Published in final edited form as:

Ear Hear. 1999 August ; 20(4): 363–371.

Some Computational Analyses of the PBK Test: Effects of Frequency and Lexical Density on Spoken Word Recognition

Ted A. Meyer and David B. Pisoni

Department of Otolaryngology (T.A.M., D.B.P.), DeVault Otologic Research Laboratory, Indiana University School of Medicine, Indianapolis, Indiana; and Department of Psychology (D.B.P.), Speech Research Laboratory, Indiana University, Bloomington, Indiana

Abstract

Objective—The Phonetically Balanced Kindergarten (PBK) Test (Haskins, Reference Note 2) has been used for almost 50 yr to assess spoken word recognition performance in children with hearing impairments. The test originally consisted of four lists of 50 words, but only three of the lists (lists 1, 3, and 4) were considered “equivalent” enough to be used clinically with children. Our goal was to determine if the lexical properties of the different PBK lists could explain any differences between the three “equivalent” lists and the fourth PBK list (List 2) that has not been used in clinical testing.

Design—Word frequency and lexical neighborhood frequency and density measures were obtained from a computerized database for all of the words on the four lists from the PBK Test as well as the words from a single PB-50 (Egan, 1948) word list.

Results—The words in the “easy” PBK list (List 2) were of higher frequency than the words in the three “equivalent” lists. Moreover, the lexical neighborhoods of the words on the “easy” list contained fewer phonetically similar words than the neighborhoods of the words on the other three “equivalent” lists.

Conclusions—It is important for researchers to consider word frequency and lexical neighborhood frequency and density when constructing word lists for testing speech perception. The results of this computational analysis of the PBK Test provide additional support for the proposal that spoken words are recognized “relationally” in the context of other phonetically similar words in the lexicon. Implications of using open-set word recognition tests with children with hearing impairments are discussed with regard to the specific vocabulary and information processing demands of the PBK Test.

The Phonetically Balanced Kindergarten (PBK) Test is an open-set test of spoken word recognition that has been widely used over the years in clinical audiology to measure speech perception skills in young children, especially children with profound hearing losses who use cochlear implants (CIs) or hearing aids (Carney, Osberger, Miyamoto, Karasek, Dettman, & Johnson, 1991; Fryauf-Bertschy, Tyler, Kelsay, Gantz, & Woodworth, 1997; Gantz, Tyler, Woodworth, Tye-Murray, & Fryauf-Bertschy, 1994; Kirk, Osberger, & Pisoni, 1995; Miyamoto, Osberger, Robbins, Myres, & Kessler, 1993; Osberger et al., 1991; Pisoni, Svirsky, Kirk, & Miyamoto, Reference Note 6; Staller, Dowell, Beiter, & Brimacombe, 1991; Waltzman, Cohen, Gomolin, Shapiro, Ozdamar, & Hoffman, 1994). The test was

Copyright: © 1999 Lippincott Williams & Wilkins, Inc.

Address for correspondence: Ted A. Meyer, M.D., Ph.D., Indiana University School of Medicine, Department of Otolaryngology, DeVault Otologic Research Laboratory, 702 Barnhill Drive, RR-044, Indianapolis, IN 46202.

Portions of the paper were presented at the 21st Annual Meeting of the Association for Research in Otolaryngology, St. Petersburg Beach, FL, February, 1998.

developed by Harriet Haskins in 1949 as part of her Master's thesis at Northwestern University to fill a need for an open-set test of speech perception that could be used for children of kindergarten age. Although Haskins' thesis was never published, three of her original four word lists are routinely used clinically.

Performance on the PBK Test is often used as the primary and "defining" criterion to identify exceptionally good users of CIs, the so-called "Star" performers (Pisoni et al., Reference Note 6). Children who display high levels of performance on the PBK Test also tend to perform well on other perceptual tests that are routinely administered as part of a standard assessment battery such as the Minimal Pairs Test (Robbins, Renshaw, Miyamoto, Osberger, & Pope, Reference Note 8), Common Phrases Test (Osberger et al., 1991), Lexical Neighborhood Test (LNT) (Kirk et al., 1995), Peabody Picture Vocabulary Test (Dunn & Dunn, 1981), and the Reynell Developmental Language Scales (Reynell & Gruber, Reference Note 7). In addition, and perhaps even more interesting, recent analyses have shown that intelligibility of produced speech, as measured by transcription scores of naive listeners, is significantly correlated to performance on the PBK Test for children who score in the top 20% of CI users on the PBK Test (Pisoni et al., Reference Note 6; Svirskey, 1996). This finding suggests that these particular children with profound hearing losses who use CIs have not only acquired the perceptual skills needed to recognize isolated words but have also developed the means to encode, represent and retrieve the sound patterns of the spoken words in memory.

Researchers and clinicians alike have routinely observed a great deal of variability among users of CIs on all of the standard assessment instruments (Carney et al., 1991; Fryauf-Bertschy et al., 1997; Gantz et al., 1994; Staller et al., 1991). In the case of the PBK Test, the range of performance for children covers almost the full scale from zero to approximately 90% correct word recognition in some cases (Carney et al., 1991; Fryauf-Bertschy et al., 1997; Gantz et al., 1994; Osberger et al., 1991; Waltzman et al., 1994). On closer inspection, the individual scores on the PBK Test tend to follow a bimodal distribution of performance: some of the children with CIs are able to perform well on the PBK open-set test, whereas other children routinely score zero on the test (Lane, 1995). Why is there so much variability among CI users on the PBK Test? Why is the PBK Test apparently so difficult for some deaf children with CIs but much more manageable for other children? These are several of the questions that only future research may be able to answer.

In examining the task demands of the PBK Test, two major factors come to mind that are worth considering in some greater detail. First, we consider the specific vocabulary items used on this test. Haskins developed the PBK lists from words taken from the well-known PB-50 lists for the adult listeners (Egan, 1948). The words chosen by Haskins were also among the 2500 words of highest frequency spoken by preschool children (The International Kindergarten Union, Reference Note 9). It has been suggested recently by Kirk et al. (1995) that children with limited vocabulary skills, such as very young children or children with CIs, may score poorly on or be unable to do the PBK Test at all because the specific words used on the test are simply too difficult and are not within the vocabulary of these children. Is this explanation of the difficulty of the PBK Test correct? How can this account be assessed?

At the present time, there are no vocabulary norms or lexical databases for children with profound hearing losses, so it is difficult to compare the words on the PBK Test with words from the vocabularies of deaf children. Kirk et al. (1995) attempted a computational analysis of the words on the PBK Test using a lexical database constructed by Logan (Reference Note 4). Logan's database contained samples of the vocabulary from normal-hearing children that was available in the CHILDES (Child Language Data Exchange System)

database (MacWhinney & Snow, 1985). Kirk et al., however, found that a large number (approximately 69%) of the words on the PBK Test were, in fact, not within Logan's database. Thus, it is possible that the children with CIs cannot recognize the words on the PBK Test simply because they are not familiar with the specific items used in this test.

In addition to the specific vocabulary items used on the PBK Test, the lexical properties of the words themselves may also play an important role in recognition. Some words may be hard to perceive because they are phonetically confusable with many other similar-sounding words (see Luce & Pisoni, 1998). And, some words may be hard to perceive because children with profound hearing losses have difficulty making fine phonetic discriminations that are needed for identification of these particular sound patterns. In their computational analyses of the vocabulary of the PBK Test, Kirk et al. (1995) found that a large number of words on the PBK Test were perceptually difficult; that is, the words came from regions of the lexicon where there were many other phonetically similar words that were higher in frequency than the target word. These observations suggest that discriminability and competition among lexical candidates that are phonetically similar to each other may also affect performance on the PBK test and may be another important factor that makes the words on this test difficult for children who use CIs (see Luce & Pisoni, 1998). A detailed explanation is provided in the Discussion. Thus, the PBK Test may be a very difficult test for children with profound hearing losses because many of the words are unfamiliar and therefore not part of their lexicon. Also, many of the words may be difficult to perceive in isolation, where, according to the Neighborhood Activation Model (NAM) (Luce, Reference Note 5, Luce & Pisoni, 1998; Luce, Pisoni & Goldinger, 1990), competition from other phonetically similar words in the child's language hinders recognition.

In addition to the specific vocabulary items used on the PBK Test, it is also possible that the task demands of an open-set word recognition test affect performance, particularly in young children with profound hearing losses who may not be able to make fine phonetic discriminations among different sound patterns and encode these into memory. An open-set word recognition test like the PBK Test has no external context or response constraints that listeners can use to help them perceive the target word other than the knowledge of sound patterns and the regularity of words in their mental lexicons. Responding to a word in isolation therefore requires that the listener encode the sound pattern into working memory, and then access a motor program for the word from information stored in the lexicon to produce an utterance on demand in a repetition task. Children may therefore have difficulty with an open-set test not only because the vocabulary is unfamiliar but because they are not able to encode, represent or access the sound pattern of these novel items from memory in what is essentially an imitation task. Either or both of these alternatives are possible accounts of why the PBK Test is difficult for young children with profound hearing losses. Because of the importance placed on open-set word recognition tests as a criterion for evaluating performance with CIs, we believe that it is critical at this time to gain an understanding of the reasons for the generally poor performance observed on the PBK test by children with hearing impairments and to analyze the specific task demands of this test.

Haskins (Reference Note 2) made several important observations that remain relevant to an understanding of speech perception by young children with profound hearing losses despite the passage of 50 yr. In particular, based on our analyses reported below, we believe it is possible to provide a principled explanation of why the PBK Test is a difficult test for some children with CIs. Several new computational analyses of the specific words used on the PBK Test were conducted, and these results provide some new insights into the primary factors that influence word recognition in open-set tests.

Although not widely known, Haskins actually constructed four separate lists of PBK words and then collected speech perception data on these materials, which were included in her thesis. However, only three of these original lists are currently being used by audiologists in the clinic and in research laboratories. Apparently, the fourth PBK list was never used because, according to Haskins, her speech perception scores with adult listeners with normal hearing showed that the words on this particular list (PBK List 2) were “more audible” than the words used on the other three lists. To the best of our knowledge, no speech perception data were ever collected from young children with either normal hearing or hearing losses, the target population that the materials were originally developed for at the time. Because of the important role that the PBK Test has played in recent years in assessments of speech perception and word recognition skills in young children, we critically examine both issues below.

In the first section of this paper, we report new computational analyses on all four of Haskins’ original word lists. The results of these analyses permit us to offer a theoretically motivated explanation for the differences in speech perception she observed among the lists of words with adult listeners. Not only can we now account for differences among the four original lists, but these new computational analyses of Haskins’s data also provide additional support for a recent model of spoken word recognition—NAM (Luce, Reference Note 5; Luce & Pisoni, 1998; Luce et al., 1990). NAM is a model of speech perception and spoken word recognition whose underlying hypothesis is that words are recognized relationally in the context of other phonetically similar words. This model attempts to account for the effects of lexical activation and competition on spoken word recognition. In particular, it is assumed that a sound pattern activates multiple lexical items in memory and this pattern of activation affects lexical discrimination and subsequent recognition performance (Luce & Pisoni, 1998).

NAM predicts that when a spoken word is perceived, it activates a set of representations of similar sounding words in memory. This set of words is called a “lexical neighborhood.” Once the neighborhood of the stimulus is activated, the word recognition system must then discriminate among the activated neighbors or candidates and decide which of the neighbors best matches the stimulus. The lexical discrimination and decision process is influenced by three factors: 1) the frequency of the target word; 2) the number of phonetically similar words (neighbors) activated in memory by the stimulus pattern; and 3) the frequencies of occurrence of the neighbors in the similarity neighborhood. Because of increased competition among activated items in memory, stimuli from densely populated lexical neighborhoods are more slowly and less accurately recognized than stimuli from sparsely populated neighborhoods. In addition, the presence of high-frequency neighbors produces increased competition among words that are similar to the target word, thereby also reducing recognition performance. According to NAM, stimulus words with few low-frequency neighbors will be identified most accurately, and stimulus words with many high-frequency neighbors will be identified least accurately.

Using NAM of spoken word recognition as our theoretical framework, we are able to offer an explanation for why some words are easy to recognize and why other words are more difficult to recognize in an open-set test format (see Luce & Pisoni, 1998). Finally, we offer some comments about the implications of these findings for assessment of speech perception by children with profound hearing losses.

Methods

Word Lists

Haskins recorded a single randomization of the four lists of the PBK Test. She then presented the four PBK lists and one of the PB-50 lists (List 13) at sequentially higher signal levels to 22 adult listeners with normal hearing. The words were presented at -7.8 , -2.8 , 2.2 , 12.2 , and 27.2 dB SL (0 dB SL was determined as the mean threshold for spondaic words). Listeners were asked to identify the words in an open-set format.

Lexical Neighborhood Database

The four PBK lists of Haskins (Reference Note 2) and List 13 of the PB-50 words (Egan, 1948) were analyzed using several techniques to compute similarity spaces for spoken words (see Pisoni, Nusbaum, Luce, & Slowiaczek, 1985). A computational analysis of the specific words used on all five lists was carried out using a computerized version of a 20,000-word Webster's Pocket Dictionary (Pisoni et al., 1985). From the pocket dictionary, we obtained word frequencies (Kucera & Francis, 1967), as well as lexical neighborhood frequencies and densities. Word frequency refers to the frequency counts from the Kucera and Francis (1967) norms. A "lexical neighborhood" of a word is defined as all words in the pocket dictionary that differ from the stimulus word by a single phoneme substitution, addition, or deletion (Greenberg & Jenkins, 1963). For example, if the stimulus is "pit," an example of a neighbor by substitution is "bit," by deletion "it," and by addition "spit." Neighborhood density refers to the number of words in the lexical neighborhood, and neighborhood frequency refers to the mean of the frequency counts (Kucera & Francis, 1967) of the words in the lexical neighborhood. The mean word frequency was calculated for all the words in each list. Because a few of the words had very high frequencies (e.g., "and" in List 2), the median word frequency was also calculated for each list. Mean and median lexical neighborhood frequencies and densities were also calculated for the words in each individual list. Two additional measures were calculated for each word, the ratio of the word frequency to the neighborhood frequency and the ratio of the word frequency to the neighborhood density. These "second-order statistics" provide measures of the relation of a specific word to its neighbors and are used to quantify the amount of lexical competition among phonetically similar words in lexical memory.

Results

Haskins stated that she felt that List 2 was "easier" than the other lists at all levels of presentation, although a significant difference among the lists emerged only at 12.2 and 2.2 dB SL. She also stated that an "item analysis" of the individual words in PBK List 2 might yield additional information on ways to improve the list, or to make it more like the other lists, but she did not indicate what kind of item analysis to do or what perceptual dimensions of spoken words might be relevant to speech perception performance or to the differences in performance she observed. The results from these intelligibility tests are shown in Table 1 (adapted from Haskins, Reference Note 2, Table VI). According to Haskins, at the middle intensity tested (2.2 dB SL), List 2 is significantly easier than Lists 1, 3, and 13. At the next higher intensity tested (12.2 dB SL), List 2 is significantly easier than all the other lists. At the highest intensity tested (27.2 dB SL), List 2 is significantly easier than only List 13, but at this level, performance is nearly perfect, and any differences among lists are confounded by ceiling effects.

In examining the data in Table 1, one can see that performance on the PBK Test is clearly related to the presentation level of the stimulus. The corresponding slopes of the articulation functions between 20 and 80% correct response are approximately 4%/dB. At the ends of

the functions (0 to 20%, 80 to 100%), performance changes little with a change in the level of presentation (approximately 1%/dB). Thus, for listeners with hearing losses, increasing the presentation level of the stimulus words by 5 dB could amount to as much as a 20% increase in performance, or very little increase in performance depending on the region of the curve the subject is working and the subject's range of "useable hearing."

Haskins also examined test-retest reliability for the word lists at only a single presentation level (12.2 dB SL). She found a large amount of improvement between the first and second repetitions of the test with the average score for List 2 increasing approximately 5.5%, and the average score for the remaining lists increasing approximately 8.8%. Significant improvement was seen for Lists 1 and 4 ($p < 0.05$), Lists 3 and 13 ($p < 0.01$), but not for List 2 ($p > 0.05$).

The results of the computational analysis of the lexical properties of the words on the PBK lists are displayed in Table 2. Here, we show both the mean and median word frequency, neighborhood frequency, and neighborhood density for the test items on the five word lists. In examining the mean word frequency, the reader will notice a very large difference in the mean frequency between List 2 (954.4) compared with the mean frequencies of the other three lists (295.6 to 689.1). On closer inspection, List 2 contains the word "and," which is one of the six most common words in the English language. The frequency of "and" is very high and skews the frequency distribution of List 2 a great deal. However, if we examine the median word frequency for the different lists, we still find that List 2 has a higher median frequency (110) than the other three lists (13 to 75) ($p < 0.001$). Thus, the words in List 2 do occur more frequently in the language than the words in the other lists, which may account, in part, for why List 2 was "more audible" than the other PBK lists.

Little difference exists in the mean neighborhood frequencies or densities between List 2 and the other test lists. The mean neighborhood frequency of List 2 (261.4) is within the range of mean neighborhood frequencies for the other three lists (196.9 to 294.6). The mean neighborhood density of List 2 (16.4 words) is slightly less than the mean neighborhood densities of the other lists (17.0 to 17.7 words). However, these differences are not very large. Thus, it appears that the mean neighborhood frequency and density of the words for the different PBK lists do not provide us much useful information about the reasons for the differences in audibility observed between the lists. However, these three measures (word frequency, neighborhood frequency, and neighborhood density) are absolute values for the words themselves without regard to the effects of other phonetically similar words. Put another way, these are "first-order" measures that do not take into account the relational properties of words to other phonetically similar words in the lexicon.

The next step in the analysis of the lexical properties of the PBK words was to generate ratios of the individual word frequencies to the neighborhood frequencies and densities of the individual words. These measures are "second-order" statistics that capture the relational properties of words to their lexical neighborhoods. Both mean and median values are listed in Table 2, but because of the effect of very high-frequency words (e.g., "and"), we will just examine the median values. Ratios were generated by dividing the frequency of the individual words by the neighborhood frequency of the individual words. The median of these ratios for List 2 was approximately 1.4, whereas the median ratios for the three other PBK lists (0.2 to 0.9) were much smaller than the median ratio for List 2. This finding demonstrates that the words in List 2 have a higher frequency of occurrence than the other words in their lexical neighborhoods, whereas the words in the remaining lists have a lower frequency of occurrence than the words in their lexical neighborhoods. Put another way, this analysis demonstrates that there is less lexical competition among phonetically similar words on PBK List 2 than the words on the other three PBK lists.

A second statistic was generated by dividing the frequency of the individual words by the neighborhood density of the individual words. The results of this analysis showed that the median ratio for List 2 (5.3) was larger than the median ratios for the other three PBK lists (0.7 to 5.0). As reported earlier, the median neighborhood densities for the different PBK lists were quite similar (range 16.4 to 17.7 words); however, this statistic provides further evidence that there is less lexical competition among phonetically similar words on PBK List 2 than the words on the other three PBK lists. Thus, words on List 2 should be less confusable than words on the other three lists. In short, these analyses of the similarity spaces or lexical neighborhoods of the words demonstrate that the words in PBK List 2 are inherently more distinctive and discriminable than the words on the three other PBK lists. If we assume that in open-set tests spoken words are recognized relationally in the context of other phonetically similar words, these computational analyses would predict that the words on PBK List 2 should be more acoustically distinctive and therefore more easily recognizable than the words on any of the other three lists. This is precisely the result that Haskins found in her thesis in 1949 almost 50 yr ago.

Discussion

The results from the present analysis of the specific words used on the PBK Test give further support to NAM (Luce, Reference Note 5; Luce & Pisoni, 1998; Luce et al., 1990). The lexical discrimination and decision process is influenced by three factors: 1) the frequency of the target word; 2) the number of phonetically similar words (neighbors) activated in memory by the stimulus pattern; and 3) the frequencies of occurrence of the neighbors in the similarity neighborhood.

The present analysis of the items on the PBK Test revealed that the median values of word frequency for the words in List 2 are considerably higher than the median values of word frequency for the words in the other lists. According to Pollack, Rubenstein, and Decker (1959), listeners are more likely to correctly identify a word with a high-frequency of occurrence than a word with a low-frequency of occurrence at unfavorable signal to noise ratios when the message set is unknown. Thus, one would expect, and NAM predicts that on the average, at unfavorable signal to noise ratios, listeners would be more likely to correctly identify a word from List 2 than a word on the other lists. Moreover, scores with List 2 should be better than scores for the other lists at points along portions of the psychometric functions generated from the data in Table 2. At the end of the functions where the sensation levels of the stimuli are very low, the effect of the frequency of the stimulus word on the identification of that word is not very important, because there is little useful information in the signal at a presentation level at which some of the words are just becoming intelligible (very low SL). At the other end of the psychometric functions, the presentation level of the stimuli is high enough so that all the words are easily intelligible, and word frequency again has little effect on identification.

The results of our analyses also revealed that the ratio of the frequency of the stimulus to the frequency of the neighborhood of the stimulus was higher for List 2 than the other lists. If the frequency of the stimulus is relatively great compared with the frequency of the neighborhood of that stimulus, then the probability of correctly identifying the stimulus should be high. Again, this effect is more important at intermediate SLs where the stimulus is neither too easy nor too difficult to identify. The ratio of the frequency of the stimulus to the density of the neighborhood was also more favorable for the words on List 2 than the words on the other three PBK lists. If the density or number of items in a lexical neighborhood is low in relation to the frequency of the stimulus, the probability of a listener correctly identifying the stimulus should be high. Thus, we find that both absolute and

relative lexical neighborhood measures obtained in the present study lend further support to NAM.

The findings reported here are of theoretical interest. First, the PBK Test was constructed to be a test of word identification for children, and especially children with hearing losses. However, Haskins measured the “equivalence” of the PBK lists using adult subjects with normal hearing, not children with normal hearing or hearing losses. Nevertheless, the PBK Test has proven to be a very useful test for children, and it has probably been one of the most important measures of open-set word recognition for children with CIs. Second, are the lexical neighborhood differences observed here important for the target group the test was designed for and under the conditions the test is normally given? The words chosen for the PBK Test were selected to be within the vocabulary of a child based on norms for preschool children (The International Kindergarten Union, Reference Note 9). In a recent study, Kluck, Pisoni, and Kirk (Reference Note 3) obtained speech perception scores from twenty-two 3-yr-olds and eight 4-yr-olds with normal hearing using recorded versions of the PBK Test and the LNT. The words were presented to the children at approximately 70 dB SPL in an auditory-only format. The 4-yr-old children performed at near-ceiling levels for both tests with performance on the PBK Test just slightly lower than performance on the LNT. Performance for the 3-yr-old children was slightly lower than performance for the 4-yr-old children on both tests. These results from the children with normal hearing provide a benchmark for children with impaired hearing and CIs. Clearly, 3- and 4-yr-old children with normal hearing are able to imitate and repeat all the test items on both lists without any difficulty.

Kluck et al. (Reference Note 3) addressed the issue of word familiarity in the different lists by asking the parents of the children whether or not their child was familiar with the word, using a seven-point scale. They found that all the words on both the PBK Test and the LNT were reported to be very familiar to the 4-yr-old children, with the mean familiarity rating of the words from the PBK Test rated 6.58 (out of 7) and the words from the LNT rated 6.87 (out of 7). The words were slightly less familiar to the 3-yr-old children (mean familiarity of PBK words: 5.95; mean familiarity of LNT words: 6.74 out of 7). Thus, it appears that the words on the PBK Test are less familiar to the 3-yr-olds than to the 4-yr-olds, but most of the words on the PBK Test are familiar to preschool-aged children.

The present findings are also of clinical interest. The PBK Test is usually given by an audiologist through either monitored live voice or a tape-recorded version of the test at approximately 70 dB SPL. The most intense level Haskins presented to her subjects was approximately 27 dB above the threshold (estimated threshold for spondaic words) for her group of subjects. If we assume that the average speech reception threshold for spondaic words is approximately 20 dB SPL (ANSI S3.6, 1989; Young, Dudley, & Gunter, 1982) for a group of young adults with normal hearing, then Haskins presented the PBK lists to her subjects at approximately 47 dB SPL. We would expect performance to be perfect or nearly perfect for adults with normal hearing at a presentation level of 70 dB SPL, regardless of the PBK list tested, and that no differences between the lists would be apparent at this level.

For children with normal hearing, we would also expect nearly perfect performance at 70 dB SPL assuming that the children were able to perform the task and actually knew the words being tested. For children with hearing losses, their performance should be related to the level at which the word lists are presented. If the child has useable hearing and is able to understand speech well, and if the PBK words are presented at a high enough level, the child should perform well (nearly 100%) on the PBK Test (and performance should be nearly identical for the four different PBK lists). If the PBK lists cannot be presented at an appropriate level, or the child does not have useable hearing, then we would expect

performance on the PBK Test to be low, and we might expect to see differences emerge between scores on PBK List 2 and the other PBK lists based on the lexical differences between the lists reported above, if List 2 were to ever be used.

Although the PBK Test may have certain shortcomings, it is certainly a difficult and probably the most important test of open-set word recognition for children with profound hearing losses both with and without CIs. The tasks involved in the correct repetition of a stimulus in an open-set test of word recognition place a number of information processing demands on these children. The listener must search and retrieve words from lexical memory by encoding phonemic differences based on information present in the speech signal without the aid of any external context or retrieval cues and then discriminate and select a pattern from memory. The listener must then use this information to form an articulatory plan or a motor program to produce a verbal or motor response to the stimulus. The various perceptual and cognitive operations needed for spoken word recognition require access to a variety of memory codes and neural representations of speech and spoken language at different levels of analysis. The speed and efficiency of these information-processing operations, particularly as they might be employed in tasks requiring transformation and mapping from perception to production, will depend to a large extent on having representations of words in memory and organizing these representations systematically in a lexicon that can be accessed efficiently to provide different sources of information about the words in the language.

We know that pediatric CI users recognize words relationally in accordance with the predictions of NAM. This was demonstrated by Kirk et al. (1995) using the LNT, which was designed especially to examine this question. It is possible that the lexical neighborhoods of children with CIs are fundamentally different from the lexical neighborhoods of children with normal hearing. We predict that because of poorer discrimination abilities, the neighborhoods of the CI users will be broader (i.e., contain more words) than the neighborhoods of children with normal hearing. For example, let us assume that a CI user has difficulty discriminating the place of articulation of a consonant. If the stimulus word presented to the listener is “pan,” that listener would have difficulty discriminating the stimulus from words that vary from the stimulus word in the place of articulation of the initial consonant (e.g., “tan” and “can”). For a listener with normal hearing, the lexical neighborhood of the word “pan” includes the words “tan” and “can” as they vary from the stimulus by a single phoneme substitution. For the CI user who has difficulty discriminating the place of articulation of consonants, the lexical neighborhood of the stimulus “pan” might also contain words in the lexical neighborhoods of “tan,” and “can.” Thus, the structure of the lexical neighborhoods of pediatric CI users may be quite different from the structure of the lexical neighborhoods of a listener with normal hearing.

Although the lexical neighborhoods of a pediatric CI user may be broader than the lexical neighborhoods of a listener with normal hearing, the average size of the receptive vocabulary of the pediatric CI user or deaf child is considerably smaller than the average size of the receptive vocabulary of a group of children of the same age with normal hearing (Geers & Moog, 1994; Miyamoto, Osberger, Robbins, Myres, Kessler, & Pope, 1992). A smaller vocabulary should lead, therefore, to lexical neighborhoods that are, in general, less dense (contain fewer words) than lexical neighborhoods for listeners with normal hearing and normal vocabulary development. At the present, we are conducting an analysis of the errors made by pediatric CI users with 2 and 3 yr of implant use on open-set tests of word recognition (Chin, Meyer, Hay-McCutcheon, Wright, & Pisoni, Reference Note 1). This analysis should give us a better understanding of the neighborhood structure of the children with CIs and provide insight into how these similarity neighborhoods change with increased implant use. As the ability of a pediatric CI user to make fine distinctions between the

different acoustic-phonetic properties of words improves, the listener's lexical neighborhood structure for a particular word should become more refined and decrease in size. Preliminary analyses of the error data provide support for this prediction.

Acknowledgments

This study was supported by NIH/NIDCD Grants DC00064, DC00111, and NIH Training Grant DC00012. We would like to thank Darla J. Sallee for clerical assistance. We would also like to thank Dr. Steven B. Chin and the reviewers for comments and suggestions on an earlier version of the manuscript.

References

1. American National Standards Institute. Specifications for audiometers. (ANSI S3.6-1989). New York: ANSI; 1989.
2. Carney, AE.; Osberger, MJ.; Miyamoto, RT.; Karasek, A.; Dettman, DL.; Johnson, DL. Speech perception along a continuum: From hearing aids to cochlear implants. In: Feigin, JA.; Stelmachowicz, PG., editors. Pediatric amplification. Omaha, NE: Boystown National Research Hospital; 1991. p. 93-113.
3. Dunn, L.; Dunn, L. Peabody picture vocabulary test—revised. Circle Pines, MN: American Guidance; 1981.
4. Egan JP. Articulation testing methods. *Laryngoscope*. 1948; 58:955–991. [PubMed: 18887435]
5. Fryauf-Bertschy H, Tyler RS, Kelsay DM, Gantz BJ, Woodworth GG. Cochlear implant use by prelingually deafened children:: The influences of age at implant and length of device use. *Journal of Speech, Language, and Hearing Research*. 1997; 40:183–199.
6. Gantz BJ, Tyler RS, Woodworth GG, Tye-Murray N, Fryauf-Bertschy H. Results of multichannel cochlear implants in congenital and acquired prelingual deafness in children: : Five-year followup. *American Journal of Otology*. 1994; 12(Suppl):1–7. [PubMed: 8572105]
7. Geers AE, Moog JS. Spoken language results: : Vocabulary, syntax, and communication. *The Volta Review*. 1994; 96:131–148.
8. Greenberg JH, Jenkins JJ. Studies in the psychological correlates of the sound system of American English. *Word*. 1963; 20:157–177.
9. Kirk KI, Pisoni DB, Osberger MJ. Lexical effects on spoken word recognition by pediatric cochlear implant users. *Ear and Hearing*. 1995; 16:470–481. [PubMed: 8654902]
10. Kucera, F.; Francis, W. Computational analysis of present-day American English. Providence, RI: Brown University Press; 1967.
11. Lane H. Acquisition of speech perception ability in prelingually DEAF children with a multi-channel cochlear implant [Letter to the Editor]. *American Journal of Otology*. 1995; 16:393–399. [PubMed: 8588637]
12. Luce PA, Pisoni DB. Recognizing spoken words: : The neighborhood activation model. *Ear and Hearing*. 1998; 19:1–36. [PubMed: 9504270]
13. Luce, PA.; Pisoni, DB.; Goldinger, SD. Similarity neighborhoods of spoken words. In: Altmann, GTM., editor. Cognitive models of speech processing: Psycholinguistic and computational perspectives. Cambridge, MA: MIT Press; 1990. p. 122-147.
14. MacWhinney B, Snow C. The child language data exchange system. *Journal of Child Language*. 1985; 12:271–296. [PubMed: 4019604]
15. Miyamoto RT, Osberger MJ, Robbins AM, Myres WA, Kessler K. Prelingually deafened children's performance with the Nucleus multichannel cochlear implant. *American Journal of Otology*. 1993; 14:437–445. [PubMed: 8122704]
16. Miyamoto RT, Osberger MJ, Robbins AM, Myres WA, Kessler K, Pope M. Longitudinal evaluation of communication skills of children with single- or multichannel cochlear implants. *American Journal of Otology*. 1992; 13:215–222. [PubMed: 1609849]
17. Osberger MJ, Miyamoto RT, Zimmerman-Phillips S, Kemink JL, Stroer BS, Firszt JB, Novak MA. Independent evaluation of the speech perception abilities of children with the Nucleus 22- channel cochlear implant system. *Ear and Hearing*. 1991; 12(Suppl):S66–S80.

18. Pisoni DB, Nusbaum HC, Luce PA, Slowiaczek LM. Speech perception, word recognition and the structure of the lexicon. *Speech Communication*. 1985; 4:75–95.
19. Pollack I, Rubenstein H, Decker L. Intelligibility of known and unknown message sets. *Journal of the Acoustical Society of America*. 1959; 31:273–279.
20. Staller SJ, Dowell RC, Beiter AL, Brimacombe JA. Perceptual abilities of children with the Nucleus 22-channel cochlear implant. *Ear and Hearing*. 1991; 12(Suppl):S34–S47.
21. Svirsky MA. Speech production and language development in pediatric cochlear implant users. *Journal of the Acoustical Society of America*. 1996; 99:2570.
22. Waltzman SB, Cohen NL, Gomolin RH, Shapiro WA, Ozdamar SR, Hoffman RA. Long -term results of early cochlear implantation in congenitally and prelingually deafened children. *American Journal of Otology*. 1994; 12(Suppl):9–13. [PubMed: 8572107]
23. Young LL, Dudley B, Gunter MB. Thresholds and psychometric functions of the individual spondaic words. *Journal of Speech and Hearing Research*. 1982; 25:586–593. [PubMed: 7162160]

Reference Notes

1. Chin, SB.; Meyer, TA.; Hay-McCutcheon, M.; Wright, GA.; Pisoni, DB. Seventh Symposium on Cochlear Implants in Children. Iowa City, IA: 1998. Understanding the structure of the mental lexicons of children who use cochlear implants.
2. Haskins, H. Unpublished master's thesis. Evanston, IL: Northwestern University; 1949. A phonetically balanced test of speech discrimination for children.
3. Kluck, M.; Pisoni, DB.; Kirk, KI. Performance of normal-hearing children on open-set speech perception tests (Progress Report No. 21). Bloomington, IN: Speech Research Laboratory, Indiana University; 1997.
4. Logan, JS. A computational analysis of young children's lexicons (Technical Report No. 8). Bloomington, IN: Speech Research Laboratory, Indiana University; 1992.
5. Luce, PA. Neighborhoods of words in the mental lexicon (Technical Report No. 6). Bloomington, IN: Speech Research Laboratory, Indiana University; 1986.
6. 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 in pediatric cochlear implant users (Progress Report No. 21). Bloomington, IN: Speech Research Laboratory, Indiana University; 1997.
7. Reynell, JK.; Gruber, CP. Reynell developmental language scales: U.S. edition. Los Angeles, CA: Western Psychological Services; 1990.
8. Robbins, AM.; Renshaw, JJ.; Miyamoto, RT.; Osberger, MJ.; Pope, MJ. Minimal pairs test. Indianapolis, IN: Indiana University School of Medicine; 1988.
9. The International Kindergarten Union, Child Study Committee. A study of vocabulary of children before entering the first grade. Washington, D.C.: 1928.

Table 1

Mean percent correct responses and standard deviations for PBK Lists 1 to 4 and PB-50 List 13. Values averaged across 22 adult listeners with normal hearing (adapted from Haskins, Reference Note 2, Table VI).

Sensation Level	PBK List 1		PBK List 2		PBK List 3		PBK List 4		PB-50 List 13	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
-7.8	3.4	3.8	7.6	8.9	6.7	6.3	6.8	5.7	5.7	5.6
-2.8	16.3	10.3	24.2	10.8	16.8	8.9	19.5	9.1	17.5	8.2
2.2	33.8	10.6	44.1	11.9	36.6	9.5	39.3	12.5	33.4	8.2
12.2	69.9	8.8	82.2	11.2	71.5	11.2	73.2	9.0	68.2	7.8
27.2	98.1	3.1	98.4	1.7	97.1	2.8	97.7	2.4	95.7	1.9

Table 2

Lexical neighborhood measures for PBK Lists 1 to 4 and PB-50 List 13. Word Frequency/Mean Neighborhood Frequency or Neighborhood Density values were calculated for the individual words in the lists. The mean and median values are presented here.

	<u>PBK List 1</u>	<u>PBK List 2</u>	<u>PBK List 3</u>	<u>PBK List 4</u>	<u>PB-50 List 13</u>
Mean (Word Frequency)	558.6	954.4	689.1	612.7	295.6
Mean (Mean Neighborhood Frequency)	294.6	261.4	231.5	272.5	196.9
Mean (Neighborhood Density)	17.0	16.4	17.5	17.7	17.3
Mean (Word Frequency/Mean Neighborhood Frequency)	4.4	8.0	2.0	3.3	1.9
Mean (Word Frequency/Neighborhood Density)	35.3	94.1	42.5	41.3	13.0
Mean (Word Frequency)	62.0	110.0	46.0	75.0	13.0
Mean (Word Frequency/Mean Neighborhood Frequency)	0.9	1.4	0.9	0.9	0.2
Mean (Word Frequency/Neighborhood Density)	3.9	5.3	2.7	5.0	0.7