RESEARCH ON SPOKEN LANGUAGE PROCESSING
Progress Report No. 22 (1998)
Indiana University

The Development of the Lexicon in Deaf Children with Cochlear Implants: New Insights into Language Acquisition[1]

Stefan Frisch,[2] and David B. Pisoni

Speech Research Laboratory
Department of Psychology
Indiana University
Bloomington, Indiana 47405

---

[1] This work was supported by NIH-NIDCD Research Grants DC00111, DC0004, DC00423 and Training Grant DC00012 to Indiana University. We would like to thank Steve Chin, Terri Kerr, Karen Kirk, Ted Meyer and Mario Svirsky for their assistance and suggestions on this work.

[2] Program in Linguistics, University of Michigan, Ann Arbor, MI.

The Development of the Lexicon in Deaf Children with Cochlear Implants: New Insights into Language Acquisition

Introduction

A cochlear implant is a surgically inserted electronic device that directly stimulates the auditory nerve based on input from an externally worn microphone and speech processor. Children who have a profound hearing loss from a very young, pre-linguistic age, and who acquire spoken language through a cochlear implant, provide a unique opportunity to study the sensory, perceptual, and cognitive bases of language acquisition. While the electrical signal received through a cochlear implant is very different from speech perceived with normal hearing, we hypothesize that the same abstract units of language and lexical organization are nonetheless part of the linguistic knowledge of children who learn spoken language through cochlear implants.

It is now well established in the field of pediatric cochlear implants that some children who learn language through a cochlear implant do exceptionally well on standardized tests of speech perception and spoken word recognition compared to other implanted children (e.g., Staller, Pelter, Brimacombe, Mecklenberg, & Arndt, 1991). A study to identify the primary factors that underlie the exceptionally good performance of these children, the so-called “Stars”, was recently carried out (Pisoni, Svirsky, Kirk, & Miyamoto, 1997). Using an “extreme groups” design, speech perception, intelligibility, and language scores were examined for two samples of pediatric cochlear implant users. The first group, the “Stars”, were children who scored in the top 20% on the Phonetically Balanced Kindergarten (PBK) test (Haskins, 1949), an open-set spoken word recognition test given to all children who use cochlear implants followed by the DeVault Otologic Research Center in the Indiana University School of Medicine. The second group, the “Control” group, scored in the bottom 20% on the PBK.

The “Stars” scored very well not only on the criterial PBK word recognition test, but also on a number of other speech perception and language tests. Most notable were the unusually high correlations among several different measures of spoken word recognition and scores on the Reynell receptive and expressive language scales (Reynell & Gruber, 1990), suggesting that one common underlying factor in these children may be the acquisition of spoken language, specifically, the development of an acoustic-phonetic lexicon which serves as the “interface” between the initial sensory input and the representation of the sound patterns of words in lexical memory. Spoken word recognition performance was also highly correlated with speech intelligibility and open-set measures of language comprehension that required children to interpret spoken language in meaningful ways.

A second study of the same population of “Stars” and “Controls” used software simulations of spoken word recognition based on behavioral data (Frisch & Pisoni, 1997). These simulations applied hypotheses about the process of spoken word recognition in normal hearing adults to the feature and word identification performance of both the “Stars” and “Controls”. These simulations showed that, while some feature distinctions are less adequately transmitted through a cochlear implant (e.g. place of articulation), children with cochlear implants can utilize the available feature and phoneme contrasts to recognize spoken words just like normal hearing adults. For example, children who use cochlear implants recognize words in the context of acoustically similar words stored in long term memory (Kirk, Pisoni, & Osberger, 1995; Luce & Pisoni, 1998). In other words, acoustically similar words are stored in proximity to one another in the phonological lexicon of normal hearing adults and in children who have learned spoken language through a cochlear implant.

Taken together, these analyses of children who are learning spoken language using a cochlear implant suggest that at least some of these children have been able to begin the normal process of spoken language acquisition, despite the degraded signal made available to them through their devices. The patterns of intercorrelation among several measures of speech perception, intelligibility, and language development suggest that these children are getting sufficient sound input through their cochlear implants to map the sound patterns of words onto meanings and build a lexicon of words. The functional organization of this lexicon appears to be analogous to that of adults with normal hearing. The development of a mental lexicon is a necessary prerequisite for constructing a grammar of the target language from the ambient language spoken in the environment. These studies demonstrate that the relevant units of language and lexical organization must be, to some degree, abstractions of the physical signal transmitted to the brain by the peripheral auditory system.

Descriptive Data

The studies we review used an extreme groups design. Two groups of pediatric cochlear implant users, the "Stars" and "Controls," were selected from a longitudinal study of 160 children with profound hearing loss who use cochlear implants followed by the DeVault Otologic Research Center in the Indiana University School of Medicine. Subjects in both groups were prelingually deafened, and they received cochlear implants because they derived no benefit from conventional hearing aids. The groups were selected based on their performance in a test of open-set word recognition, the PBK (Haskins, 1949). This is an extremely difficult test for pediatric cochlear implant users, and we show below that good performance on this test is indicative of well-developed speech and language skills overall. The "Stars" had performance in the top 20% of scores and the "Controls" had performance in the bottom 20% of scores among children in the longitudinal study at 2 years post-implant.

Table 1

Demographic information for the populations studied.

	Stars (N = 27)	Controls (N = 23)
Mean Age at Onset	0.3	0.8
Mean Age at Implantation	5.8	4.4
Length of Deprivation	5.5	3.6
Mean Chronological Age	6.7	5.4
Mean Length of Implant Use	0.9	1.0
Communication Mode:		
Oral Communication	N = 19	N = 8
Total Communication	N = 8	N = 15
Mean Non-verbal IQ	107	101

Demographic information for these two groups is shown in Table 1 (taken from Pisoni et al., 1997). These groups differ on chronological age, age of implantation, length of deprivation, and on the communication mode of their instructional environment. These factors have been shown to be related to the speech perception performance of prelingually deafened children (Osberger, Robbins, Todd, Riley, Kirk, & Carney, 1996; Fryauf-Bertschy, Tyler, Kelsay, Gantz, & Woodworth, 1997), though the differences between the “Stars” and “Controls” make conflicting predictions. It is not clear from the demographic information that one group should be superior in spoken word recognition. The “Stars” and “Controls” do not differ in non-verbal IQ. Data presented below are based on clinical assessment for these children at 1 year post-implant, one year *before* the assessment of PBK performance that was used to define the groups.

Children who learn spoken language using a cochlear implant are a particularly important and interesting population to study. For these children, the target language to be acquired is the same as for children without hearing loss. In particular, they must learn to produce the gestures, segments, and words of the target language in order to communicate. However, their perceptual input is very different. The electrical stimulation of the auditory nerve provided by the cochlear implant is unlike aural hearing. To the extent that children using cochlear implants and children without hearing loss can acquire the same linguistic knowledge and underlying grammatical system, we will have evidence that some aspects of the linguistic system are independent of phonetic detail. In other words, the linguistic system is sufficiently robust to be acquired despite the degraded and incomplete input provided by the implant. If we can find systematic differences between the linguistic knowledge of children who use cochlear implants and children without hearing loss, then we will have evidence that other aspects of the linguistic system are crucially dependent on audition. Thus, the pediatric cochlear implant population is an important new source of data relevant to the status of grammar as innately preprogrammed or learned entirely from exposure to language (e.g., Prince & Smolensky, 1997; Seidenberg, 1997).

Correlations

Pisoni et al. (1997) studied the pattern of intercorrelation between scores on a variety of clinical assessments of speech and language ability. In this section, we first review the behavioral tests used to assess speech and language ability at the DeVault Otologic Research Center. We then present behavioral data for the “Stars” and “Controls” and correlations among the scores on these tests within each group.

Assessment Tools

Scores on a variety of behavioral tests are collected for each child in the longitudinal study at regular intervals. These scores are used to assess development in speech perception, spoken word recognition, language development, and speech intelligibility.

Speech Perception Tests. Speech perception is assessed through two tests. The first, the Minimal Pairs Test (MPT; Robbins, Renshaw, Miyamoto, Osberger, & Pope, 1998) assesses feature identification using a two-choice picture pointing task. Words on the test differ on a single dimension of phonological contrast (e.g., *bear* and *pear* differ only on voicing) and are presented using a mesh screen so that the child must rely on audition alone. The second test, the Common Phrases Test (CPT; Osberger et al., 1991), assesses spoken language comprehension. In this test, the child must answer simple questions or obey commands. This test is given in three different modality conditions: auditory only (face covered), visual only (lip reading), and auditory plus visual.

Spoken Word Recognition Tests. Word recognition is assessed using the PBK test (Haskins, 1949), and a second test known as the Lexical Neighborhoods Test (LNT; Kirk et al., 1995). These open-set tests are administered auditorily. Words on the PBK test contain a representative selection of English phonemes based on frequencies in running speech. The LNT contains words that were chosen to differ on the degree to which they are phonetically confusable with other words. The “Easy” words are words that have very few “neighbors”, words that contain similar phonemes. The “Hard” words have many other words to which they are phonemically similar. The LNT contains monosyllabic and multisyllablic words.

Speech Intelligibility Tests. The intelligibility of speech produced by the children is assessed using either the Beginner's Intelligibility Test (BIT; Osberger, Robbins, Todd, & Riley, 1994) or the Monsen sentences (Monsen, 1983), depending on the abilities of the children. Both tests require the child to repeat a sentence based on the clinician's spoken model, so the tests involve imitation and repetition rather than spontaneous sentence generation. These utterances are transcribed by naive adult listeners who are unfamiliar with deaf speech, and intelligibility is scored based on the number of words correctly transcribed.

Vocabulary Test. The Peabody Picture Vocabulary Test (PPVT; Dunn, 1965) is used to assess vocabulary knowledge. This test requires the child to select the picture corresponding to a word presented both auditorily and visually (either written, or in sign, depending on the child's age and preferred mode of communication).

Language Tests. Language skills are assessed using the Reynell Developmental Language Scales (Reynell & Gruber, 1990). In this test, the child manipulates and describes objects based on questions and instructions of varied length and grammatical complexity given by the clinician. Performance on this test is compared to normative data to obtain both receptive and expressive language quotients. Children who know sign are presented with simultaneous oral and signed instructions, and they can respond either orally or with sign.

Behavioral Data

The performance of the "Stars" and "Controls" on the battery of speech and language tests reveals consistent differences between the two groups of subjects on tests which depend on perception and production of spoken language (Pisoni et al., 1997). These differences suggest that, after 1 year of implant use, the "Stars" have made significant progress in developing the core phonological abilities and skills necessary for speech perception and production. We propose that these abilities hinge upon the development of an acoustic-phonetic mental lexicon which links perceptual, articulatory, and semantic information about lexical items in long-term memory.

Figures 1 and 2 show performance on the speech perception tests: the Minimal Pairs Test and Common Phrases Test. For the Minimal Pairs Test, chance performance is 50%. The "Stars" perform significantly above chance levels in identifying Manner, Vowel Place, and Vowel Height. Their performance on Voicing and Place contrasts is not statistically significant, though a similar trend is present. For the "Controls", none of the contrasts are identified at levels significantly above chance. The "Stars" also outperform the "Controls" on the Common Phrases Test, indicating that their performance/ability with spoken language is not restricted to the perception of isolated words. Since the performance of the "Stars" in visual only (lipreading) and audio-visual conditions is clearly above that of the "Controls", we can conclude that the "Stars" have also developed an awareness of the articulatory-acoustic mapping. Not only are the "Stars" better at perceiving the acoustic events corresponding to words, but also this information is encoded along with the articulatory correlates of those acoustic events in both the auditory and visual domains.

Insert Figures 1 and 2 about here

Figures 3 and 4 show performance on the tests of open-set spoken word recognition: the PBK and the LNT. Performance on the LNT is broken down into performance on monosyllabic words and multisyllablic words (MLNT), and for each of these groups by “Easy” and “Hard” words. Figure 3 shows performance scored by words correct, and Figure 4 shows performance scored by phonemes correct.

Both figures show that performance on these tests at 1 year post-implant is much better for the “Stars” than the “Controls”. The “Controls” were unable to recognize any words on these tests, and due to their very poor performance, “Hard” word lists were not even administered for any of these children. Note that performance for the “Stars” is better for “Easy” words than “Hard” words, and also better for longer (multisyllablic) words than shorter words. This is the pattern found for normal hearing adults (e.g., Luce & Pisoni, 1998).

Insert Figures 3 and 4 about here

Figure 5 shows speech intelligibility scores for the “Stars” and “Controls”. This figure shows that the superior performance of the “Stars” on speech-related tasks is not limited to speech perception and receptive aspects of language. The “Stars” show evidence of perception to production transfer of their linguistic knowledge. Once again, this suggests the presence of an abstract system, a mental lexicon, which unites perceptual and articulatory representations of spoken words.

Insert Figure 5 about here

Figures 6 and 7 show PPVT vocabulary scores and Reynell developmental language scores for the “Stars” and “Controls”. Here, we see no difference between the “Stars” and “Controls”. These tests show that both the “Stars” and “Controls” have acquired language to some extent. However, these tests are not specific to *spoken* language. For children who are learning a signed language in a Total Communication environment, sign is used in addition to spoken language on these tests. The difference between the “Stars” and “Controls” is thus a difference in oral/aural language ability and not some more fundamental developmental difference in linguistic ability or overall mental capacity.

Insert Figures 6 and 7 about here

If we examine performance on the PPVT and Reynell tests more closely, differences unrelated to the distinction between the “Stars” and “Controls” (which is based on performance on the PBK test at 2 years post-implant) can be found. First, there is a considerable effect of communication mode on performance on the PPVT. For the “Stars”, the children in Oral Communication programs average 0.54 language quotient, while children in Total Communication programs average 0.75. For the “Controls” the children in Oral Communication programs average 0.55, while the children in Total Communication programs average 0.76. Thus, it appears that the children in Total Communication programs either have a better developed vocabulary, or they are more effective in making their knowledge known by using a combination of speech and sign. Children in Oral Communication programs are given this test using a combination of speech and written materials. For the Reynell developmental language tests, the pattern is more complex. There is no significant effect of group (“Stars” vs. “Controls”) or communication mode (Oral vs. TC) on expressive language ability. For receptive language ability, there is no effect of communication mode for the “Stars” (0.53 for Oral, 0.59 for TC) but there is a significant difference for the “Controls” (0.36 for Oral, 0.65 for TC). Again, these differences may reflect the way these language tests are administered to the children using their preferred mode of communication.

The behavioral data show that the “Stars” and “Controls” are very similar along several important dimensions of linguistic ability; in particular, vocabulary and overall receptive and expressive communication skills. However, on tests which focus on the ability to perform tasks targeting spoken language ability, the “Stars” are considerably more adept than the “Controls”. In perception, the “Stars” have better performance in speech feature discrimination, spoken word recognition, and sentence comprehension. The “Stars” also produce more intelligible speech than the controls, as transcribed by untrained listeners. These overall patterns suggest the “Stars” have taken steps toward acquiring a spoken language that the “Controls” have not.

Based on these behavioral differences, it appears that successful cochlear implant users, like the “Stars”, have developed an acoustic-phonetic lexicon which links the perceptual, articulatory, and semantic representations of lexical items that is structured much like that of adults with no hearing loss (Frisch & Pisoni, 1997; Kirk et al., 1995; Pisoni et al., 1997). In the following section, we discuss the correlation analysis of Pisoni et al. (1997), that examined individual differences to test this hypothesis more carefully.

Correlations Among Behavioral Measures

The group data presented in the previous section suggest that a variety of spoken language tasks draw on a common underlying mechanism: the mental lexicon. Stronger confirmation of this hypothesis can be found by examining individual differences in performance among the “Stars” and “Controls”. Not only are the “Stars” better than the “Controls” at spoken language tasks, but individual differences in performance by the “Stars” across these tests are correlated. No such correlations are found for the “Controls”, indicating that the vocabulary and language skills that they do have may be unrelated to their facility with spoken language. However, we should also caution that the absence of correlations for the “Controls” could also be due to floor effects and low variances in the observed measures.

Tables 2 and 3 show correlations among test scores within each group, the “Stars” and “Controls”, reported by Pisoni et al. (1997). A (-) indicates that too few data points were available for the particular test pairs for a correlation to be made. This was frequently the case for the “Controls” where the LNT word recognition tests were often not administered due to the child’s inability to perform the task at all. Correlations that are significant at the 0.01 level are indicated by boldface.

There are a number of significant positive correlations among scores for the “Stars” in Table 2, and many of the scores which are not significant are also large positive correlations. These correlations in individual variation among children in the “Stars” group provide converging evidence that spoken word recognition, speech perception, and speech production abilities are all linked by a common underlying mechanism. The scores of the “Stars” on spoken language tasks are also strongly correlated with vocabulary knowledge and language ability, indicating that the spoken language skills of these children provide the foundation of their linguistic competence.

Table 2

Correlations among speech and language measures within the "Stars" group. Statistically significant correlations ($p < 0.01$) are bolded.

		Minimal Pairs			CPT			LNT		MLNT		PPVT	Reynell	
		Man	Voi	Pl	Aud	Vis	A/V	Easy	Hard	Easy	Hard		Rec	Exp
CPT	Aud	**0.58**	0.38	0.04										
	Vis	**0.80**	0.51	0.18										
	A/V	**0.86**	0.53	0.16										
LNT	Easy	0.34	0.20	0.16	**0.81**	0.41	0.42							
	Hard	0.51	0.58	-0.06	**0.85**	0.57	0.56							
MLNT	Easy	0.53	0.34	0.06	**0.83**	0.60	0.56							
	Hard	0.38	0.33	-0.11	0.70	0.62	0.33							
PPVT		**0.56**	0.46	0.07	**0.69**	0.40	0.50	0.62	0.63	0.57	0.35			
Reynell	Rec	**0.77**	0.69	0.20	**0.82**	0.64	0.64	**0.86**	**0.81**	**0.84**	0.66	**0.81**		
	Exp	**0.78**	0.61	0.31	**0.85**	**0.79**	0.67	**0.83**	**0.82**	**0.87**	**0.76**	**0.68**		
Intell		0.55	0.53	0.41	0.65	**0.87**	0.43	**0.89**	**0.80**	**0.87**	0.72	0.45	**0.80**	**0.85**

Table 3

Correlations among speech and language measures within the "Controls" group. Statistically significant correlations ($p < 0.01$) are bolded. A (-) indicates an insufficient amount of data for a correlation.

		Minimal Pairs			CPT			LNT		MLNT		PPV	Reynell	
		Man	Voi	Pl	Aud	Vis	A/V	Easy	Hard	Easy	Hard		Rec	Exp
CPT	Aud	-0.51	-0.05	0.46										
	Vis	-0.19	-0.70	-0.19										
	A/V	-0.38	-0.46	-0.03										
LNT	Easy	-	-	-	-	-	-							
	Hard	-	-	-	-	-	-							
MLNT	Easy	-	-	-	-	-	-							
	Hard	-	-	-	-	-	-							
PPVT		0.20	0.23	-0.25	-0.46	-0.47	0.04	-	-	-	-			
Reynell	Rec	0.08	-0.63	0.01	-	-	0.33	-	-	-	-	**0.69**		
	Exp	-0.28	-0.49	0.33	-	-	0.36	-	-	-	-	0.56		
Intell		0.19	-0.11	-0.09	0.04	0.25	0.07	-	-	-	-	-0.01	-0.39	-0.13

By comparison, correlations among scores for the "Controls" in Table 3 are much smaller and inconsistent: some are positive and some are negative. The only statistically significant correlation, between Reynell receptive language quotient and PPVT vocabulary score, reflects the differences in communication mode discussed above. Recall that the "Controls" in Oral Communication programs had much lower receptive language quotients than the “Controls” in Total Communication programs. Also, the children in TC programs had better vocabulary scores than the children in Oral programs. There is too little data to examine correlations independently by communication mode, so it is not clear whether this correlation is anything other than a reflection of the group trend. Overall then, the correlation analysis provides little evidence for any common mechanism underlying the performance of the “Controls” on tests of speech and language ability.

Based on the individual differences, Pisoni et al. (1997) conclude that the “Stars” have made considerable progress in acquiring spoken English. In particular, the “Stars” have developed a spoken language lexicon. Pisoni et al. hypothesize that individual differences in the level of detail in which words are represented in the acoustic-phonetic lexical space underlie differences in performance on tests of speech feature identification, sentence comprehension, and spoken word recognition. These differences also appear to indirectly influence speech intelligibility and measures of overall linguistic development. Finally, these differences are connected to overall vocabulary knowledge, suggesting that the level of lexical development is closely tied to the number of words that are known. Gathercole, Willis, Baddeley, and Emslie (1994) come to a similar conclusion based on individual differences in the abilities of adults and children without hearing loss in a variety of speech tasks.

Lexical Development

The performance of the “Stars” and “Controls” on behavioral tests of spoken language ability and correlations between performance on different tasks points toward the crucial role that the mental lexicon plays in linking perceptual, gestural, and semantic information about words. In particular, it was suggested above that the “Stars” have developed an acoustic-phonetic mental lexicon organized much like the lexicon of adults with normal hearing. In this section, we examine the nature of lexical organization in successful pediatric cochlear implant users. Recently, Frisch and Pisoni (1997) proposed a computational model of spoken word recognition by pediatric cochlear implant users. This model has two important assumptions. First, it is assumed that pediatric CI users’ performance on the Minimal Pairs Test can be used to predict their ability to recognize phonemes in spoken word recognition. Second, it is assumed that spoken word recognition consists of two separate independent stages, phoneme identification and lexical access. As a first test of whether the phonological structure of the lexicon can be used to predict spoken word recognition performance by the “Stars” and “Controls”, they considered a model with *no lexical knowledge*. At 1 year post-implant, a model of spoken word recognition with no lexicon is able to predict the spoken word recognition performance of the “Controls”, but it under-predicts the actual performance of the “Stars”. These simulations provide converging evidence that the lexicon plays an important role in spoken language processing for the “Stars.”

Predicting Phoneme Identification

Frisch and Pisoni (1997) used performance on the Minimal Pairs Test to predict phoneme identification in the task of open-set spoken word recognition. In the Minimal Pairs Test, the possible contrasts among segments in English are grouped into five broad categories: place of articulation, manner of articulation, voicing, vowel place, and vowel height.

In order to make predictions of phoneme confusions from the Minimal Pairs Test, it is necessary to decide explicitly whether a failure to identify some contrast will result in confusion between a particular pair of phonemes. For example, if place of articulation for a particular pair of consonants cannot be discriminated, are the remaining manner and voicing features sufficient to differentiate the consonants? In linguistic theory, this is referred to as “distinctness” (Stanley, 1967). In order to answer this question, we utilize a set-theoretic model of the phonemic inventory, originally developed in Broe (1993), that makes explicit the contrastiveness of phonemic representations using natural classes.

We can use the concept of a natural class to predict phoneme confusions when contrasts are not discriminable. For example, suppose that no vowel height features are discriminable. In a hypothetical five vowel inventory {a, e, i, o, u}, the height features define distinct natural classes, {a}, {e, o}, and {i, u}. The place features define different classes, {e, i} and {a, o, u}. Together, the height and place features individuate the vowels. With combinations of vowel place and vowel height features, each of the vowels is distinct. With no vowel height features, /e/ and /i/ are confusable, and /a/, /o/, and /u/ are confusable. Different confusions arise if no vowel place features are discriminable. In this case, /e/ and /o/ are confusable, and /i/ and /u/ are confusable.

For simplicity, Frisch and Pisoni (1997) assume that when several phonemes are confusable, each is equally likely to be identified as the "perceived" phoneme. A more accurate model might incorporate some sort of weighting among confusions (see Luce & Pisoni, 1998). Assuming equally likely confusions, the probability of any particular confusion is inversely related to the size of the set of possible confusions. Once the confusable phonemes for some dimension is known, the pattern of confusability can be represented using a confusion matrix. The confusion matrix provides a convenient display for the probability of confusions in an entire inventory when a particular contrast is lost.

Tables 4 and 5 show confusion matrices for the five vowel inventory when vowel height or vowel place features are removed. The intended phoneme is given in the left column; the perceived phoneme is given across the top. Each number in the table represents the probability that the intended phoneme is heard as the perceived phoneme. A phoneme is always assumed to be confusable with itself. In other words, there is a possibility that the correct phoneme will be identified by chance even if a contrast is not detectable.

Table 4

Confusion matrix for the five vowel inventory when vowel height features are removed.

		Perceived				
Intended		i	e	a	o	u
	i	0.5	0.5	0	0	0
	e	0.5	0.5	0	0	0
	a	0	0	0.33	0.33	0.33
	o	0	0	0.33	0.33	0.33
	u	0	0	0.33	0.33	0.33

Table 5

Confusion matrix for the five vowel inventory when vowel place features are removed.

	Perceived					
Intended		i	e	a	o	u
	i	0.5	0	0	0	0.5
	e	0	0.5	0	0.5	0
	a	0	0	1	0	0
	o	0	0.5	0	0.5	0
	u	0.5	0	0	0	0.5

Confusions for more complicated phoneme inventories, like the full set of vowels and consonants in English, can be predicted using linguistic features in exactly the same way. For larger inventories with many phonemes and features, the natural classes and algorithms discussed in Broe (1993) are particularly useful. Unlike the five vowel inventory, confusions for the full inventory of English are much more difficult to determine by simple visual inspection of the feature matrix. Details of the features used for English in the simulations reported here can be found in Frisch and Pisoni (1997).

Distinctive features and natural classes provide several insights into loss of contrast when a feature is categorically absent from the representation. From a perceptual perspective, this corresponds to a lack of ability to identify the perceptual cues that signal the contrast. However, behavioral measures, like the performance of children who use cochlear implants on the Minimal Pairs Test, rarely reveal a complete absence of distinctness. Instead, identification is imperfect or unreliable for particular contrasts. For our simulation, imperfect identification is treated as a probabilistic determinant of whether or not a particular contrast is perceived. Performance scores for each dimension of contrast on the Minimal Pairs Test can then be used to define a probability distribution of confusions for each phoneme. The method for generating this distribution will now be presented.

Since the Minimal Pairs Test is a two-alternative forced-choice task, the first step in converting performance on such a task to a prediction of open-set word recognition performance is to transform the two-alternative word identification score into an estimate of open-set feature identification performance. Open-set performance is assumed to be equivalent to closed-set performance, with the advantage of the restricted choices in a closed-set task factored out. The equation for estimating open-set performance is given in (1). This equation merely scales the performance from the closed-set range of chance performance to 100% to the open-set range of 0% to 100% (Black, 1957). For example, two-alternative closed-set performance of 50% corresponds to 0% open-set performance (chance); closed-set performance of 75% corresponds to 50% open-set performance.

(1) estimated open-set = (observed closed-set - chance) ÷ (1 - chance)

The Minimal Pairs Test provides identification scores for individual feature contrasts. The simplest assumption for identification of several contrasts, used by Frisch and Pisoni (1997), is that the identification of each contrast is independent. Under this assumption, the probability of simultaneously identifying multiple contrasts is the product of the individual contrast probabilities. For consonants, there are three contrasts and for vowels there are two. Assuming that the probability of identifying place, manner, and voicing are p, m, and v, respectively, and that the probability of identifying vowel place and vowel height are vp and vh, the distribution of possibilities for feature detection in vowels and consonants is given in Table 6.

Table 6

Probability distribution of identification performance.

Consonants		Vowels	
Contrasts	Probability	Contrasts	Probability
None	(1-p)×(1-m)×(1-v)	None	(1-vp)×(1-vh)
Place	p×(1-m)×(1-v)	Vowel Place	vp×(1-vh)
Manner	(1-p)×m×(1-v)	Vowel Height	(1-vp)×vh
Voicing	(1-p)×(1-m)×v	All	vp×vh
Place & Manner	p×m×(1-v)		
Place & Voicing	p×(1-m)×v		
Manner & Voicing	(1-p)×m×v		
All	p×m×v		

For example, with performance of 60% on Place, 70% on Manner, and 80% on Voicing on the Minimal Pairs Test, the predicted probability of identifying Place and Manner but not Voicing would be $p \times m \times (1-v) = 0.2 \times 0.4 \times (1-0.6) = 0.032$. The probability of identifying no features would be $(1-0.2) \times (1-0.4) \times (1-0.6) = 0.192$.

Note that the use of natural classes to predict confusions when a single contrast is lost apply equally well when multiple contrasts are lost. A unique set of distinct natural classes can be generated with any set of features removed, and the resulting minimal natural classes represent the phonemes that are no longer distinct when those features are lost.

Using the estimated probabilities of different outcomes shown in Table 6, we can generate a probabilistic distribution of contrast confusions for each phoneme. The probabilities for each possible outcome can then be used as weights for combining categorical confusion matrices of the sort given in Tables 4 and 5. The result is a probabilistic confusion matrix that takes into account the statistical reliability of feature identification for individual features and their combinations. This distribution is used to predict phoneme confusions in open-set word recognition.

The Phoneme Confusion Model

Given a model of phoneme confusion for open-set tasks, the simplest model of spoken word recognition applies the phoneme confusion model on each phoneme in the word and treats the result as the recognized word. This model does not employ lexical knowledge at all, as the perceptual result is not matched to an internalized representation in the mental lexicon. This model is certainly unrealistic as a model of normal spoken word recognition (although see the proposals of Boothroyd, 1997). It was shown above that open-set word recognition is influenced by competition of phonemically similar candidates in the mental lexicon, in both normal hearing adults and children who use cochlear implants (Kirk et al., 1995; Luce & Pisoni, 1998). This finding has been replicated several times for normal hearing adults (see Luce, Pisoni, & Goldinger, 1990) but has only recently been demonstrated in the pediatric cochlear implant population. We discussed above the claim by Pisoni et al. (1997) that the "Stars" have internalized a mental lexicon of spoken words which is crucial for subsequent language development. The "Controls" may not have developed such an internalized store. Thus, it may be the case that the Phoneme Confusion Model is appropriate for the "Control" group but not the "Stars" or normal hearing adults.

A sample application of the phoneme confusion model will now be presented, to give the reader a better feel for the predictions and results. The example is based on the average performance on the Minimal Pairs Test for the "Stars" group at two years post-implant. Mean two-alternative feature identification scores for this group are: Place 71.1%, Manner 80.1%, Voicing 70.3%, Vowel Place 92.7%, Vowel Height 94.2%. Fifty iterations of open-set phoneme confusions for the word *please* /pliz/ were generated based on these identification scores. The first fifteen simulated words are given in (2). The simulation perceived the entire sequence of phonemes /pliz/ correctly in 6 out of 50 iterations (12% correct). When scored by phonemes correct, 125 out of 200 phonemes were perceived correctly (62.5%). Qualitatively, the simulation performed much as expected given the Minimal Pairs Test identification scores and the assumption that word recognition can be modeled only using phoneme identification. The vowel was rarely confused (42 correct out of 50, 84%). For the initial stop /p/, manner was most frequently preserved (34 out of 50, 68%) followed by voicing (32 out of 50, 62%) and place (30 out of 50, 60%).

(2) φνιζ
πλιΔ
ρλιζ
δμιζ
βλιΖ
πϖιΣ
μλεν
πλιΖ
πνιζ
πδισ
πζο^{I}ζ
ωλιζ
Σλιζ
τλεζ
Σλαζ

Simulation and Results

Mean actual and predicted performance for each group on the PBK and LNT tests scored in percent correct phonemes and words is given in Figures 8 and 9. Recall that the LNT hard word lists were not given to the "Controls" as they were unable to perform the task.

When scored by phonemes correct, the model makes a good prediction of performance on the LNT for the "Stars". We take this as support for the general approach of using closed-set contrast identification performance to predict open-set segment identification (however, see Sommers, Kirk, & Pisoni, 1997, for evidence that there are very important differences in the patterns of performance between open-set and closed-set response formats for word recognition). The model over-predicts performance on the PBK for both groups and on the LNT for the "Controls."

When scored by words correct, the simulation under-predicts the performance of the "Stars" on the LNT. The model gives reasonable predictions of performance on the PBK for both groups, and on the

LNT for the “Controls.” These results are completely consistent with the conclusions above and also with claims made elsewhere in the literature. In particular, Kirk et al. (1995) and Frisch and Pisoni (1997) concluded that the words on the PBK and LNT do not differ greatly in their lexical characteristics, but they do differ in their familiarity to young children with cochlear implants. Children are more likely to be familiar with the words on the LNT than the PBK. Thus, we might expect that recognizing words on the PBK essentially becomes a task of non-word repetition for these children, in which there is no lexical knowledge used. For the LNT, by contrast, the words are known and so lexical information aids in the word recognition process. Since the Phoneme Confusion Model has no lexical information, we might expect it to provide a good model of word recognition for the PBK, but a poor one for the LNT. This is precisely the patterns shown in Figure 9 for the “Stars”.

Insert Figures 8 and 9 about here

For the “Controls,” we have claimed there is little evidence for the existence of a lexicon of spoken words as part of their linguistic knowledge. Figure 9 shows the Phoneme Confusion Model provides a reasonable prediction of the very poor word recognition performance by the “Controls.” These simulations suggest that, for words on the LNT word lists, successful pediatric cochlear implant users are comparing the degraded information they receive from their cochlear implant to some set of internalized phonological representations of lexical items in memory. Unlike the Phoneme Confusion Model of word recognition, actual listeners can accommodate “near misses”, which are modeled here as perceptual mistakes in the recognition of a phoneme as in [bliz] or [pliΣ], and retrieve the appropriate item, /pliz/.

Conclusions and Implications

Taken together, our analyses of prelingually deaf children who are learning spoken language through a cochlear implant suggest that at least some of these children have been able to begin the normal process of language acquisition, despite the degraded and impoverished physical signal made available to them through their devices. These studies demonstrate that the relevant units of language and lexical organization must be, to some degree, abstractions of the physical signal transmitted to the brain by the peripheral auditory system.

It is well-known in the field of psycholinguistics that much can be learned about a system by examining in great detail cases where the system breaks down. Children who use cochlear implants are not always able to learn spoken language. Differences in performance on a wide range of behavioral tests between the “Stars” and “Controls” highlight the importance of learning a repertoire of abstract phonetic/phonological units that act as the interface between the perception and production of speech. Children who are able to accomplish this task have mastered the tools of spoken language and become proficient perceivers and producers. The fact that language acquisition is not automatic for all children who use a particular implant and processing strategy suggests that the information transmitted by a cochlear implant is barely sufficient for the patterns of spoken language to be learned. If further research can pinpoint whether the source of individual differences is peripheral or central, then children who use cochlear implants may provide important answers to the nature versus nurture debate. If it is found that individual differences in nerve survival, number of electrodes, or some other component of the peripheral auditory system is linked to successful language acquisition with a cochlear implant, then it would appear that there is an innate maturational mechanism for language which need only receive enough input to activate and organize the linguistic system of the child. However, if children with comparable sensory inputs can differ drastically on their ability to acquire language, it may be that perceptual learning and experience play a crucial role. In this case, factors related to the type of clinical intervention, communication mode, and the specific learning environment in general may be the best predictors of successful spoken language acquisition. In either case, it is clearly important to now

consider differences between populations and individuals to develop an accurate model of language acquisition and linguistic development. Children who use cochlear implants are a new and growing population which can provide important new insights into the nature of the neural representation of language (e.g., the lexicon) and its relation to language acquisition.

References

Black, J. W. (1957). Multiple-choice intelligibility tests. *Journal of Speech and Hearing Disorders* 22(2): 213-235.

Boothroyd, A. (1997). Auditory capacity of hearing-impaired children using hearing aids and cochlear implants: Issues of efficacy and assessment. *Scandinavian Audiology Supplement* 26(Suppl 46): 17-25.

Broe, M. (1993). Specification theory: The treatment of redundancy in generative phonology. Unpublished Ph.D. dissertation, University of Edinburgh.

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

Frisch, S. & Pisoni, D. B. (1997). Predicting spoken word recognition performance from feature discrimination scores in pediatric cochlear implant users: A preliminary analysis. In *Research on Spoken Language Processing Progress Report No. 21* (pp. Bloomington, IN: Speech Research Laboratory, Indiana University. 261-288. Also submitted to *Ear and Hearing*.

Fryauf-Bertschy, H., Tyler, R. S., Kelsay, D., Gantz, B. J., & Woodworth, G. G. (1997). Cochlear implant use by prelinguistically deafened children: The influences of age at implant and length of device use. *Journal of Speech, Hearing, and Language Research* 40: 183-199.

Haskins, H. A. (1949). *A phonetically balanced test of speech discrimination for children*. Unpublished Master's Thesis, Northwestern University.

Kirk, K. I., Pisoni, D. B., and Osberger, M. (1995). Lexical effects on spoken word recognition by pediatric cochlear implant users. *Ear and Hearing* 16(5): 470-481.

Gathercole, S. E., Willis, C. S., Baddeley, A. D., & Emslie, H. (1994). The children's test of nonword repetition: A test of phonological working memory. *Memory* 2: 102-127.

Luce, P. A. and Pisoni, D. B. (1998). Recognizing spoken words: The Neighborhood Activation Model. *Ear and Hearing* 19: 1-36.

Luce, P. A., Pisoni, D. B., and Goldinger, S. D. (1990). Similarity neighborhoods of spoken words. In G. Altmann (ed.), *Cognitive models of speech processing: Psycholinguistic and computational perspectives.* Cambridge, MA: MIT Press.

Monsen, R. B. (1983). The oral speech intelligibility of hearing-impaired talkers. *Journal of Speech and Hearing Disorders* 48: 286-296.

Nusbaum, H. C., Pisoni, D. B., and Davis, C. K. (1984). Sizing up the Hoosier Mental Lexicon: Measuring the familiarity of 20,000 words. In D. Pisoni (ed.), *Research on Spoken Language Processing Progress Report No. 10*. Bloomington, IN: Indiana University. 357-376.

Osberger, M. J., Robbins, A. M., Todd, S. L., & Riley, A. I. (1994). Speech intelligibility of children with cochlear implants. *Volta Review* 96: 169-180.

Osberger, M. J., Robbins, A. M., Todd, S. L., Riley, A. I., Kirk, K. I., & Carney, A. E. (1996). Cochlear implants and tactile aids for children with profound hearing impairment. In F. H. Bess, J. S. Gravel, & A. M. Tharpe (eds.) *Amplification for children with auditory deficits*. Nashville, TN: Bill Wilkerson Center Press. 283-308.

Pisoni, D.B., Svirsky, M.A., Kirk, K.I., and Miyamoto, R.T. (1997). Looking at the "Stars": A first report on the intercorrelations among measures of speech perception, intelligibility, and language in pediatric cochlear implant users. Paper presented at the 5th International Cochlear Implant Conference, May 1997, New York. Also in *Research on Spoken Language Processing Progress Report No. 21*. Bloomington, IN: Speech Research Laboratory, Indiana University. 51-91.

Prince, A. & Smolensky, P. (1997). Optimality: From neural networks to universal grammar. *Science* 275: 1604-1610.

Reynell, J. K. & Gruber, C. P. (1990). *Reynell Developmental Language Scales: U.S. Edition*. Los Angeles, CA: Western Psychological Services.

Robbins, A. M., Renshaw, J. J., Miyamoto, R. T., Osberger, M. J., and Pope, M. L. (1988). *Minimal pairs test*. Indianapolis, IN: Indiana University School of Medicine.

Seidenberg, M. S. (1997). Language acquisition and use: Learning and applying probabilistic constraints. *Science* 275: 1599-1603.

Sommers, M. S., Kirk, K. I., & Pisoni, D. B. (1997). Some considerations in evaluating spoken word recognition by normal hearing, noise-masked normal hearing, and cochlear implant listeners I: The effects of response format. *Ear and Hearing* 18: 89-99.

Staller, S. J., Pelter, A. L., Brimacombe, J. A., Mecklenberg, D., & Arndt, P. (1991). Pediatric performance with the Nucleus 22-Channel Cochlear Implant System. *American Journal of Otology* 12: 126-136.

Stanley, R. (1967). Redundancy rules in phonology. *Language* 43: 393-436.

Figure 1: Average Minimal Pairs Test scores for the “Stars” and “Controls.”

Figure 2: Average Common Phrases Test scores for the “Stars” and the “Controls” in three modality conditions.

Figure 3: Average PBK and LNT scores in percent words correct for the “Stars” and “Controls.” The (*) indicates no data for the “Controls” on “Hard” word lists.

Figure 4: Average PBK and LNT scores in percent phonemes correct for the “Stars” and “Controls.” The (*) indicates no data for the “Controls” on “Hard” word lists.

Figure 5: Average speech intelligibility for the “Stars” and “Controls.”

Figure 6: Average vocabulary quotient on the PPVT for the “Stars” and “Controls.”

Figure 7: Average Reynell receptive and expressive language quotients for the “Stars” and “Controls.”

Figure 8: Comparison of observed and predicted performance for the “Stars” and “Controls” scored by phonemes correct.

Figure 9: Comparison of observed and predicted performance for the “Stars” and “Controls” scored by words correct.