichannel cochlear implant. *J Speech Hear Res* 35:892–902.

Fryauf-Berstchy H, Tyler RS, Kelsay DMR, Gantz BJ, Woodworth GG. (1997). Cochlear implant use by prelingually deafened children: the influences of age at implant and length of device use. *J Speech Hear Res* 40:183–199.

Greensberg JH, Jenkins JJ. (1964). Studies in the psychological correlates of the sound system of American English. *Word* 20:157–177.

Haskins H. (1949). *A Phonetically Balanced Test of Speech Discrimination for Children.* Unpublished master's thesis, Northwestern University, Evanston, IL.

Hirsh IJ, Davis H, Silverman SR, Reynolds EG, Eldert E, Benson RW. (1952). Development of materials for speech audiometry. *J Speech Hear Disord* 17:321–337.

Hodgson W. (1985). Testing infants and young children. In: Katz J, ed. *Handbook of Clinical Audiology.* Baltimore: William & Wilkins, 642–663.

Hood JD, Poole JP. (1980). Influence of the speaker and other factors affecting speech intelligibility. *Audiology* 19:434–455.

Hudgins CV, Hawkins JE, Karlin JE, Stevens SS. (1947). The development of recorded auditory tests for measuring hearing loss for speech. *Laryngoscope* 57:57–89.

Jerger S. (1984). Speech audiometry. In: Jerger J, ed. *Pediatric Audiology.* San Diego: College Hill Press, 71–94.

Kirk KI. (1996). *Lexical Discrimination and Age at Implantation at Implantation: A First Report.* Paper presented at the Acoustical Society of America, Indianapolis, IN.

Kirk KI, Pisoni DB, Osberger MJ. (1995). Lexical effects on spoken word recognition by pediatric cochlear implant users. *Ear Hear* 16:470–481.

Landauer TK, Streeter LA. (1973). Structural differences between common and rare words: failure of equivalence assumptions for theories of word recognition. *J Verb Learn Verb Behav* 12:119–131.

Logan JS. (1992). *A Computational Analysis of Young Children's Lexicons.* Research on Speech Perception Technical Report No. 8. Bloomington, IN: Speech Research Laboratory, Indiana University.

Luce PA. (1986). *Neighborhoods of Words in the Mental Lexicon.* Research on Speech Perception Technical Report No. 6. Bloomington, IN: Speech Research Laboratory, Indiana University.

Luce PA, Carrell TD. (1981). *Creating and Editing Waveforms Using WAVES.* Research in Speech Perception Technical Report No. 7. Bloomington, IN: Speech Research Laboratory, Indiana University.

Luce PA, Pisoni DB. (1998). Recognizing spoken words: The Neighborhood Activation Model. *Ear Hear* 19:1–36.

Luce PA, Pisoni DB, Goldinger SD. (1990). Similarity neighborhoods of spoken words. In: Altmann GTM, ed. *Cognitive Models of Speech Processing: Psycholinguistic and Computational Perspectives.* Cambridge, MA: MIT Press.

MacWhinney B, Snow C. (1985). The child language data exchange system. *J Child Lang* 12:271–296.

Miyamoto RT, Osberger MJ, Robbins AM, Myres WA, Kessler K. (1993). Prelingually deafened children's performance with the Nucleus multichannel cochlear implant. *Am J Otol* 14:437–445.

Pisoni DB, Geers A. (1998). *Working Memory in Deaf Children with Cochlear Implants: Correlations between Digit Span and Measures of Spoken Language.* Paper presented at the Seventh Symposium on Cochlear Implants in Children, Iowa City, IA.

Pisoni DB, Nusbaum HC, Luce PA, Slowiaczek LM. (1985). Speech perception, word recognition and the structure of the lexicon. *Speech Communication* 4:75–95.

Pisoni DB, Svirsky MA, Kirk KI, Miyamoto RT. (1997). *Looking at the "Stars": A First Report on Intercorrelations among Measures of Speech Perception, Intelligibility, and Language in Pediatric Cochlear Implant Users.* Research on Spoken Language Processing Progress Report No. 21. Bloomington, IN: Speech Research Laboratory, Indiana University, 51–92.

Reynell JK, Huntley M. (1985). *Reynell Developmental Language Scales—Revised, Edition 2.* Windsor, England: NFER-Nelson Publishing.

Robbins AM, Renshaw JJ, Osberger MJ. (1995). *Common Phrases Test.* Indianapolis, IN: Indiana University School of Medicine.

Sehgal ST, Kirk KI, Svirsky M, Miyamoto RT. (1998). The effects of processor strategy on the speech perception performance of pediatric Nucleus multichannel cochlear implant users. *Ear Hear* 19:149–161.

Shea JJ III, Domico EH, Lupfer M. (1994). Speech perception after multichannel cochlear implantation in the pediatric patient. *Am J Otol* 15:66-70.

Skinner MW, Clark GW, Whitford LA, Seligman PM, Staller SJ, Shipp DB, Shallop JK, Everingham C, Menapace CM, Arndt PL, Antogenelli T, Brimacombe JA, Pijl S, Daniels P, George CR, McDermott HJ, Beiter AL. (1994). Evaluation of a new spectral peak coding strategy for the Nucleus 22 Channel Cochlear Implant System. *Am J Otol* 15:15–27.

Smith K, Hodgson W. (1970). The effects of systematic reinforcement on the speech discrimination responses of normal and hearing-impaired children. *J Auditory Res* 10:110–117.

Staller SJ, Beiter AL, Brimacombe JA, Mecklenberg DJ, Arndt P. (1991). Pediatric performance with the Nucleus 22-channel cochlear implant system. *Am J Otol* 12(Suppl):126–136.

Tobias JV. (1964). On phonemic analysis of speech discrimination tests. *J Speech Hear Res* 7:98–100.

Waltzman SB, Cohen NL, Gomolin RH, Green JE, Shapiro WH, Hoffman RA, Roland JT Jr. (1997). Long-term results of early cochlear implantation in congenitally and prelingually deafened children. *Am J Otol* 18:342–349.

APPENDIX

Stimulus Materials

Lexical Neighborhood Test (LNT) (Kirk et al, 1995)

List 1

Easy Words: juice, good, drive, time, hard, gray, foot, orange, count, brown, home, old, watch, need, food, dance, live (/lɨv/), stand, six, cold, push, stop, girl, hurt, cow

Hard Words: thumb, pie, wet, fight, toe, cut, pink, hi, song, fun, use, mine, ball, kick, tea, book, bone, work, dad, game, lost, cook, gum, cap, meat

List 2

Easy Words: down, truck, mouth, pig, give, school, boy, put, three, farm, fish, green, catch, break, house, sit, friend, jump, bird, swim, hold, want, snake, more, white

Hard Words: ear, hand, dry, zoo, goat, toy, call, sing, cut, wrong, bed, fat, man, run, hot, read (/rid/), grow, bag, cake, seat, nine, sun, bath, ten, ride

Multisyllabic Lexical Neighborhood Test (MLNT) (Kirk et al, 1995)

List 1

Easy Words: children, animal, monkey, finger, pocket, apple, morning, sugar, alright, about, because, crazy

Hard Words: butter, lion, money, jelly, yellow, purple, hello, carry, corner, heaven, measles, ocean

List 2

Easy Words: water, banana, glasses, airplane, window, tiger, cookie, again, another, almost, broken, china

Hard Words: puppy, pickle, button, summer, bottom, finish, bunny, belly, couple, under, naughty, really