Published in final edited form as:

Otol Neurotol. 2003 July ; 24(4): 612–620.

Modeling Open-Set Spoken Word Recognition in Postlingually Deafened Adults after Cochlear Implantation: Some Preliminary Results with the Neighborhood Activation Model

Ted A. Meyer, Stefan A. Frisch, David B. Pisoni, Richard T. Miyamoto, and Mario A. Svirsky

Department of Otolaryngology–Head and Neck Surgery, University of Iowa Hospitals and Clinics, Iowa City, Iowa, U.S.A.

Abstract

Hypotheses—Do cochlear implants provide enough information to allow adult cochlear implant users to understand words in ways that are similar to listeners with acoustic hearing? Can we use a computational model to gain insight into the underlying mechanisms used by cochlear implant users to recognize spoken words?

Background—The Neighborhood Activation Model has been shown to be a reasonable model of word recognition for listeners with normal hearing. The Neighborhood Activation Model assumes that words are recognized in relation to other similar-sounding words in a listener's lexicon. The probability of correctly identifying a word is based on the phoneme perception probabilities from a listener's closed-set consonant and vowel confusion matrices modified by the relative frequency of occurrence of the target word compared with similar-sounding words (neighbors). Common words with few similar-sounding neighbors are more likely to be selected as responses than less common words with many similar-sounding neighbors. Recent studies have shown that several of the assumptions of the Neighborhood Activation Model also hold true for cochlear implant users.

Methods—Closed-set consonant and vowel confusion matrices were obtained from 26 postlingually deafened adults who use cochlear implants. Confusion matrices were used to represent input errors to the Neighborhood Activation Model. Responses to the different stimuli were then generated by the Neighborhood Activation Model after incorporating the frequency of occurrence counts of the stimuli and their neighbors. Model outputs were compared with obtained performance measures on the Consonant-Vowel Nucleus-Consonant word test. Information transmission analysis was used to assess whether the Neighborhood Activation Model was able to successfully generate and predict word and individual phoneme recognition by cochlear implant users.

Results—The Neighborhood Activation Model predicted Consonant-Vowel Nucleus-Consonant test words at levels similar to those correctly identified by the cochlear implant users. The Neighborhood Activation Model also predicted phoneme feature information well.

Conclusion—The results obtained suggest that the Neighborhood Activation Model provides a reasonable explanation of word recognition by postlingually deafened adults after cochlear implantation. It appears that multichannel cochlear implants give cochlear implant users access to

© 2003, Otology & Neurotology, Inc.

Address correspondence and reprint requests to Ted A. Meyer, M.D., Ph.D., Department of Otolaryngology–Head and Neck Surgery, University of Iowa Hospitals and Clinics, 200 Hawkins Drive, 2nd Floor Pomerantz Pavilion, Iowa City, IA 52246, U.S.A.; ted-meyer@uiowa.edu.

Presented at the 2001 American Neurotology Society Meeting, May 12–13, 2001, Palm Desert, California, U.S.A.

their mental lexicons in a manner that is similar to listeners with acoustic hearing. The lexical properties of the test stimuli used to assess performance are important to spoken-word recognition and should be included in further models of the word recognition process.

Keywords

Cochlear implantation; Open set; Word recognition

Understanding spoken words and constructing lexical representations of these words in a manner that allows one to communicate effectively are very complex processes. Nevertheless, by the time we reach adulthood, humans acquire a working database, a mental lexicon, containing many thousands of words. We are able to access words from our lexicons with great success under a wide variety of listening situations: in isolation; in running speech; if spoken by a woman, man, or child; whether presented slowly or fast; and in noisy or otherwise degraded listening conditions.

Even after losing the ability to hear acoustically, many adults achieve open-set spoken word recognition with a cochlear implant (CI) (1–5). Presumably, adult CI users who lost their hearing after developing normal language maintain a central representation of language that has a structure similar to the representation they had when they heard acoustically. The CI provides enough stimulation to allow listeners to access this lexical information.

Much research has focused on measuring improvements in overall word recognition performance. As technology improves, and as the criteria for implantation change, average word recognition performance continues to increase for CI users (2,3). A significant amount of work has been undertaken to identify variables that are associated with variability in outcome performance between individual CI users (6–9). In contrast, little research has attempted to explain either the perceptual mechanisms CI users may be using to understand words and speech within a psycholinguistic framework. In studies examining consonant confusions made by listeners with varying degrees of sensorineural hearing loss (SNHL), Bilger and Wang (10,11) reported that listeners with similar patterns of hearing loss can demonstrate vastly different levels of performance. Furthermore, the confusions made by the individual subjects illustrate that different listeners may be focusing on different aspects of the stimuli in attempts to improve performance.

CI users demonstrate even greater variability in their abilities to recognize phonemes, words, and speech than do listeners with SNHL who hear acoustically. Perhaps some of the variability in word recognition is attributable to large differences in the number of surviving spiral ganglion cells for different CI users. Given the greatly varying number of surviving neural elements, CI users may use different cognitive or linguistic strategies to maximize performance. Can we explain these perceptual mechanisms and intersubject differences? The variability in performance and possibly in perceptual mechanisms used to understand speech cannot be addressed simply by measuring performance longitudinally and reporting gains in the percentage of words correctly identified over time with a CI.

In a related line of research, several authors have examined relations between phoneme perception and word recognition within the same stimulus. Boothroyd and Nitttrouer (12,13) described relations between phoneme and word recognition within the same stimulus in terms of a “context model.” A power-law relation exists between the number of words correctly identified and the number of phonemes in those words correctly identified. The relation is captured in the following equation

$$P(C)_w = [P(C)_p]^j$$

where $P(C)_w$ is the percentage of words correctly identified by the listener, and $P(C)_p$ is the percentage of phonemes in those words correctly identified. The exponent j is an uncertainty factor that is related to the difficulty of the words for the listeners (context). For example, given a list of three-phoneme nonwords, the exponent j would be equal to the number of phonemes, three. Given context for a list of three-phoneme words that are completely familiar to the listeners, the exponent would be substantially less than 3. The model has been previously fit to data obtained from young adults (12), older adults and children (13), adult CI users (postlingual profound SNHL) (14), and children with CIs (prelingual profound SNHL) (SA Frisch, TA Meyer, DB Pisoni, submitted).

Frisch et al. (SA Frisch, TA Meyer, DB Pisoni, submitted) found that for children with profound prelingual SNHL who received the Nucleus 22-channel CI, the choice of speech-processing strategy (multipeak [MPEAK] (2), spectral peak [SPEAK] (3)) affects perceptual ability but not the context factor, j . In other words, $P(C)_w$ and $P(C)_p$ were better for the SPEAK users than for the MPEAK users, but the exponent j was virtually identical for the two groups. These authors also found that word list affects context. The listeners scored the same percentage of phonemes correctly on different tests of monosyllabic words, but they did more poorly with words that were more lexically difficult. In other words, as lexical uncertainty increased, the exponent in Equation 1 increased as predicted. Therefore, the power-law relation holds not only for adults and children with normal hearing but also for adults and children with CIs.

We propose to examine phoneme and word recognition through a different and novel approach. We use a model of spoken-word recognition to not only predict word recognition performance but also to actually determine or generate open-set spoken word recognition from performance on closed-set phoneme recognition tasks. Through mathematical modeling, we aim to provide further insights into the cognitive mechanisms involved in word recognition with a CI. By gaining a better understanding about these mechanisms, we may be able to explain some of the variability in performance between different CI users and between the different processing strategies. This line of research may eventually further improvements in speech-processing strategies for CI users.

In the current study, we determine whether access to the mental lexicon provided by the CI is in any way similar to that provided by “normal acoustic hearing” by applying a current model of word recognition to phoneme and word recognition. This approach is conceptually different from the research described in the preceding paragraphs. Instead of simply obtaining scores on different tests and performing regression analysis to measure relations between different variables, we used performance on closed-set phoneme recognition tasks to generate responses to an open-set word recognition task with the Neighborhood Activation Model (NAM). Model output is compared with responses obtained from the CI users for the individual phonemes in the stimulus words and the words themselves. In this modeling process, the ability of the model to predict or generate the same mistakes as the CI users is as important as the ability of the model to predict or generate correct responses. Again, through mathematical modeling, we aim to provide some new insights into possible mechanisms used by CI users to understand spoken words.

Computational models have been developed by several researchers (15–17) to explain how we perceive words. One particular model, TRACE (15), proposes that as we hear more phonemes in a word, the number of possible choices narrows until there is only one

remaining response. This is not a perfect system. When noise is added to the model, or if any of the phonemes are perceived incorrectly, the model generates inaccurate predictions of subjects' performance.

NAM proposed by Luce and Pisoni (16,17) was implemented to predict and generate open-set spoken word recognition from phoneme confusion matrices. NAM hypothesizes that spoken word recognition incorporates both bottom-up and top-down processing. Bottom-up processing refers to phoneme identification and top-down processing refers to the listener making decisions about what was heard and what response to the stimulus is to be made. Words are identified on the basis of the confusability of the individual phonemes and the relations of the stimulus word to other words that are phonetically similar (a lexical neighborhood). The underlying assumption of NAM can be summarized with a mathematical description. With NAM, the probability of correctly identifying a particular three-phoneme Consonant-Vowel Nucleus-Consonant (CVC) stimulus word ($c_1 v_1 c_2$) can be described as follows

$$p(\text{ID}) = \frac{p(c_1|c_1) * p(v_1|v_1) * p(c_2|c_2) * \text{freq}_s}{p(c_1|c_1) * p(v_1|v_1) * p(c_2|c_2) * \text{freq}_s + \sum_i p(c_i|c_1) * p(v_i|v_1) * p(c_i|c_2) * \text{freq}_i}$$

where $p(x_m | x_n)$ is the probability of responding phoneme (m) when phoneme (n) was presented. freq_s is the word usage frequency of the stimulus word, and freq_i is the word usage frequency of word $_i$ that lies in the lexical neighborhood of the stimulus word (e.g., Kucera and Francis (18)).

Luce and Pisoni have defined a lexical neighborhood of a particular word as group of words that can be confused with the target word. An “easy” word is a word of high usage frequency that has few words that are lexically similar. That is, an easy word is a common word from a sparse neighborhood. A hard word is a word with low usage frequency that has many words that are lexically similar. Under identical listening conditions, easy words are more recognizable than hard words.

The following example may help to make NAM easier to understand. Take two simple words, “bat” and “less.” For a particular listener, if we assume under a particular listening condition that the phoneme pairs “b” and “l,” “a” and “e,” and “t” and “s” are equally identifiable, then the two words “bat” and “less” should be recognized equally often. However, NAM predicts that the frequency of occurrence of a word is important to recognition. According to the word counts of Kucera and Francis (18), “less” (438 per 1 million words) is more common than “bat” (18 per 1 million words). NAM also predicts that the number of words that can be confused with the stimulus impacts the correct identification of the stimulus. For this example, a neighborhood of a stimulus word is defined as the words that are one phoneme away from the stimulus words by way of a single phoneme addition, deletion, or substitution. “Bat” has 33 neighbors (e.g., at, bet, cat), whereas “less” has only 19 neighbors (e.g., loss, bless, let). The neighbors of “bat” are on average more common words (mean frequency of occurrence, 512.2) than the neighbors of “less” (mean frequency of occurrence, 59.6). Therefore, “less” is a relatively common word in a sparse neighborhood of less common words, whereas “bat” is less common and lies in a relatively dense neighborhood of more common words. Computationally, “less” is easier to perceive than “bat.”

As a working model, NAM has been very successful in providing a theoretic framework to explain word recognition. For example, Charles-Luce et al. (19,20) have demonstrated that

lexical neighborhood properties of the individual syllables of spondaic words (e.g., baseball, hotdog) influence the identification of the other syllable. Differences in the intelligibility of the individual lists of the Phonetically Balanced Kindergarten test (H Haskins, unpublished dissertation) have been explained with NAM (21). The underlying assumptions of NAM also predict differences in reaction times to naming stimuli. Increased uncertainty about the stimulus words causes less accurate performance and longer decision times (22,23).

The overall goal of this project is to learn more about the underlying sensory and perceptual mechanisms used by CI users in understanding spoken words. Specifically, we aim to predict open-set spoken word recognition for a given CI user on the basis of the user's closed-set phoneme perception performance. To achieve this goal, we used a computational model, NAM. The model is explanatory; it does not simply attempt to correlate performance on open-set and closed-set tasks. Instead, NAM provides strict and verifiable hypotheses about how CI users use their ability to discriminate and identify different words.

Why is it important to develop computational models of word recognition by CI users? Although CIs represent the most successful example of a neural prosthesis that is clinically useful, the wide range of performance that is observed in patients who use these devices continues to baffle clinicians and scientists alike. Even when we consider only postlingually-deafened patients (i.e., those who became deaf after acquiring language), we find some fortunate patients who can easily converse on the telephone with unknown speakers, and others who only obtain a limited amount of lipreading enhancement. Attempts to correlate speech perception or word recognition performance with strictly peripheral factors (such as the number of surviving spiral ganglion cells in the cochlea) or demographic factors (such as duration of profound deafness) have met with limited success. It is our contention that a more thorough understanding of speech perception by CI users is highly desirable (and may even be essential) for facilitating future breakthroughs in CI design. The current study also hypothesizes that even a relatively simple task such as the identification of a single word may involve more central information-processing skills than a purely segmental, bottom-up approach would suggest. In other words, listeners in general and CI users in particular may identify words on the basis of the acoustic characteristics of the signal, but they also bring to the task their knowledge of the lexicon, phonotactics, and word frequency. Moreover, matching of acoustic inputs to candidate responses may involve the use of short-term memory. In consequence, individual differences in any of these higher order skills may explain some of the observed variability in clinical results obtained with CIs. The current study represents a small first step in the development of computational models that may be extremely useful to tease apart the contribution of sensory and cognitive factors to word recognition and speech perception by this clinical population.

PATIENTS AND METHODS

Participants

Twenty-six adults participated in the project. All of the participants lost their hearing as adults (postlingual profound SNHL) and used their CI for at least 11 months. Listeners used either the Nucleus (SPEAK strategy; Cochlear Corporation, Lane Cove NSW, Australia (3)) or the Clarion (CIS strategy; Advance Bionics Corporation, Sylmar, California, U.S.A. (24)) CI. Participants were paid for their time.

Apparatus

Digitized samples of the stimuli were stored on a personal computer. They were output through the computer's built-in sound card. An Acoustics Research speaker was used to

present the stimuli to the subjects acoustically at 70 dB sound pressure level (SPL). Testing was performed in a sound-attenuated booth (IAC).

Vowel Perception

We examined confusions to nine vowels in an /hVd/ format (Iowa medial vowel test (25)). The vowels are listed in Table 1. The stimuli were presented to the implant users in a nine-alternative, forced-choice paradigm. We obtained responses to at least 15 repetitions of each vowel.

More than nine vowels and diphthongs exist in American English. The online dictionary (26) used in the current study includes 15 separate vowels. The 9×9 vowel confusion matrices were expanded to 15×15 matrices by assuming similar levels of confusions across the different vowel features. The 15 vowels and their associated features along the three dimensions (height, place, and tense) are listed in Table 1. An example of an expanded vowel confusion matrix is shown in Table 2.

Consonant perception

We examined confusions to 24 consonants in an /aCa/ format (Iowa medial consonant test (25)). Again, the stimuli were presented through a computer from compact disk recordings. The consonants and their associated features are listed in Table 3. The samples were presented acoustically at 70 dB SPL to the implant users in a 24-alternative, forced-choice paradigm. We obtained responses to at least nine repetitions of each consonant.

Open-set spoken word recognition

Responses to the CNC monosyllabic word test (27) were examined. The test consists of 10 lists of 50 words. Each subject heard a single presentation of the words on at least three lists (150 words). The words were presented acoustically at 70 dB SPL to the implant users in a forced-response paradigm. The CI users were asked to respond with a CNC to the stimulus. There were essentially no nonresponses to the CNC words.

Modeling spoken word recognition with NAM

Experiment 1—According to NAM, words are identified on the basis of the confusability of the individual phonemes and on the relation of the stimulus word to other similar sounding words. We used a 1,338-word subset of CNC words from an online version of *Webster's Pocket Dictionary* (26) to assess the confusions that could be made by the CI users. All 1,338 CNC words in the online dictionary were considered as possible responses (denominator of Equation 2) if their phoneme strings could be obtained from the responses in the vowel and consonant confusion matrices. Word frequency counts were obtained from Kucera and Francis (18). To implement NAM, the probabilities from the vowel and consonant confusion matrices for the individual CI users and the word usage frequencies of the CNC words and their neighbors were used to calculate $p(\text{ID})$ for the various CNC stimuli. This calculation provides a baseline to demonstrate how well the model predicts overall performance.

Experiment 2—In addition to calculating $p(\text{ID})$, the next step in the modeling process was to devise a computational representation of NAM to generate different possible responses to the CNC stimuli. The model was divided into two stages: 1) bottom-up processing, or phoneme perception; and 2) top-down processing, which refers to the listener making decisions about what was heard and what response to the stimulus is to be made. Each CNC word was presented to the computational model 100 times. Bottom-up processing, the confusability of the individual phonemes, was simulated with the model outputting CNC

strings with the distribution of phonemes from the consonant and vowel confusion matrices obtained from the individual CI users (28) (M Broe, unpublished dissertation). Top-down processing, the relations of the stimulus word to other words that are phonetically similar, was simulated by calculating $p(\text{ID})$ for the different CNC strings output by the first stage of the model with Equation 2. Again, all 1,338 CNC words in the online dictionary were considered as possible responses (denominator of Equation 2). A few of the generated strings from the first step of the model were nonwords. Nonwords have a frequency of occurrence of 0, and according to Equation 2, if $\text{freq}_k = 0$, then $p(\text{ID})$ for these responses equals 0%. We therefore assumed that nonword responses were not possible and they were discarded. Computationally, this was quite simple to implement, and we feel that it was appropriate because the CI users do not respond with nonwords in these tasks.

For the real-word CNC strings, after $p(\text{ID})$ was calculated with Equation 2, the percentages of words and phonemes correctly identified were tallied. Information transmission (IT) analysis (29) was then carried out for the different phonemes. Percent IT for the phonemes in the responses generated to the CNCs by NAM was calculated and compared with percent IT calculated from the responses to the CNCs obtained from the individual CI users. If a phoneme feature (e.g., consonant voicing) was identified correctly on every response, IT for that feature would be 100%. What this means is that for every voiced consonant (e.g., b, d, g, m) in the stimulus, the response to the stimulus must be a voiced consonant (e.g., b, d, g, m). The response consonant does not have to be correct. It just has to be a voiced consonant. If the responses to a particular phoneme feature were random, IT for that feature would be 0%. IT allows us to examine more than just the percentage of words and phonemes correctly identified by the listeners or by NAM, it also provides insights into the pattern of confusions made by CI users and NAM.

RESULTS

Linear regression analysis

Phoneme and word recognition abilities were quite varied across the different listeners. For the Iowa medial consonants, performance ranged from 6% correct, slightly more than chance performance, to 73% correct. For the Iowa medial vowels, performance ranged from 21% to 98% correct. For the CNC test (open-set words), performance ranged from 0% correct to 70% correct. Percentage of correct responses for the individual CI users for the consonants, vowels, CNC test, and predictions from NAM are listed in Table 4. We performed a multiple linear regression analysis for three of the variables of interest in this study: the percentage of correct responses on the open-set CNC test as a function of the percentage of correct responses on the consonant and vowel perception tasks. The relation between the variables is illustrated in the following equation

$$\text{CNC} = -16.6 + .67 * P(\text{C})_{\text{con}} + 0.40 * P(\text{C})_{\text{vow}}; R^2 = 0.89, F_{2,23} = 90.9, p < 0.001$$

where CNC is the percentage of correct CNC words obtained, $P(\text{C})_{\text{con}}$ is the percentage of correct consonants obtained, and $P(\text{C})_{\text{vow}}$ is the percentage of correct vowels obtained. A strong linear relation was obtained between the closed-set phoneme and the open-set CNC word scores. However, this relation does not address possible cognitive mechanisms that CI users may be using to recognize words, nor does it address any of the particular confusions made by the CI users.

A strong linear relation was also found between the CNC scores obtained by the individual CI users and the word scores predicted from NAM by calculating $p(\text{ID})$ for the individual

CNC words with Equation 2 by using the confusion matrices for the individual CI users. The data are shown in Figure 1 and the regression equation is shown in the following equation

$$\text{CNC}_{\text{obtained}} = 9.7 + 0.98 * \text{CNC}_{\text{predicted}}; R^2 = 0.82, F_{1,24} = 111.5, p < 0.001$$

The slope of the regression line is near 1.0, and the correlation is very high ($r = 0.91$). This finding suggests that NAM does quite well at predicting the overall correct number of responses for the individual CI users. However, this regression analysis still does not provide detailed insights into the type of errors made by the listeners or the pattern of errors generated by the model. In the current study, we presented each CNC word to the CI users only once and used Equation 2 to calculate $p(\text{ID})$ for the CNC words. Therefore, the subjects only heard one presentation of each CNC word, and NAM generated a probability of a listener identifying the word correctly. This probability may be more meaningful if it were compared with multiple presentations of each CNC to the individual CI users to determine whether the model recognizes words in the same percentages as the individual listeners. We hope to undertake this experiment in future research endeavors.

Information transmission analysis

Information transmission was calculated for several phoneme features of the consonants (voicing, manner, and place) and vowels (tense, height, and place) of the CNC words. Computations were carried out for both the responses from the CI users and the responses generated by NAM. Results for the initial consonants and medial vowels of the CNC words are discussed. IT for the final consonants was very similar to the results obtained for the initial consonants and will not be discussed. Relations between consonant feature IT generated by NAM and the consonant feature IT obtained by the CI users are shown in Figure 2. NAM does fairly well in obtaining levels of consonant feature information similar to the individual CI users. Many of the data points lie near the diagonal, and the regression equations have high correlations ($0.74 < r < 0.92$). The intercepts of the regression equations are between 15 and 25, and the slopes range from 0.54 to 0.77. NAM tends to overestimate a listener's ability to perceive voicing cues at the upper end of the performance scale and underestimate voicing perception at the low and middle portions of the scale. NAM does better at generating similar levels of IT for consonant manner and place of articulation for these subjects.

Relations between vowel feature IT generated by NAM and the vowel feature IT obtained from the CI users are shown in Figure 3. NAM also does quite well in obtaining levels of vowel feature information similar to the individual CI users. For the medial vowels, most of the data points lie near the diagonal, and the regression equations have very high correlations ($0.93 < r < 0.95$). The intercepts of the regression equations are near 0.0 (−4.1–9.4), and the slopes are fairly close to 1.0 (0.74–0.83). NAM slightly underestimates the CI users' abilities to perceive vowel tense cues across the entire performance range. NAM generates similar levels of IT for vowel height and place of articulation.

DISCUSSION

The present analysis was carried out to simulate the process of lexical access in spoken-word recognition using closed-set phoneme perception abilities as predictors. Lexical access and the relative frequency of occurrence of a stimulus word when compared with words in its neighborhood are undoubtedly important to word recognition by CI users (30). These variables were incorporated into NAM, which provides a good fit between the predicted and obtained word scores.

IT analysis was also performed. NAM not only predicted overall performance well but also provided levels of IT similar to those obtained from actual CI users. Overall, NAM made more accurate predictions of subjects' responses to medial vowels than initial consonants. This may be explained in part because closed-set medial vowel confusion matrices were used to predict open-set medial vowels, whereas closed-set medial consonant confusion matrices were used to predict open-set initial and final consonants. The acoustic cues associated with different consonants are strongly dependent on word position, initial, medial, or final (31). Thus, examining confusions for initial and final consonants in a closed-set task may improve our ability to predict and generate responses in an open-set CNC task. Also, Loizou et al. (32) have recently shown that surrounding phonemes influence the perception of the target phonemes by CI users. Expanding the closed-set task to include different phonetic environments should improve the modeling predictions.

At the very least, these results strongly suggest that a purely bottom-up account of word recognition by CI users is not easily tenable. Instead, it seems that the listener's knowledge about lexical properties of words (such as word frequency) is actually used to enhance word recognition performance. This enhancement is particularly important in the case of CI users because of the limitations in their auditory input. Furthermore, the present results suggest that the large individual differences in clinical outcomes of cochlear implantation may be attributable to a variety of factors, including not only peripheral ones such as differences in neural survival or in the ability to discriminate stimulation channels but also more central factors that are related to the cognitive mechanisms used to process sensory information.

Beyond their intrinsic scientific interest, mathematical models such as the one examined here may be used as clinical tools. For example, instead of simply measuring improvements in performance over time, we should take great strides to understand how children with CIs develop language with their CIs. One published opinion has even stated that "not a single case has been reported in the literature of a child acquiring language because of an implant" (33). Svirsky et al. (34) have recently shown that the rate of language development by children with CIs was significantly greater than language development expected for deaf children who do not receive implants. They concluded that CIs provide the better performers among children with CIs with enough information to be developing oral linguistic skills based on the input from their CIs. At this time, many of the clinical tools used to assess word recognition and phoneme perception in adults are difficult or impossible to use in children, partly because of their incomplete language development. Further test and model development for use with children may help researchers determine how these children are forming their mental lexicons.

CONCLUSION

The current study suggests that adult CI users with postlingual profound SNHL use word frequency of occurrence and word similarity information to identify spoken words in a manner that is fundamentally similar to the way listeners with normal hearing recognize spoken words. Continued development of these models should lead to further insights into the neural processing and cognitive mechanisms involved in word recognition and speech perception with a CI.

Acknowledgments

The authors thank Heidi S. Neuburger, M.S., C.C.C.-A., for assistance with data collection.

This work was supported by American Academy of Otolaryngology–Head and Neck Surgery, National Organization for Hearing Research (T.A.M.); Deafness Research Foundation, National Institute on Deafness and

Other Communication Disorders grant R01-DC03937 (M.A.S.); and National Institute on Deafness and Other Communication Disorders grant T32-DC00012 (D.B.P.).

REFERENCES

1. Brown CJ, Abbas PJ, Bertschy M, et al. Longitudinal assessment of physiological and psychophysical measures in cochlear implant users. *Ear Hear.* 1995; 16:439–449. [PubMed: 8654899]
2. Skinner MW, Holden LK, Holden TA, et al. Performance of post-linguistically deaf adults with the Wearable Speech Processor (WSP III) and Mini Speech Processor (MSP) of the Nucleus Multi-Electrode Cochlear Implant. *Ear Hear.* 1991; 12:3–22. [PubMed: 2026285]
3. Skinner MW, Clark GM, Whitford LA, et al. Evaluation of a new spectral peak (SPEAK) coding strategy for the Nucleus 22 channel cochlear implant system. *Am J Otol.* 1994; 15:15–27. [PubMed: 8572106]
4. Tyler RS, Moore BC, Kuk FK. Performance of some of the better cochlear-implant patients. *J Speech Hear Res.* 1989; 32:887–911. [PubMed: 2601319]
5. Tyler RS, Parkinson AJ, Woodworth GG, Lowder MW, Gantz BJ. Performance over time of adult patients using the Ineraid or Nucleus cochlear implant. *J Acoust Soc Am.* 1997; 102:508–522. [PubMed: 9228814]
6. Blamey PJ, Pyman BC, Gordon M, et al. Factors predicting post-operative sentence scores in postlingually deaf adult cochlear implant patients. *Ann Otol Rhinol Laryngol.* 1992; 101:342–348. [PubMed: 1562140]
7. Gantz BJ, Woodworth GG, Knutson JF, Abbas PJ, Tyler RS. Multivariate predictors of audiological success with multichannel cochlear implants. *Ann Otol Rhinol Laryngol.* 1993; 102:909–916. [PubMed: 8285510]
8. Blamey P, Arndt P, Bergeron F, et al. Factors affecting auditory performance of postlinguistically deaf adults using cochlear implants. *Audiol Neurotol.* 1996; 1:293–306. [PubMed: 9390810]
9. Rubinstein JT, Parkinson WS, Tyler RS, Gantz BJ. Residual speech recognition and cochlear implant performance: effects of implantation criteria. *Am J Otol.* 1999; 20:445–452. [PubMed: 10431885]
10. Bilger RC, Wang MD. Consonant confusions in patients with sensorineural hearing loss. *J Speech Hear Res.* 1976; 19:718–748. [PubMed: 1003952]
11. Wang MD, Bilger RC. Consonant confusions in noise: a study of perceptual features. *J Acoust Soc Am.* 1973; 54:1248–1266. [PubMed: 4765809]
12. Boothroyd A, Nittrouer S. Mathematical treatment of context effects in phoneme and word recognition. *J Acoust Soc Am.* 1988; 84:101–114. [PubMed: 3411038]
13. Nittrouer S, Boothroyd A. Context effects in phoneme and word recognition by young children and older adults. *J Acoust Soc Am.* 1990; 87:2705–2715. [PubMed: 2373804]
14. Rabinowitz WM, Eddington DK, Delhorne LA, Cuneo PA. Relations among different measures of speech reception in subjects using a cochlear implant. *J Acoust Soc Am.* 1993; 92:1869–1881. [PubMed: 1401531]
15. McClelland JL, Elman JL. The TRACE model of speech perception. *Cognit Psychol.* 1986; 18:1–86. [PubMed: 3753912]
16. Luce, PA. Technical Report No. 6. Bloomington, IN: Speech Research Laboratory, Indiana University; 1986. Neighborhoods of words in the mental lexicon.
17. Luce PA, Pisoni DB. Recognizing spoken words: the neighborhood activation model. *Ear Hear.* 1998; 19:1–36. [PubMed: 9504270]
18. Kucera, F.; Francis, W. Computational Analysis of Present-Day American English. Providence, RI: Brown University Press; 1967.
19. Charles-Luce, J.; Luce, PA.; Cluff, MS. Retroactive influence of syllable neighborhoods. In: Altmann, GTM., editor. *Cognitive Models of Speech Processing: Psycholinguistic and Computational Perspectives.* Cambridge: MIT Press; 1990. p. 173–184.
20. Cluff MS, Luce PA. Similarity neighborhoods of spoken two-syllable words: retroactive effects on multiple activation. *J Exp Psychol Hum Percept Perform.* 1990; 3:551–563. [PubMed: 2144570]

21. Meyer TA, Pisoni DB. Some computational analyses of the PBK test: effects of word frequency and lexical density on spoken word recognition. *Ear Hear.* 1999; 20:363–371. [PubMed: 10466571]
22. Luce PA, Goldinger SD, Auer ET, Vitevitch MS. Phonetic priming, neighborhood activation, and PARSYN. *Percept Psychophys.* 2000; 62:615–625. [PubMed: 10909252]
23. Vitevitch MS, Luce PA, Pisoni DB, Auer ET. Phonotactics, neighborhood activation, and lexical access for spoken words. *Brain Lang.* 1999; 68:306–311. [PubMed: 10433774]
24. Wilson BS, Finley CC, Lawson DT, Wolford RD, Eddington DK, Rabinowitz WM. Better speech recognition with cochlear implants. *Nature.* 1991; 352:236–238. [PubMed: 1857418]
25. Tyler, RS.; Preece, JP.; Lowder, MW. Laser Videodisc and Laboratory Report. Iowa City: Department of Otolaryngology–Head and Neck Surgery, University of Iowa; 1987. The Iowa audiovisual speech perception laser video disc.
26. Nusbaum, HC.; Pisoni, DB.; Davis, CK. Progress Report No. 10. Bloomington, IN: Speech Research Laboratory, Indiana University; 1984. Sizing up the Hoosier mental lexicon: measuring the familiarity of 20,000 words.
27. Peterson G, Lehiste I. Revised CNC lists for auditory tests. *J Speech Hear Dis.* 1962; 27:62–70.
28. Frisch SA, Pisoni DB. Modeling spoken word recognition performance by pediatric cochlear implant users using feature identification. *Ear Hear.* 2000; 21:578–589. [PubMed: 11132784]
29. Miller GA, Nicely PE. An analysis of perceptual confusions among some English consonants. *J Acoust Soc Am.* 1955; 27:301–315.
30. Kirk KI, Pisoni DB, Osberger MJ. Lexical effects on spoken word recognition by pediatric cochlear implant users. *Ear Hear.* 1995; 16:470–481. [PubMed: 8654902]
31. Stevens, KN. *Acoustic Phonetics.* Cambridge: MIT Press; 1998.
32. Loizou PC, Poroy O, Dorman M. The effect of parametric variations of cochlear implant processors on speech understanding. *J Acoust Soc Am.* 2000; 108:790–802. [PubMed: 10955646]
33. Lane H, Bahan B. Ethics of cochlear implantation in young children: a review and reply from a DEAF-WORLD perspective. *Otolaryngol Head Neck Surg.* 1990; 119:297–313. [PubMed: 9781982]
34. Svirsky MA, Robbins AM, Kirk KI, Pisoni DB, Miyamoto RT. Language development in profoundly deaf children with cochlear implants. *Psychol Sci.* 2000; 11:153–158. [PubMed: 11273423]

FIG. 1.

Plot of CNC scores. Obtained scores (y axis) and scores generated with NAM (x axis).

Dashed line along the diagonal (slope = 1.0) represents identical observed and generated scores. Regression equation shown in the figure.

FIG. 2.

Plots of information transmission for the consonant features, voicing, manner, and place of articulation of the initial consonants in the CNC words. Obtained scores (*y* axes) and scores generated by NAM (*x* axes). *Dashed lines along the diagonal* (slope = 1.0) represent identical observed and generated scores. Regression equations shown in the plots. This analysis provides detailed insight into the type of errors made by the listeners and the pattern of errors generated by the model.

FIG. 3.

Plots of information transmission for the vowel features, height, tense, and place of articulation for the medial vowels of the CNC words. Obtained scores (y axes) and scores generated by NAM (x axes). *Dashed lines along the diagonal* (slope = 1.0) represent identical observed and generated scores. Regression equations shown in the plots.

TABLE 1

Vowel feature codes

		Height	Tense	Place
@	bat	1	1	1
c	bought	2	2	2
E	bet	2	1	1
R	bird	3	2	3
i	beet	3	2	1
I	bit	3	1	1
U	look	3	1	2
^	cut	2	2	3
u	tooth	3	2	2
e	fate	2	2	1
a	lock	1	2	2
o	note	2	2	2
Y	fire	4	3	4
W	loud	4	3	2
O	coil	5	3	4

Height: 1, low; 2, mid; 3, high; 4, low-rising; 5, mid-rising.

Tense: 1, lax; 2, tense; 3, diphthong.

Place: 1, front; 2, back; 3, mid; 4, back-fronting.

TABLE 2

Example of a vowel confusion matrix for subject 4: 9 medial vowels, 15 repetitions^a

	@	c	E	R	i	I	U	^	u	e	a	o	Y	W	O
@	0.594	0.000	0.000	0.264	0.000	0.066	0.066	0.000	0.000	0.000	0.010	0.000	0.000	0.000	0.000
c	0.000	0.631	0.000	0.057	0.000	0.000	0.115	0.000	0.000	0.000	0.063	0.073	0.000	0.061	0.000
E	0.000	0.000	0.618	0.000	0.000	0.309	0.000	0.062	0.000	0.012	0.000	0.000	0.000	0.000	0.000
R	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
i	0.000	0.000	0.000	0.000	0.847	0.000	0.000	0.000	0.000	0.153	0.000	0.000	0.000	0.000	0.000
I	0.000	0.000	0.000	0.000	0.133	0.867	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
U	0.000	0.000	0.000	0.000	0.000	0.000	0.933	0.000	0.067	0.000	0.000	0.000	0.000	0.000	0.000
^	0.000	0.000	0.000	0.000	0.000	0.000	0.600	0.400	0.000	0.000	0.000	0.000	0.000	0.000	0.000
u	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.813	0.000	0.063	0.063	0.000	0.061	0.000
e	0.000	0.000	0.012	0.000	0.154	0.000	0.000	0.000	0.000	0.819	0.000	0.015	0.000	0.000	0.000
a	0.010	0.063	0.000	0.000	0.000	0.000	0.000	0.000	0.063	0.000	0.738	0.063	0.000	0.062	0.000
o	0.000	0.073	0.000	0.000	0.000	0.000	0.000	0.000	0.063	0.015	0.063	0.724	0.000	0.062	0.000
Y	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.973	0.018	0.009
W	0.000	0.061	0.000	0.000	0.000	0.000	0.000	0.000	0.061	0.000	0.061	0.061	0.018	0.733	0.004
O	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.009	0.004	0.987

^aMatrix expanded to 15 on the basis of feature information categories shown in Table 1. Responses to the original nine vowels (@, c, E, R, i, I, U, ^, and u) are above and left of the lines. Estimated responses to the remaining six vowels (e, a, o, Y, W, and O) are to the right and below the lines. Rows, stimuli; columns, responses.

TABLE 3

Consonant feature codes

		Voicing	Manner	Place
b	<u>b</u> ook	1	1	1
C	<u>c</u> hoice	0	6	4
d	<u>d</u> og	1	1	2
J	<u>j</u> ump	1	6	4
f	<u>f</u> ar	0	2	1
g	<u>g</u> oat	1	1	3
h	<u>h</u> elp	0	2	5
k	<u>k</u> ick	0	1	3
l	<u>l</u> augh	1	4	2
m	<u>m</u> an	1	3	1
n	<u>n</u> ice	1	3	2
G	<u>r</u> ing	1	3	3
p	<u>p</u> et	0	1	1
r	<u>r</u> ed	1	7	2
s	<u>s</u> ip	0	2	2
S	<u>sh</u> ip	0	2	4
t	<u>t</u> ell	0	1	2
T	<u>th</u> in	0	2	2
D	<u>th</u> em	1	2	2
v	<u>v</u> an	1	2	1
w	<u>w</u> et	1	5	1
y	<u>y</u> awn	1	5	4
z	<u>z</u> ip	1	2	2
Z	<u>a</u> zure	1	2	4

Voicing: 0, unvoiced; 1, voiced.

Manner: 1, stop; 2, fricative; 3, nasal; 4, lateral; 5, glide; 6, affricate; 7, rhotic.

Place: 1, labial; 2, alveolar; 3, velar; 4, alveopalatal; 5, laryngeal.

TABLE 4

Percent correct obtained on the closed-set medial consonants, closed-set medial vowels, and open-set CNC words (percent correct word score generated by NAM)

Subject	P(C) consonants	P(C) vowels	P(C) CNC words	P(C) CNC words simulated-NAM
1	17.40	35.60	4.00	5.50
2	53.70	72.60	62.70	39.80
3	77.40	98.50	68.00	72.90
4	75.50	78.50	66.00	71.50
5	54.00	94.00	68.00	31.90
6	31.00	73.30	24.70	13.80
7	22.80	33.30	14.00	5.60
8	46.80	80.00	56.00	37.10
9	63.00	70.10	36.00	35.50
10	70.80	96.30	58.00	57.60
11	27.80	30.40	14.00	7.90
12	17.10	51.10	24.00	10.30
13	7.60	21.10	0.00	1.20
14	38.90	44.40	17.30	17.00
15	59.30	82.20	62.70	45.90
16	22.20	53.30	10.00	11.00
17	28.20	53.30	17.30	8.90
18	69.00	90.40	70.70	57.70
19	41.00	56.30	41.00	23.90
20	56.00	83.70	58.70	38.30
21	12.50	80.00	12.00	6.10
22	49.50	71.10	46.00	32.40
23	6.00	26.10	0.70	0.10
24	27.30	63.00	35.30	15.50
25	51.40	88.10	53.30	36.80
26	73.10	91.90	66.70	68.50