RESEARCH ON SPOKEN LANGUAGE PROCESSING
Progress Report No. 25 (2001-2002)
Indiana University

Imitation of Nonwords by Hearing-Impaired Children with Cochlear Implants: Segmental Analyses[1]

Caitlin M. Dillon, Miranda Cleary, David B. Pisoni and Allyson K. Carter

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

---

[1] This work was supported by NIH NIDCD Research Grant DC00111 and NIH NIDCD Training Grant DC00012 to Indiana University. We would like to thank Dr. Ann Geers, Dr. Rosalie Uchanski and the research team at the Center for Applied Research on Childhood Deafness, Central Institute for the Deaf, St. Louis, MO 63110 for their invaluable help in this project. We are also grateful to Rose Burkholder, Cynthia Clopper and Jeff Reynolds for their help on this project.

Imitation of Nonwords by Hearing-Impaired Children with Cochlear Implants: Segmental Analyses

Abstract. The phonological processing skills of 24 pre-lingually deaf 8- and 9-year-old experienced cochlear implant users were measured using a nonword repetition task. The children heard recordings of 20 nonwords, one at a time, and were asked to repeat each pattern as accurately as possible. Detailed segmental analyses of the consonants in the children's imitation responses were carried out. Overall, 39% of the consonants were imitated correctly. Coronals were produced correctly more often than labials or dorsals. There was no difference in the proportion of correctly reproduced stops, fricatives, nasals, and liquids, or voiced and voiceless consonants. The similarity of the children's utterances to the model nonwords was also measured using a perceptual ratings task, in which normal hearing adult listeners rated the accuracy of the imitation responses in comparison to the targets. The imitation responses were also scored using a sequence comparison algorithm to compute linguistic distances and obtain similarity scores for the responses. The results obtained from the computational analyses were similar to "traditional" segmental accuracy scores. The segmental scores, perceptual ratings, and distances were strongly correlated with each other, and with the children's scores on other speech and language outcome measures of the component phonological processes involved in nonword repetition. In general, the children's nonword repetition performance was not correlated with their demographic characteristics. The results of this study indicate that experienced pediatric cochlear implant users are able to utilize their phonological knowledge of the ambient language to imitate and immediately reproduce novel sound patterns. Furthermore, the present findings demonstrate that the component processing skills tapped by the nonword repetition task are strongly related to variation in other measures of spoken word recognition, language comprehension, phonological working memory, and speech production.

Introduction

The remarkable ability of children as young as two years of age to spontaneously imitate the speech of adult models has helped researchers in developing theories of child language acquisition (e.g., Slobin & Welsh, 1973). Similarly, elicited nonword repetition tasks have been used by researchers to provide new insights into the language learning skills of adults and to study children with various language-learning difficulties (Edwards & Lahey, 1998). Studies have revealed that nonword repetition accuracy appears to be correlated with such skills as adults' ability to learn foreign-language lexical items (Papagno, Valentine, & Baddeley, 1991) and children's ability to learn the nonword names of toys (Gathercole & Baddeley, 1990). In the present study, we examined the nonword repetition performance of 24 children who were experienced cochlear implant users. The nonwords were a subset of the 40 nonwords in the Children's Test of Nonword Repetition (CNRep), a test designed to assess individual differences in phonological working memory in young normal-hearing children (Gathercole & Baddeley, 1996; Gathercole, Willis, Baddeley, & Emslie, 1994). The children were asked to listen to a nonword pattern and repeat it back aloud after a single auditory-only exposure. They were alerted in advance that the stimuli would be unfamiliar, and were told to imitate the items to the best of their ability. Such a task is complex because it requires the participant to successfully complete several auditory, cognitive, and articulatory processes, without relying on visual cues or exposure to previous tokens. After four or five years of experience with a cochlear implant, we speculated that many of these children possessed a phonological system sufficient to allow them to produce nonword imitations that closely resembled the targets. Furthermore, we expected that individual differences in the children's performance on a nonword imitation task such as the one used in the present study would be reflected in individual differences in their performance on other tasks measuring the component processes of speech perception and production, including working memory.

In this study, we measured the imitation responses obtained in the nonword repetition task in several ways. First, we transcribed the children's imitations and examined their accuracy. We calculated the percent of nonwords imitated completely correctly, the percent of target consonants imitated correctly and the percent of target consonants imitated with the correct manner, place, and voicing features. Second, we collected perceptual ratings from normal-hearing adult listeners as to the accuracy of the imitation responses compared to the target nonwords. Finally, we developed a computer program to calculate 'distance' scores that indicated how close the imitations were to their targets. An imitation's proximity to the target was based on whether or not the imitation matched the target in terms of its place, manner, and voicing features. We investigated the relationship between the children's nonword repetition performance as measured by each of the three methods, and their demographic characteristics as well as their scores on other speech and language measures. We also computed correlations between the children's accuracy scores, perceptual ratings and distance scores in order to investigate the relationship between these three different methods of scoring nonword repetition performance. Preliminary results from 14 of the children in this study were reported in Dillon and Cleary (2000) and Cleary, Dillon, and Pisoni (2002). Additional findings are reported below.

Previous studies of speech production performance in pediatric CI users have varied in their focus and approach. Speech samples have been analyzed from individual pediatric CI users (Chin, Pisoni, & Svec, 1994) and from groups of subjects (Kirk, Diefendorf, Riley, & Osberger, 1995). The speech samples have been spontaneous (Osberger et al., 1991), elicited (Dawson et al., 1995), and imitative (Sehgal, Kirk, Svirsky, Ertmer, & Osberger, 1998). Target stimuli for imitation tasks have included English words or sentences (e.g., Tye-Murray, Spencer, Bedia, & Woodworth, 1996) and nonwords (e.g., Tobey, Geers, & Brenner, 1994), varying in length, syllable structure, and segmental content. Imitation responses have been analyzed in a variety of ways. Researchers have analyzed the suprasegmental characteristics of the speech samples, such as intonation, duration, and intensity (Tobey et al., 1991; Tobey & Hasenstab, 1991; Tobey et al., 1994); the frequency with which certain segments and features are produced, regardless of target (Hesketh, Fryauf-Bertschy, & Osberger, 1991; Osberger et al., 1991; Serry, Blamey, & Grogan, 1997); the consistency with which certain segments and features are produced by each subject (Tobey & Hasenstab, 1991); as well as the segmental or featural accuracy of the response (Tobey et al., 1991; Geers & Tobey, 1992; Chin et al., 1994). When segments or features have been the focus of study, either consonants (Chin, Kirk, & Svirsky, 1997), vowels (Ertmer, Kirk, Sehgal, Riley, & Osberger, 1997), or both (Tobey et al., 1994) have been analyzed. The production of these sounds is sometimes scored according to the position of the target segment within the word, yielding comparisons between the accuracy of word-initial versus word-final consonants (Geers & Tobey, 1992).

Suprasegmental analyses of the present set of imitation responses have already been reported in Carter, Dillon, and Pisoni (in press). In that study, we found that 64% of imitations contained the correct syllable number and 61% had correct placement of primary stress. We also found significant correlations between the children's imitation performance in terms of each of these two suprasegmental properties on the one hand, and their performance on other speech, language and working memory measures on the other hand. In addition, the children's performance in terms of syllable number and primary stress placement was strongly correlated with imitation accuracy ratings provided by normal-hearing adults. This finding was consistent with the Hudgins and Numbers' (1942) conclusion that the speech rhythm (or suprasegmental properties) of 192 8- to 20-year-old deaf students was correlated with their speech intelligibility. Hudgins and Numbers also found a similar strong correlation between speech intelligibility and consonant production. Their findings encouraged us to focus on the consonant production accuracy of the nonword imitation responses obtained in the present study.

Two speech production studies that involved similar analyses to those reported in the present study were carried out by Dawson et al. (1995) and Sehgal et al. (1998). Dawson et al. reported the results of a study of 10 children who had become deaf before 5 years of age and had used CIs for 1;1 to 4;5 years. These children completed the Test of Articulation Competence (Fisher & Logemann, 1971), in which pictures of target words were used to elicit consonants in initial, medial, and final word positions. While Dawson et al. found individual differences between the children, they found that overall, in terms of manner of articulation, affricates were produced correctly less often than stops, nasals, fricatives, and glides. In regard to place of articulation, front consonants (labials, labio-dentals, and lingua-dentals) were produced correctly more often than middle (alveolars and palatals) or back consonants (velars and glottals). They also found that 60% of voiceless consonants and 54% of voiced consonants were produced correctly.

Sehgal et al. (1998) studied the consonant productions of 10 pre-lingually deaf children who had used a CI for 1 to 3 years. Their task involved imitation of 60 different nonsense syllables, containing 20 different consonants. Each syllable was imitated 3 times by each child. They found that the children produced the bilabial place feature correctly more often than any other place, followed by alveolar, then dental, and lastly velar. Affricates were produced correctly less often than consonants with other manners of articulation, but no significant difference was revealed among the stops, nasals, fricatives, and glides. Across all of their consonant productions, the children produced the manner feature correctly 62% of the time, the place feature 65% of the time, and the voicing feature 62% of the time.

The results of Dawson et al. (1995) and Sehgal et al. (1998) together indicate that children who use CIs produce labial consonants more accurately than consonants with other places of articulation, and they produce affricates correctly less often than consonants with any other manner of articulation Similarly, in a study of the spontaneous speech of 14 pre-lingually deaf children who had used CIs for 1 year, Osberger et al. (1991) found that the children correctly produced bilabial stops and nasal /m/ most often, followed by alveolar and velar stops, followed by fricatives, then liquids, and lastly glides.

Based on these results, we expected that the children in our study would produce labials correctly more often than consonants with other places of articulation, and possibly that the target stops in our study would be produced correctly more often than fricatives, and fricatives more often than liquids. In the following pages, we present the results of our study of the children's imitation accuracy of the consonants in the target nonwords, the children's scores in terms of the distance of their responses from the target nonwords, and their accuracy based on perceptual ratings by adult listeners. We report our findings regarding the relationship between each of these three methods of evaluating the children's imitations and their demographic characteristics. In addition, we describe the results of correlational analyses between the children's nonword repetition performance and their performance on several other speech and language outcome measures.

Method

Subjects

The CI users were 24 children who participated in the Central Institute for the Deaf's 'Cochlear Implants and Education of the Deaf Child' project in either 1999 or 2000 (Geers et al., 1999). While 88 children participated in the nonword repetition task used in this study, some of the participants did not produce an imitation response to all of the 20 target stimuli. Figure 1 shows the number of children who provided a number of imitations from 0 to 20. In the present study, we report on the nonword repetition performance of the 24 children who provided an imitation response to all 20 target stimuli. While these 24 children's overall performance was slightly better than that of the larger group as a whole, the results presented below demonstrate that there was nevertheless a wide range of individual variability among the 24 children.

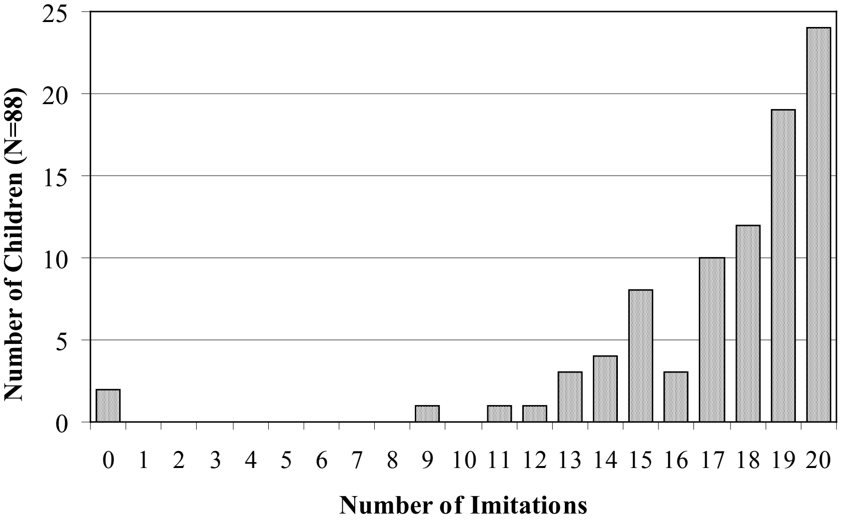

Figure 1. The number of imitations produced by the original 88 children who participated in the nonword repetition task during the 1999 and 2000 CID summer programs.

Demographic information on the 24 children is shown in Table 1. Responses from 15 males and 9 females were analyzed. Nineteen of the participants were congenitally deaf; the other five children were deaf by the time they were 3 years old. The average duration of deafness prior to implantation was 3.0 years (*SD* = 1.1, range 0.7 to 5.4 years). At the time of testing, the children in the group had used a cochlear implant for an average of 5.4 years (*SD* = 0.8, range 3.8 to 6.6 years). The average chronological age of the children was 8.8 years (*SD* = 0.5, range 8.2 to 9.9 years). Children who primarily used oral communication and children who primarily used total communication (TC) were included in the group. The average Communication Mode score was 4.4 (*SD* = 1.4, range = 2 to 6). This score is the average of scores assigned at five intervals: prior to implantation, the first year after implantation, the second year after implantation, the third year after implantation, and the current year of testing. At each interval, a ranking using the following scale was assigned to each child: 1 point for TC with emphasis on sign, 2 points for TC with equal emphasis on sign and speech, 3 for TC with emphasis on speech, 4 for cued speech, 5 for auditory-oral communication, and 6 for auditory-verbal communication. Therefore, a score of 3.5 or lower indicates that the child's method of communication was primarily TC, while a score of 3.6 or higher indicates that the child's communication setting was primarily oral (Geers et al., 1999). Accordingly, 18 of the children in the present study used oral communication and 6 of the children used total communication. All of the children who participated in the nonword repetition task were prelingually deafened and 23 were users of a Nucleus 22 cochlear implant and the SPEAK coding strategy. One child used a Clarion cochlear implant.

Child ID #	Gender	Age at onset of deafness (in months)	Duration of deafness (in years)	Duration of CI use (in years)	Chronological Age (in years)	Communication Mode Score
00102	m	0	3.3	5.8	9.1	5.2
00105	m	0	3.9	4.8	8.7	5.0
00109	m	0	3.3	5.0	8.2	3.6
00110	m	0	4.7	4.8	9.5	3.4
00207	f	0	2.1	6.1	8.3	4.8
00208	m	0	2.1	6.1	8.3	4.6
00213	f	36	2.0	4.3	9.4	6.0
00307	m	18	0.7	6.0	8.2	5.2
00309	m	0	4.0	4.8	8.7	2.6
00314	m	0	4.5	5.4	9.9	2.0
99101	m	0	3.4	5.6	9.0	5.4
99103	f	0	2.2	6.5	8.7	6.0
99104	m	0	2.9	6.6	9.5	5.0
99105	f	0	3.2	5.3	8.5	4.6
99108	f	0	2.4	5.9	8.3	5.8
99205	f	0	3.0	5.4	8.3	4.2
99207	m	0	3.9	4.5	8.4	4.0
99211	m	10	2.7	4.7	8.2	2.2
99214	m	24	1.7	4.5	8.2	2.0
99301	m	18	1.6	6.0	9.0	6.0
99304	m	0	2.6	6.4	9.0	6.0
99305	f	0	3.3	5.1	8.4	5.6
99307	f	0	5.4	3.8	9.1	4.2
99312	f	0	3.2	6.5	9.7	2.4
Mean (SD):		**4.4 (9.7)**	**3.0 (1.1)**	**5.4 (0.8)**	**8.8 (0.5)**	**4.4 (1.4)**

Table 1. Demographic information for the 24 children analyzed.

Stimulus Materials

All of the 40 nonword stimuli on the CNRep test consisted of sound sequences that are phonotactically permissible in English but lack semantic content. The subset of 20 nonwords used for this study were chosen by eliminating the 20 items that showed the least amount of variance in scores obtained previously in our lab from younger normal-hearing children (Carlson, Cleary, & Pisoni, 1998). We also eliminated some nonwords that were essentially common real words attached in an unfamiliar manner to a standard affix. Five nonwords remained at each of four lengths: 2, 3, 4, and 5 syllables. Each of the nonwords is shown with its phonemic transcription in Table 2.

The nonwords used in the present study were originally designed to assess individual differences in phonological working memory in young normal-hearing children (CNRep; Gathercole & Baddeley, 1996; Gathercole et al., 1994). While the nonwords were balanced in terms of the number of syllables contained in each, the stimuli were not balanced in terms of such linguistic characteristics as CV structure, consonant or vowel features, or stress patterns. Nevertheless, as shown in Table 3, the consonants contained in the target nonwords include consonants with a range of manner features (stops, fricatives, an affricate, nasals, and liquids), consonants with all three gross places of articulation (labial, coronal, and dorsal), and both voiced and voiceless consonants.

Number of syllables	Target nonword orthography	Target nonword transcription
2	ballop	ˈbæ.ləp
	prindle	ˈprɪn.dl̩
	rubid	ˈruˌbɪd
	sladding	ˈslæ.diŋ
	tafflist	ˈtæ.flɪst
3	bannifer	ˈbæ.nəˌfɚ
	berrizen	ˈbɛ.ɹəˌzɪn
	doppolate	ˈdɑ.pəˌle͡ɪt
	glistering	ˈglɪ.stɚ.iŋ
	skiticult	ˈskɪ.ɾəˌkʌlt
4	comisitate	kəˈmi.səˌte͡ɪt
	contramponist	kənˈtɹæm.pəˌnɪst
	emplifervent	ɛmˈplɪ.fɚˌvɛnt
	fennerizer	ˈfɛ.nɚˌa͡ɪ.zɚ
	penneriful	pəˈnɛ.ɹəˌfʌl
5	altupatory	ælˈtu.pəˌtɔ.ri
	detratapillic	diˈtræ.ɾəˌpɪ.lɪk
	pristeractional	ˈprɪ.stɚˌæk.ʃə.nl̩
	versatrationist	ˈvɚ.səˌtɹe͡ɪ.ʃəˌnɪst
	voltularity	ˈvɑl.tʃʊˌlɛ.ɹəˌti

Table 2. The 20 nonwords used in the present study (see Carlson et al., 1998), adapted from Gathercole et al. (1994).

		Labial	Coronal	Dorsal
Stop	voiceless	9 /p/	17 /t/	6 /k/
	voiced	4 /b/	5 /d/, 2 /ɾ/	1 /g/
Fricative	voiceless	5 /f/	9 /s/, 2 /ʃ/	---
	voiced	3 /v/	2 /z/	---
Affricate	voiceless	---	1 /t͡ʃ/	---
Nasal	voiced	3 /m/	10 /n/	2 /ŋ/
Liquid	voiced	---	14 /l/, 17 /r/	---

Table 3. The 112 target consonants in the 20 nonwords.

Procedure

The nonword stimuli from the CNRep were originally recorded by a British talker. For the present study, they were re-recorded by a female speaker of American English (Carlson et al., 1998) and presented auditorily to the children via a desktop speaker (Cyber Acoustics MMS-1) at approximately 70 dB SPL. In a few cases, the signal level was increased at the child's request. Each child heard the nonword stimuli played aloud one at a time, in random order. The children were told that they would hear a ‘funny word’, and were instructed to repeat it back as well as they could. Their imitation responses were recorded via a head-mounted microphone (Audio-Technica ATM75) onto digital audio tape using a TEAC DA-P20 tape deck. The DAT tapes were later digitized and segmented into individual sound files. Each imitation response was independently transcribed by two trained transcribers, the first and second authors. Intertranscriber agreement between the two authors was 93%. Discrepancies were resolved by a third transcriber.

Results and Discussion

In the present study, we measured the children’s nonword repetition performance in three ways. First, we calculated several “traditional” accuracy scores for the imitations based on the whether or not they were completely correct, contained the correct consonants, or contained consonants with the correct features (manner, place, and voicing). Second, we collected behaviorally-based perceptual ratings from adult listeners as to the accuracy of the imitations. Third, we computed several distance scores that provided a measure of the extent to which the imitations differed from the targets, using three different distance matrices that were based on consonant features. These three methods of measuring the children’s nonword repetition performance, and the results of each, are described in detail below.

Traditional Accuracy Analyses

Scoring. Previous studies have assessed nonword repetition responses using a binary scoring procedure (e.g., Avons, Wragg, Cupples, & Lovegrove, 1998; Gathercole, 1995). The examiners credited the children with either 1 point or 0 points for each response. Any error, even if the error involved only a single segment (i.e., phoneme), usually resulted in no credit. Provisions have sometimes been made for predictable patterns of immature articulation in very young children. However, the children with CIs in the present study frequently made segmental errors, so that out of the 480 imitation responses, only 5% of the imitations would have received full credit using this binary scoring procedure. The traditional scoring procedure was therefore not suitable for use in the present study. Instead, we calculated a binary accuracy score for each segment. That is, if the imitation segment was the same English phoneme as its corresponding target segment, then it received 1 point. Three additional accuracy scores were also computed for the consonants. For these scores, points were assigned based on the featural accuracy of the imitation segment. All of the accuracy scores are described in detail below.

(1) <u>Segment Score:</u> An imitation segment was counted as correct and given 1 point if the segment was correctly reproduced. For example, for a target /p/, if a child produced a /p/, he/she was given 1 point. The production of any other phoneme received 0 points.

(2) <u>Manner Feature Score:</u> An imitation consonant was counted as correct and given 1 point if the consonant was correct in terms of manner of articulation. For example, for a target /p/, which is a stop, if a child produced any stop consonant, such as [p], [b], [t] or [d], he/she was given 1 point. If, for a target /p/, a child produced a fricative, affricate, or any segment whose manner was other than a stop, he/she received 0 points.

(3) <u>Voice Feature Score</u>: An imitation consonant was counted as correct and given 1 point if the consonant was correct in terms of voicing. For example, for a target /p/, which is voiceless, if a child produced any voiceless consonant, he/she was given 1 point. If, for a target /p/, a child produced an imitation response with a voiced segment, he/she received 0 points.

(4) Place Feature Score: An imitation consonant was counted as correct and given 1 point if the place feature of the consonant was correct in terms of the three gross places of articulation (labial, coronal, and dorsal). For example, for a target /p/, which is a labial, if a child produced any labial consonant, such as [p], [b], [f] or [v], he/she received 1 point. If, for a target /p/, a child produced a coronal or dorsal, he/she received 0 points.

The segmental accuracy scores for all of the segments in a given imitation were then summed to calculate an overall segmental accuracy score for each imitation. In addition, the segmental accuracy scores assigned to all of the imitations produced by a given child were averaged in order to calculate an average segmental accuracy score for each child. The accuracy scores described in this section will be referred to as the "traditional" accuracy scores.

Results. As stated above, the children only produced 5% of the nonwords correctly without mistakes. Results of our analyses of the percent of consonant segments and percent of consonant features reproduced correctly are presented below.

Overall, 39% of the consonants were reproduced correctly. As shown in Figure 2, 56% of the target consonants were imitated correctly in terms of manner, 61% were correct in terms of place, and 66% were correct in terms of voicing. That is, the children produced voicing correctly most often, and imitated place correctly more often than manner. This finding is consistent with the results reported by Chin et al. (1997) in a study of 9 hearing-impaired children who had used a cochlear implant for an average of 5 years. They used of the Goldman-Fristoe Test of Articulation, which requires the child to name 44 real English words, shown in pictures presented to the child. The words in the Goldman-Fristoe Test contain each of the English consonants at least once in word-initial, word-medial and word-final positions. Chin et al. found that the children in their study produced the voicing feature accurately more often than the place or manner features (voicing = 53%, place = 48%, manner = 40%).

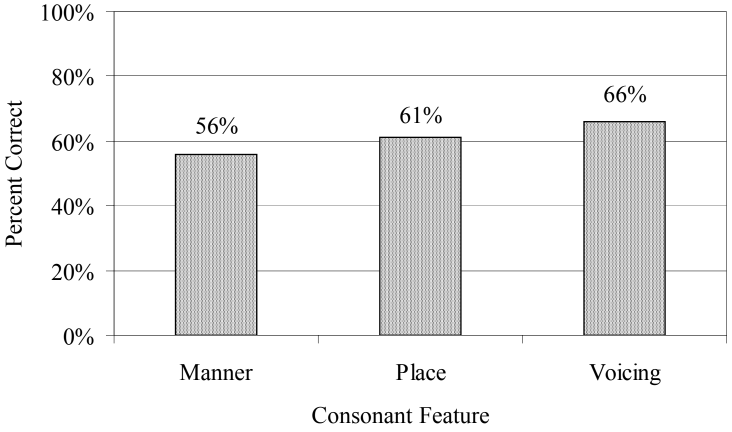

Figure 2. The percent of target consonants that were imitated correctly overall, and in terms of place, in terms of manner, and in terms of voicing.

To gain further insight into whether the children had more difficulty imitating certain feature values more than others, we examined the children's imitations of each of the features more closely. For example, to investigate the manner feature, we compared the proportion of target stops, fricatives, nasals, liquids, and glides imitated with the correct manner feature. For the place feature, we calculated the proportion of labials imitated as labials, the proportion of target coronals imitated as coronals, and the proportion of target dorsals imitated as dorsals. Similarly, for the voicing feature, we calculated the proportion of target voiceless consonants and target voiced consonants imitated with the correct voicing. The results of these calculations are shown in Figures 3, 4, and 5, respectively. In the results presented below, the target affricate was not included because there was only one.

As shown in Figure 3, the manner feature was imitated with similar levels of accuracy across the four target manners of stop, fricative, nasal and liquid (59%, 58%, 55%, and 51%, respectively). A one-way ANOVA revealed that the manner feature of the target consonant did not significantly affect the number of imitations produced with the correct manner.

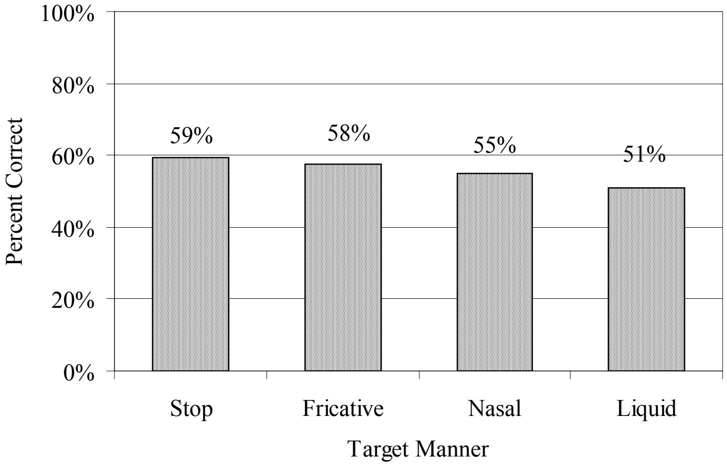

Figure 3. Percent of target consonants imitated correctly in terms of manner.

As shown in Figure 4, the children correctly imitated the place feature of a greater proportion of target coronals than labials or dorsals (66% versus 52% and 40%, respectively). A one-way ANOVA revealed a significant main effect of target place on whether or not the imitation consonant was produced with the correct place ($F(2, 109) = 12.1$, $p < .001$). Post-hoc Tukey tests indicated that target coronal consonants were reproduced with the correct place more often than target labials or dorsals ($p < .01$ and $p < .001$, respectively). However, there was no significant difference between the number of target labials and dorsals reproduced with the correct place.

As shown in Figure 5, the children produced imitations with the correct voicing for an equal proportion of target voiceless and target voiced consonants (66%). A one-way ANOVA confirmed that there was not a significant difference between the number of target voiceless consonants reproduced with the correct voicing, and the number of target voiced consonants reproduced with the correct voicing.

Because we found that the imitation accuracy scores differed depending on the specific place feature of the target (Figure 4), we further investigated the accuracy scores for the segments, grouping them according to both their word position and their manner and place features. We found again the result shown in Figure 4, that coronals were imitated correctly more often than labials or dorsals, and this pattern held up across target segments of various places, manners and word positions.

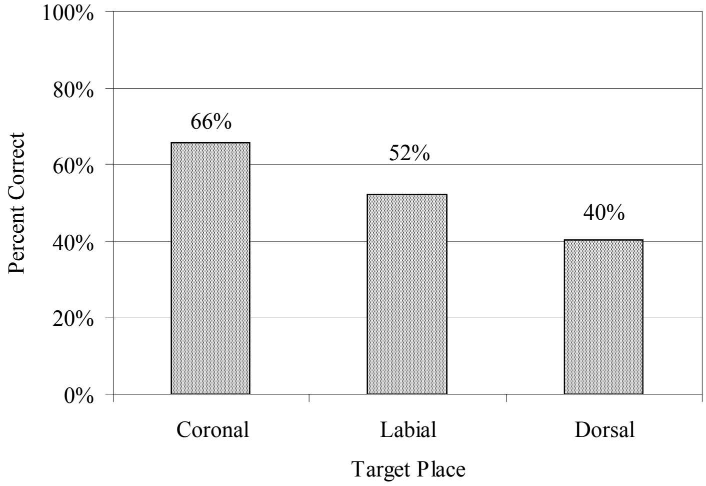

Figure 4. Percent of target consonants imitated correctly in terms of place.

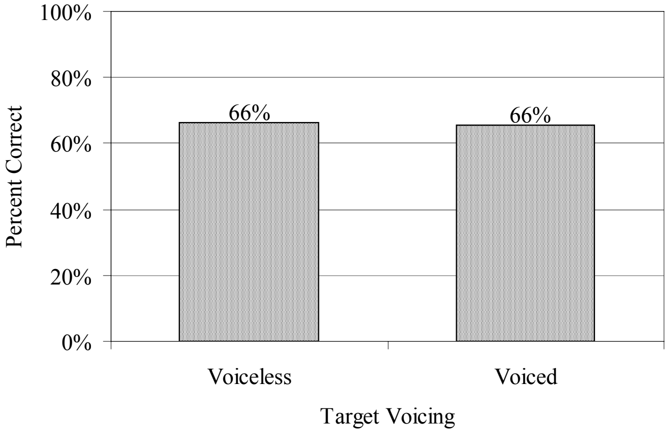

Figure 5. Percent of target consonants imitated correctly in terms of voicing.

We also used the segmental accuracy scores for each response to complete analyses by item and by subject. For the item analysis, we computed mean segmental accuracy scores for each target nonword (averaged across children). We also calculated the mean segmental accuracy score for each target syllable length (2, 3, 4, and 5). The results of the item analysis, shown in Figure 6, revealed that the segmental accuracy scores differed depending on the target nonword, and that the shorter target nonwords (2- and 3-syllables) tended to be imitated with greater accuracy than the longer target nonwords (4- and 5-syllables). A one-way ANOVA revealed a significant main effect of syllable length ($F(3, 16) = 4.3$, $p < .05$). Post-hoc Tukey tests indicated that there was a significant difference between the percentage of consonants reproduced correctly for 2-syllable words compared to 5-syllable words ($p < .05$).

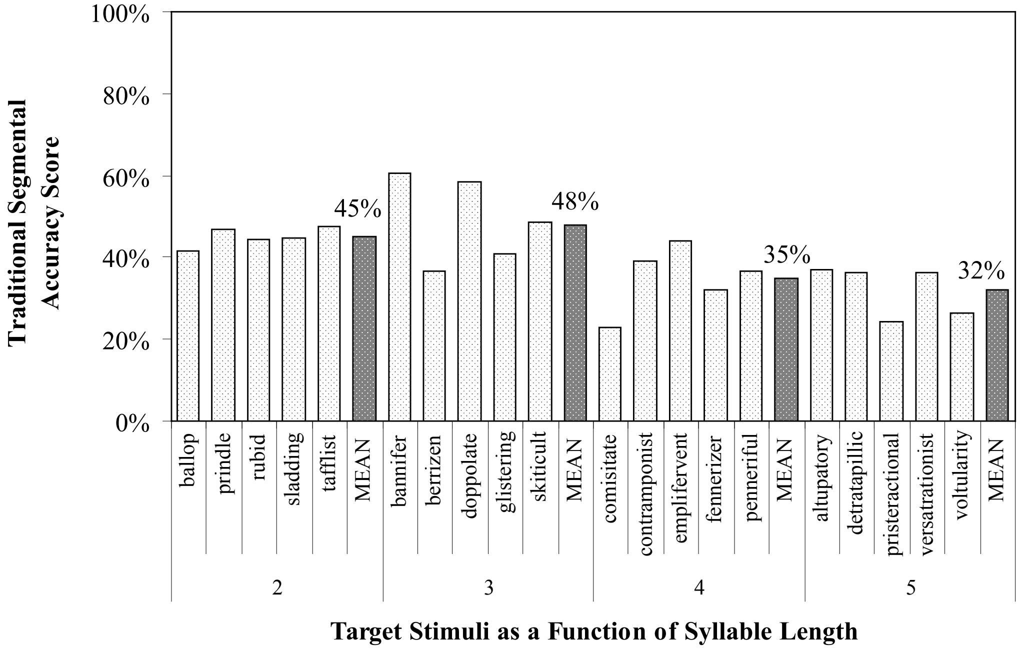

Figure 6. Mean traditional segmental accuracy score for each target nonword and for each target syllable length, averaged across children.

For the subject analysis, we computed mean segmental accuracy scores for each child (averaged across target nonwords), shown in Figure 7. The individual children's segmental accuracy scores exhibited a large degree of variation, ranging from 8% to 76%.

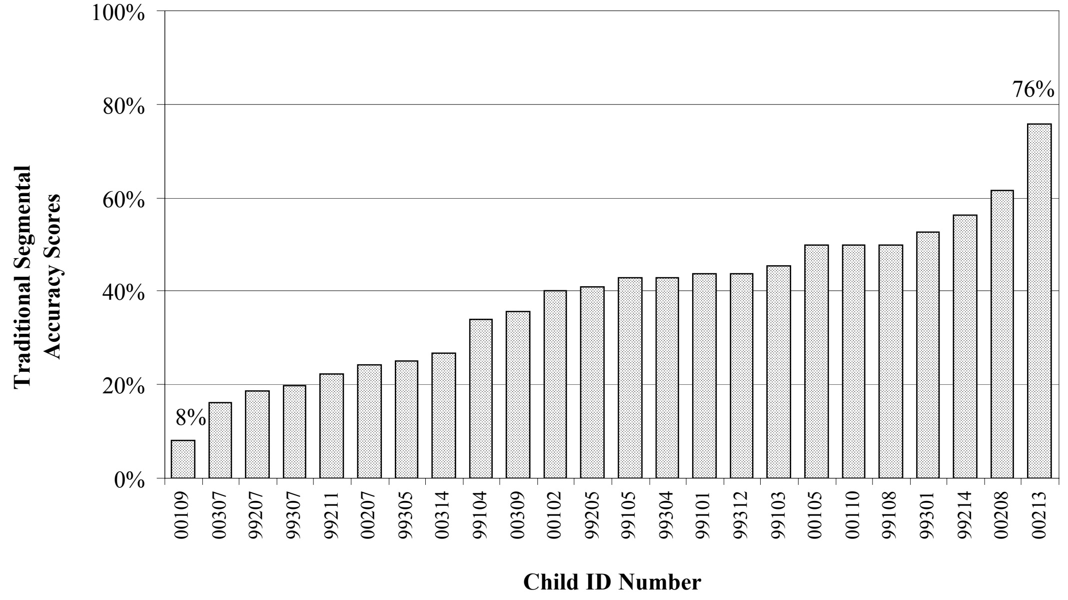

Figure 7. Mean traditional segmental accuracy score for each child, averaged across target nonwords.

Correlational Analyses. We were also interested in whether or not there were significant correlations between the children's accuracy scores and their demographic characteristics. Specifically, we computed correlations between the children's segmental and featural accuracy scores and the following demographic variables: age at onset of deafness, duration of deafness prior to implantation, age at implantation, duration of CI use, age at time of testing, gender, number of active electrodes, and degree of exposure to an oral-only communication environment (based on Communication Mode scores). As shown in Table 4, none of the demographic variables were significantly correlated with any of the traditional accuracy scores. We suspect that the relative homogeneity of the demographic characteristics of the children in this study might have prevented significant correlations. In the present study, 18 of the 24 children participated in oral communication programs, while only 6 of the children participated in total communication programs. In a recent larger study including the children described in the present paper, we collected perceptual ratings of the nonword responses from 76 children who use cochlear implants (Dillon, Burkholder, & Pisoni, this issue). We found moderate correlations between the children's average perceptual ratings and both age at onset of deafness and duration of deafness, and a strong correlation between the children's average perceptual ratings and their communication mode scores. The children whose educational environments emphasized oral communication received higher ratings for their nonword responses than the children involved in total communication programs.

	Correlation r values			
Demographic Variables	Segmental Accuracy	Manner Accuracy	Place Accuracy	Voicing Accuracy
Age at onset of deafness	+.40	+.33	+.29	+.27
Duration of deafness	-.26	-.20	-.16	-.07
Age at implantation	+.06	+.08	+.05	+.17
Duration of cochlear implant use	+.09	+.08	+.11	+.01
Age at time of testing	+.24	+.24	+.30	+.30
Gender	+.11	+.11	+.16	+.08
Number of active electrodes	-.05	-.13	+.03	.00
Degree of exposure to oral-only communication	+.27	+.23	+.16	+.17

Table 4. Correlations between the 24 children's demographic characteristics and their mean traditional segmental, manner, place, and voicing accuracy scores.

Additionally, we were interested in the extent to which the children's performance on the nonword imitation task would correlate with separate measures of the component processes involved in the imitation of a nonword stimulus. Although the nonword repetition task used in the present study may appear to be relatively simple at first glance, it in fact involves multiple component processes: auditory and phonological encoding, short-term storage of the target item in working memory, and articulatory planning and production. In order to be able to imitate a nonword pattern, a child needs to perform reasonably well in each of these component processes. The fact that the children in this study also participated in tasks that measured their performance on these component processes as part of another concurrent study (Geers et al., 1999) provided an unusual opportunity to assess the contribution of these component processes. Thus, correlations between the children's scores on several of these assessment tasks and their nonword imitation scores are reported below and shown in Table 5. The assessment tasks are described below.

	Correlation *r*-values			
Outcome and Process Measures of Performance	Segmental Accuracy	Manner Accuracy	Place Accuracy	Voicing Accuracy
Word Recognition				
Word identification, closed-set, pointing response				
Word Intelligibility by Picture Identification (WIPI)	+.72**	+.74**	+.80**	+.82**
Word identification, open-set, spoken repetition				
Lexical Neighborhood Test, Easy Words (LNTe)	+.80**	+.76**	+.82**	+.76**
Word identification, open-set, spoken repetition				
Lexical Neighborhood Test, Hard Words (LNTh)	+.80**	+.76**	+.80**	+.80**
Word identification, open-set, spoken repetition				
Lexical Neighborhood Test, Multisyllabic Words (mLNT)	+.77**	+.77**	+.82**	+.82**
Sentence identification, open-set repetition of target sentence				
Bamford-Kowal-Bench (BKB)	+.80**	+.80**	+.86**	+.85**
Language Comprehension				
Closed set, receptive language comprehension of words and sentences				
Test of Auditory Comprehension of Language-Revised (TACL-R)	+.69**	+.71**	+.70**	+.74**
Phonetic Feature Discrimination				
Perception of speech pattern contrasts				
Video Game for Assessing speech pattern contrasts (VIDSPAC)	+.34	+.36	+.38	+.40
Working Memory				
Forward Digit Span				
Wechsler Intelligence Scale for Children (WISC-III) Auditory Digit Span Subtest, Forward Recall of Digit-Name Lists	+.64**	+.67**	+.62**	+.64**
Backward Digit Span				
Wechsler Intelligence Scale for Children (WISC-III) Auditory Digit Span Subtest, Backward Recall of Digit-Name Lists	+.30	+.37	+.45*	+.50*
Speech Production				
Speech Intelligibility				
McGarr Sentence Intelligibility Test	+.82**	+.85**	+.87**	+.88**
Speaking Rate				
McGarr Mean Duration of 3-Syllable Sentences (log, msec)	-.63**	-.67**	-.68**	-.68**
McGarr Mean Duration of 5-Syllable Sentences (log, msec)	-.81**	-.85**	-.90**	-.94**
McGarr Mean Duration of 7-Syllable Sentences (log, msec)	-.84**	-.85**	-.91**	-.92**

Table 5. Correlations between the 24 children's scores on several outcome and process measures, and their mean traditional segmental, manner, place, and voicing accuracy scores. $^{**}p \leq .001$, $^{*}p \leq .05$

Several of the assessment tasks were tests of speech perception and spoken word recognition. The Word Intelligibility by Picture Identification (WIPI) test is a closed set test of spoken word identification involving a pointing response (Ross & Lerman, 1979). The Lexical Neighborhood Test (LNT) is an open-set test of spoken word identification consisting of 100 monosyllabic words divided into four lists of 25 words each (Kirk, Pisoni, & Osberger, 1995). Two of the lists contain words that are 'lexically easy' (i.e., are phonetically similar to very few other words) and two of the lists contain words that are 'lexically hard' (i.e., are phonetically confusable with many other words). A child is typically tested on one 'easy' word list and one 'hard' word list, with separate percent-correct scores obtained for each list. The Multisyllabic Lexical Neighborhood Test (MLNT) is analogous to the LNT but uses multisyllabic words of 2 or 3 syllables. The Bamford-Kowal-Bench Sentence List Test (BKB) is an open-set task involving spoken repetition of a target sentence (Bench, Kowal & Bamford, 1979). As shown in Table 5, there were strong correlations between the children's 4 traditional accuracy scores (segmental, manner, place, voicing) and their scores on the WIPI ($r = +.72$, $+.74$, $+.80$, $+.82$, respectively; $p < .001$), LNTe ($r = +.80$, $+.76$, $+.82$, $+.76$, respectively; $p < .001$), LNTh ($r = +.80$, $+.76$, $+.80$, $+.80$, respectively; $p < .001$), mLNT ($r = +.77$, $+.77$, $+.82$, $+.82$, respectively; $p < .001$), and BKB ($r = +.80$, $+.80$, $+.86$, $+.85$, respectively; $p < .001$). These results indicate that children who scored higher on measures of spoken word recognition tended to produce more correctly imitated consonants.

The battery of tests administered by CID also included the Test of Auditory Comprehension of Language Revised (TACL-R), a language comprehension measure that assesses children's receptive vocabulary, morphology, and syntax (Carrow-Woolfolk, 1985). The children's 4 traditional accuracy scores were each highly correlated with their TACL scores ($r = +.69$, $+.71$, $+.70$, $+.74$; $p < .001$). These correlations with the TACL-R indicate that better performance on the nonword repetition task used in the present study corresponds to higher language comprehension scores in terms of receptive vocabulary, morphology, and syntax.

A measure of phonetic feature discrimination, VIDSPAC, was also administered to the children during the CID summer program. The VIDSPAC is a video game test that was specifically designed to measure hearing-impaired children's ability to perceive speech feature contrasts (Boothroyd, 1997). The children's VIDSPAC scores were not significantly correlated with their traditional accuracy scores.

A measure of working memory was also obtained from the children using the WISC Digit Span Supplementary Verbal Sub-test of the Wechsler Intelligence Scale for Children, Third Edition (WISC-III) (Wechsler, 1991). This memory span task has both a 'digits forward' subsection and a 'digits backward' subsection. For the forward digit span task, a child listens to and repeats lists of digits as spoken live-voice by the experimenter at a rate of approximately one digit per second (WISC-III Manual, Wechsler, 1991). Two lists are administered at each list length, beginning with two digits. The list length is increased one digit at a time until the child fails to correctly repeat both lists administered at a given length. The child receives points for correct repetition of each list, with no partial credit. The backward digit span task differs from the forward digit span task only in that the children are asked to repeat back the digits that they heard in backwards order, starting with last digit that was presented to them, and finishing with the first digit that was presented. There were strong correlations between the children's 4 traditional accuracy scores and their forward digit spans ($r = +.64$, $+.67$, $+.62$, $+.64$; $p \leq .001$). This result suggests that a longer digit span as measured by the WISC task corresponds to higher scores on the nonword repetition task used in the present study. The correlations between the children's backward digit spans and 2 of the 4 traditional accuracy scores (segmental and manner) were not significant, but there were moderate correlations between backward digit spans and both the place and voicing accuracy scores ($r = +.45$, $+.50$, respectively; $p < .05$). Backward digit span is often considered a measure of executive function, and is usually found not to correlate with measures that tap into the same cognitive processes as forward digit span, which is a measure of phonological coding and verbal rehearsal (see Engle, 2002; Lezak, 1983; Li & Lewandowsky, 1995; Rosen & Engle, 1997).

As part of the larger study at CID, a measure of speech intelligibility was also obtained from each child using the McGarr Sentence Intelligibility Test (McGarr, 1983). This test involves eliciting sentences containing either 3, 5 or 7 syllables in length. The child was provided with spoken and/or signed models of each sentence as well as the printed text of each sentence, and was prompted to speak as intelligibly as possible. The children's utterances were recorded and later played back to naïve listeners who were asked to transcribe the children's speech using standard orthography. This provided an objective measure of speech intelligibility. Each child's productions were also submitted to an acoustic analysis. Included among the various acoustic measures was a simple measure of sentence duration. Pisoni and Geers (2000) reported that CI children's speaking rate on the McGarr sentences, particularly, the longer seven syllable sentences, was strongly correlated with measures of working memory as well as with speech intelligibility. For this reason, in the present study, we examined the relationships between nonword repetition performance and McGarr Intelligibility. We also examined the relations between nonword repetition performance and sentence duration (duration being inversely related to speaking rate). We found strong correlations between the children's 4 traditional accuracy scores and McGarr speech intelligibility ($r = +.82, +.85, +.87, +.88$; $p < .001$). Thus, the children who produced more intelligible speech on the McGarr task also tended to reproduce more consonants correctly in their nonword repetitions. We also found strong negative correlations between the children's 4 traditional accuracy scores and McGarr speaking rates for the 3-syllable sentences ($r = -.63, -.67, -.68, -.68$; $p < .001$), for the 5-syllable sentences ($r = -.81, -.85, -.90, -.94$; $p < .001$), for the 7-syllable sentences ($r = -.84, -.85, -.91, -.92$; $p < .001$). These negative correlations suggest that the children who spoke more slowly in the McGarr task tended to perform more poorly on the nonword repetition task.

In summary, these results show that children who correctly reproduced a greater number of consonants or consonant features also tended to score higher on a wide range of outcome measures that assess the component processes involved in nonword repetition. That is, they tended to have higher scores for measures of spoken word recognition, language comprehension, phonological working memory, and speech production.

Perceptual Ratings

Procedure. As part of a larger study, we also obtained perceptual ratings for each child's nonword repetition performance using these same utterances (see Dillon, Burkholder, Cleary & Pisoni, this issue). This perceptual measure consisted of repetition accuracy ratings for each child's productions, gathered from monolingual English-speaking normal-hearing adult listeners who had no experience with the speech of deaf or hearing-impaired persons. The behaviorally-based perceptual ratings were obtained from a play-back experiment in the following manner: on each of 280 randomized trials, the listener heard a target stimulus followed by 1 second of silence and then by a child's imitation response. The listener was asked to rate the target-imitation pair on a seven-point scale using the following endpoint labels: 1 = 'totally fails to resemble the "target" utterance', 7 = 'perfectly accurate rendering of the "target" utterance, ignoring differences in pitch'. An average rating per imitation was calculated, from which an average rating per child was also calculated.

Results. An item analysis and a subject analysis were also carried out for the perceptual ratings. The results of the item analysis are shown in Figure 8. We found that the mean perceptual ratings differed across target nonwords, with the mean ratings ranging from 2.3 for *emplifervent* to 4.8 for *prindle*. The mean perceptual ratings for 2- and 3-syllable targets tended to be higher than the mean ratings for 4- and 5-syllable targets. A one-way ANOVA revealed that overall, imitations of longer target nonwords received lower ratings than imitations of shorter target nonwords ($F(3, 16) = 4.2$, $p < .05$). However, post-hoc Tukey tests failed to reach significance.

We found a wider range of variation in the subject analysis than in the item analysis. As shown in Figure 9, the children's individual mean perceptual ratings ranged from 1.6 to 5.7. The wide range in the mean ratings for the 24 children demonstrates that, despite the fact that all of these children produced imitation responses to all of the target stimuli in the nonword repetition task, they still exhibited a great deal of individual variability.

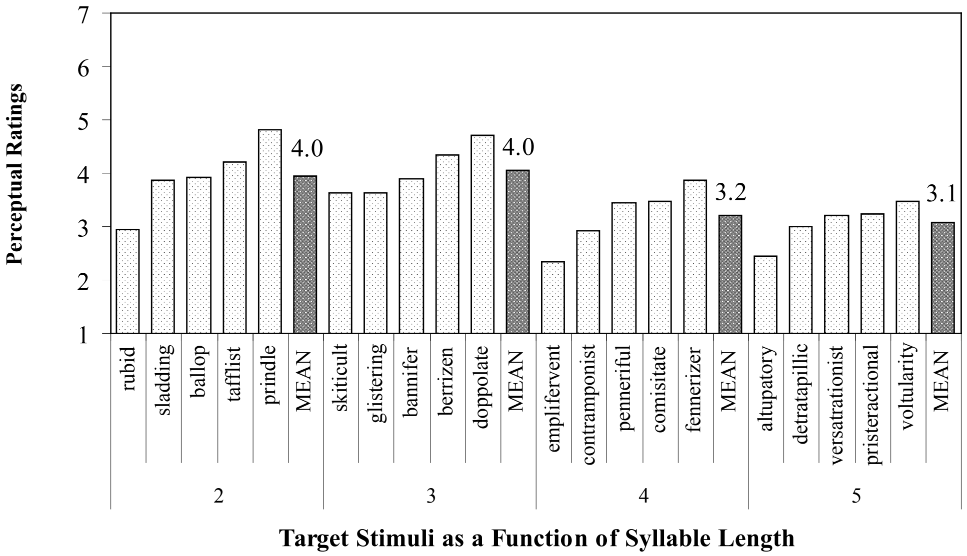

Figure 8. Mean perceptual rating for each target nonword and for each target syllable length, averaged across children.

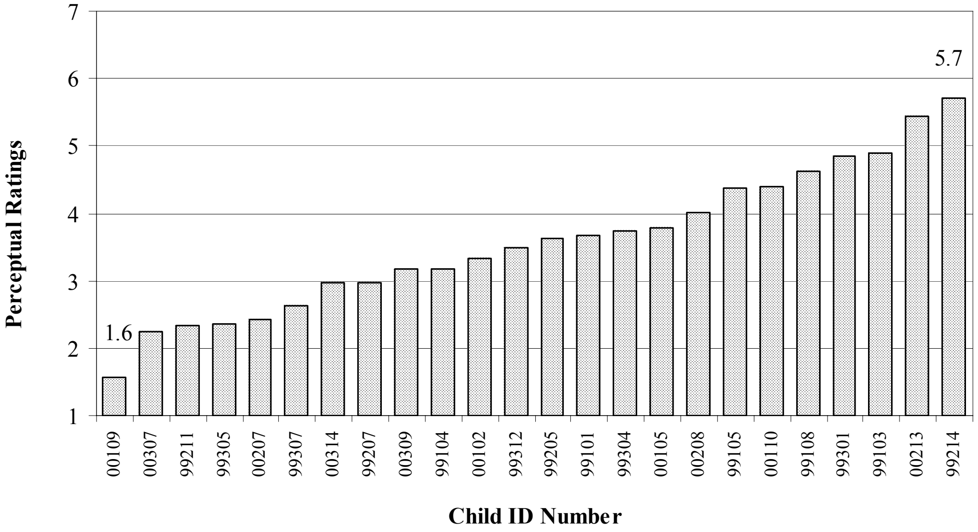

Figure 9. Mean perceptual rating for each child, averaged across target nonwords.

Correlational Analyses. We completed correlational analyses of the children's mean perceptual ratings, computing correlations between the perceptual ratings on the one hand and the children's demographic characteristics and scores on other speech and language measures on the other hand. We also examined the extent to which the children's perceptual ratings correlated with the "traditional" accuracy scores described above. Strong correlations between the ratings and the accuracy scores would indicate that the traditional segmental scoring method captured at least some of the characteristics used by normal-hearing adult listeners when evaluating the accuracy of the children's nonword repetition responses.

As shown in Table 6, we found that the children's mean perceptual ratings were moderately correlated with their age at onset of deafness ($r = +.44$, $p < .05$). This indicates that children who became deaf at later ages produced nonword repetition responses that were more accurate, based on the behaviorally-based perceptual ratings, than children who became deaf at younger ages or who were congenitally deaf. None of the other demographic variables was significantly correlated with the perceptual ratings. We suspect that at least some of the correlations between the other demographic variables and the perceptual ratings would have reached significance with a less homogeneous group of children.

	Correlation *r* values
Demographic Variables	Perceptual Ratings
Age at onset of deafness	+.44*
Duration of deafness	- .27
Age at implantation	+.07
Duration of cochlear implant use	+.04
Age at time of testing	+.18
Gender	+.14
Number of Active Electrodes	+.01
Degree of exposure to oral-only communication	+.21

Table 6. Correlations between the 24 children's demographic characteristics and their mean perceptual ratings. *$p < .05$

We also ran correlations between the children's mean perceptual ratings and their scores on measures of the component processes involved in nonword repetition. As shown in Table 7, there were strong correlations between the perceptual ratings and all of the children's spoken word recognition scores on the WIPI, LNTe, LNTh, mLNT, and BKB ($r = +.71$, +.74, +.68, +.76, +.73, respectively; $p < .001$); their language comprehension scores on the TACL-R ($r = +.62$, $p = .001$); and their phonological working memory span as measured by the forward digit span task ($r = +.78$, $p < .001$). The McGarr Intelligibility scores were also strongly correlated with the perceptual ratings ($r = +.77$, $p < .001$), and the McGarr sentence durations were negatively correlated with the perceptual ratings ($r = -.62$, -.83, -.84 for the 3-, 5-, and 7-syllable sentences, respectively; $p \leq .01$). As with the traditional accuracy scores discussed above, the perceptual ratings were not significantly correlated with the children's VIDSPAC scores nor with their backward digit spans.

Outcome and Process Measures of Performance	Correlation *r*-values Perceptual Ratings
Word Recognition	
Word identification, closed-set, pointing response	
Word Intelligibility by Picture Identification (WIPI)	+.71**
Word identification, open-set, spoken repetition	
Lexical Neighborhood Test, Easy Words (LNTe)	+.74**
Word identification, open-set, spoken repetition	
Lexical Neighborhood Test, Hard Words (LNTh)	+.68**
Word identification, open-set, spoken repetition	
Lexical Neighborhood Test, Multisyllabic Words (mLNT)	+.76**
Sentence identification, open-set repetition of target sentence	
Bamford-Kowal-Bench (BKB)	+.73**
Language Comprehension	
Closed set, receptive language comprehension of words and sentences	
Test of Auditory Comprehension of Language-Revised (TACL-R)	+.62**
Phonetic Feature Discrimination	
Perception of speech pattern contrasts	
Video Game for Assessing speech pattern contrasts (VIDSPAC)	+.35
Working Memory	
Forward Digit Span	
Wechsler Intelligence Scale for Children (WISC-III) Auditory Digit Span Subtest, Forward Recall of Digit-Name Lists	+.78**
Backward Digit Span	
Wechsler Intelligence Scale for Children (WISC-III) Auditory Digit Span Subtest, Backward Recall of Digit-Name Lists	+.37
Speech Production	
Speech Intelligibility	
McGarr Sentence Intelligibility Test	+.77**
Speaking Rate	
McGarr Mean Duration of 3-Syllable Sentences (log, msec)	-.62**
McGarr Mean Duration of 5-Syllable Sentences (log, msec)	-.83**
McGarr Mean Duration of 7-Syllable Sentences (log, msec)	-.84**

Table 7. Correlations between the 24 children's scores on several outcome and process measures, and their mean perceptual ratings. $**p \leq .001$

'Distance' Scores from Computational Analyses

Scoring. Because the process of scoring the nonword imitations for segmental accuracy is time-consuming, we were interested in automating the scoring process. In order to do this, we developed a MATLAB (Mathworks, 2000) program based on a sequence comparison algorithm described by Sankoff and Kruskal (1983). The transcription of each imitation response was compared with a transcription of the original target non-word. In order to compare the two sequences of segments, they first had to be appropriately aligned. For a relatively accurate imitation that contained the same number of segments as the target, such an alignment is straightforward: the first imitation segment can be aligned with the first target segment, the second imitation segment can be aligned with the second target segment, and so on. However, many of the nonword responses contained either additional segments that did not have corresponding segments in the target (i.e., insertions), or involved omission of one or more target segments (i.e., deletions). The program therefore needed to compute the optimal alignment between the stimulus and response before computing distance scores. In order to compute this optimal alignment, we used a stimulus-response confusion matrix. (A portion of a confusion matrix is shown in Table 9.) The confusion matrix consisted of a table in which all target segments were listed in the top row, and all possible responses (i.e., all English phonemes) were listed in the leftmost column. We also included an additional column to represent inserted segments and an additional row to represent deleted segments. Each cell in the confusion matrix contained a weight which represented the distance between a given target segment and a given response segment. In this way, the matrix contained all of the possible phoneme substitution, insertion and deletion weights. Because each weight represented the distance between two segments, larger weights represented greater distances. Therefore, a perfect imitation of a target segment was assigned a weight of “0.” Additionally, because we were primarily interested in investigating consonants in the present study, any imitation vowel given in response to any target vowel was assigned a weight of “0.” In the case of substitutions, deletions and insertions, the weights, or distances between each stimulus and response were calculated in three different ways. These three methods of calculating the weights will be referred below to as “Greenberg and Jenkins” (GJ), “Shipman and Zue” (SZ) and “McLennan” (McL).

Greenberg and Jenkins (GJ). The GJ distances were based on Greenberg and Jenkins’ (1964) method of calculating the distance between two segments based on their similarity in terms of linguistic features (place, manner, and voicing). In order to calculate the GJ distances in the present study, we compared each target segment to each potential response segment. Each pair of segments was assigned a distance weight of 0-3 points, based on whether or not the two segments were the same or different in terms of place, manner, and voicing. The pair was assigned one point for each of the features in which they differed. Thus, if the two segments had the same place, manner, and voicing, then the GJ distance between them was 0. If they differed only in terms of one of the three features, then the GJ distance between them was 1. If they differed in terms of 2 features, then the GJ distance between them was 2. If they differed in terms of all 3 features, then the GJ distance between them was 3. Table 8 shows the English consonants grouped according to their linguistic features. Three gross place features (labial, coronal, and dorsal) are listed in the top row, and 6 manners of articulation are listed in the leftmost column. The English consonant inventory includes both voiceless and voiced stops, fricatives, and affricates. In Table 8, the voiceless consonants of a given place and manner appear above the dotted line in each cell, and the voiced consonants appear below the dotted line. The linguistic features of the consonants shown in Table 8 were used to form the GJ distance matrix.

		Labial	Coronal	Dorsal
Stop	voiceless	p	t	k, ʔ
	voiced	b	d, ɾ	g
Fricative	voiceless	f	θ, s, ʃ	h
	voiced	v	ð, z, ʒ	---
Affricate	voiceless	---	t͡ʃ	---
	voiced	---	d͡ʒ	---
Nasal	voiced	m	n	ŋ
Liquid	voiced	---	l, r	---
Glide	voiced	w	j	---

Table 8. The English consonants used in this study, grouped according to their manner of articulation (stop, fricative, affricate, nasal, liquid, glide), their place of articulation (labial, coronal, or dorsal), and their voicing.

A sample portion of the GJ distance matrix is shown in Table 9. For example, the GJ distance between the stop /p/ and the nasal /m/ is 2, because /p/ and /m/ differ in terms of 2 features. These two segments have the same place of articulation (labial), but /p/ is voiceless and a stop while /m/ is voiced and a nasal consonant.

Imitation Segment	Target Segment									
	p	**b**	**m**	**f**	**v**	**t**	**d**	**a**	**i**	…
p	0	1	2	1	2	1	2	4	4	
b	1	0	1	2	1	2	1	4	4	
m	2	1	0	2	1	3	2	4	4	
f	1	2	2	0	1	2	3	4	4	
v	2	1	1	1	0	3	2	4	4	
t	1	2	3	2	3	0	1	4	4	
d	2	1	2	3	2	1	0	4	4	
a	4	4	4	4	4	4	4	0	0	
i	4	4	4	4	4	4	4	0	0	
…										

Table 9. A sample portion of the GJ distance matrix.

Shipman and Zue (SZ). The SZ distances were based on a sonority scale developed by Shipman and Zue (1982). The SZ scale used in the present study is shown in Table 10. The categories on this scale are based primarily on manner contrasts. The SZ scale shown in Table 10 was used to calculate the SZ distances, shown in the confusion matrix in Table 11. In order to calculate the SZ distances, each target segment was compared to each potential imitation segment. A pair of segments was then assigned a distance weight of 0-5 points, based on how far the two segments were from each other on the SZ scale shown in Table 10. For example, the SZ distance between the stop /p/ and the nasal /m/ is 3, as shown in Table 11, because stops and nasals are three ranks apart from each other on the SZ scale. Note that the SZ scale treats syllabic consonants as vowels.

Category	Rank	Phonemes
stops	1	p, b, t, d, k, g, ʔ
weak fricatives (non-sibilants)	2	f, θ, v, ð, h
strong fricatives (sibilants)	3	s, ʃ, z, ʒ, t͡ʃ, d͡ʒ
nasals	4	m, n, ŋ
glides and liquids	5	w, j, l, r
vowels and syllabic consonants	6	i, ɪ, e͡ɪ, ɛ, æ, a, o͡ʊ, ɔ, ʌ, ʊ, u, ɔ͡ɪ, a͡ɪ, e͡ɪ, a͡ʊ, ə, m̩, n̩, l̩

Table 10. The scale based on Shipman and Zue (1982) that was used to calculate the SZ distance weights in the present study.

Imitation Segment	Target Segment									
	p	**b**	**m**	**f**	**v**	**t**	**d**	**a**	**i**	…
p	0	0	3	1	1	0	0	5	5	
b	0	0	3	1	1	0	0	5	5	
m	3	3	0	2	2	3	3	2	2	
f	1	1	2	0	0	1	1	4	4	
v	1	1	2	0	0	1	1	4	4	
t	0	0	3	1	1	0	0	5	5	
d	0	0	3	1	1	0	0	5	5	
a	5	5	2	4	4	5	5	0	0	
i	5	5	2	4	4	5	5	0	0	
…										

Table 11. A sample portion of the SZ distance matrix.

McLennan (McL). The McL distances were based on a sonority scale developed by McLennan (2001). McLennan's scale was based on Shipman and Zue's scale and is shown in Table 12. McL differs from SZ in that its categories are based on both manner and voicing contrasts. The McL scale also treats syllabic consonants as consonants rather than as vowels. The McL scale in Table 12 was used to calculate the McL distances, shown in the confusion matrix in Table 13. In order to calculate McL distances, each target segment was compared to each potential imitation segment. Each pair of segments was assigned a distance weight of 0-10 points based on how far the two segments were from each other on the McL scale. For example, the McL distance between the stop /p/ and the nasal /m/ is 7, because stops and nasals are 7 levels apart from each other on the McL scale.

Category	Rank	Phonemes
voiceless stops	1	p, t, k, ʔ
voiced stops	2	b, d, ɾ, g
affricates	3	t͡ʃ, d͡ʒ
voiceless sibilant fricatives	4	s, ʃ
voiceless fricatives	5	f, θ, h
voiced sibilant fricatives	6	z, ʒ
voiced fricatives	7	v, ð
nasals	8	m, n, ŋ, m̩, n̩
liquids	9	l, r, l̩
glides	10	w, j
vowels	11	i, ɪ, e͡ɪ, ɛ, æ, a, o͡ʊ, ɔ, ʌ, ʊ, u, ɔ͡ɪ, a͡ɪ, e͡ɪ, a͡ʊ, ə

Table 12. The scale based on McLennan (2001) that was used to calculate the McL distance weights in the present study.

Imitation Segment	Target Segment									
	p	**b**	**m**	**f**	**v**	**t**	**d**	**a**	**i**	...
p	0	1	7	4	6	0	1	10	10	
b	1	0	6	3	5	1	0	9	9	
m	7	6	0	3	1	7	6	3	3	
f	4	3	3	0	2	4	3	6	6	
v	6	5	1	2	0	6	5	4	4	
t	0	1	7	4	6	0	1	10	10	
d	1	0	6	3	5	1	0	9	9	
a	10	9	3	6	4	10	9	0	0	
i	10	9	3	6	4	10	9	0	0	
...										

Table 13. A sample portion of the McL distance matrix.

The program used two input matrices to calculate a distance score for each imitation response. One input matrix contained the target nonwords and all of the responses, and the other was one of the three distance matrices described above. The program used a sequence comparison algorithm (Sankoff & Kruskal, 1983) to align the segments in an imitation with segments in the target in numerous ways, and calculate the distance between the response and its target. It then selected an optimal alignment from the numerous possible ways of aligning an imitation with its target. The optimal alignment was that which resulted in the minimal distance, that is, the least distance between the target and the imitation. When multiple alignments produced the same scores, the alignment selected as optimal was that alignment that had a greater number of substitutions (rather than deleted or inserted segments); if a further tie-breaker was necessary, the alignment with a greater number of insertions (rather than deletions) was selected as optimal. The output of the program was also in the form of matrices. One matrix contained the optimal alignment of each imitation with its target non-word. This matrix was used to calculate an average distance score per child.

Results. We carried out both item and subject analyses for the GJ, SZ, and McL distance scores which were similar to the item and subject analyses that we computed for the traditional accuracy scores. Because the McL distances for each segment can range from 0 to 10 while the GJ distances for each segment can only range from 0 to 4 and the SZ distances can only range from 0 to 5, the McL distance scores per imitation were higher than the GJ and SZ scores.

The item analyses for each of the three distance scores are shown in Figures 10, 11, and 12, respectively. The mean GJ scores for the target items ranged from 5.9 for the target *doppolate* to 22.4 for the target *pristeractional*. The mean SZ scores ranged from 5.9 for the target *doppolate* to 21.3 for both *pristeractional* and *contramponist*. The mean McL scores ranged from 11.4 for *bannifer* to 40.4 for *pristeractional*. For all three of the distance measures, the mean distance scores for 2- and 3-syllable targets tended to be lower than the mean distance scores for 4- and 5- syllable targets. Three one-way ANOVA's revealed that there was a significant main effect of syllable length on the distance scores (GJ: $F(3, 16) = 7.0$; SZ: $F(3, 16) = 7.9$; McL: $F(3, 16) = 7.9$; $p < .01$). Post-hoc Tukey tests revealed that 2- and 3-syllable targets had lower distance scores than 5-syllable targets ($p < .01$), for all 3 distance scores.

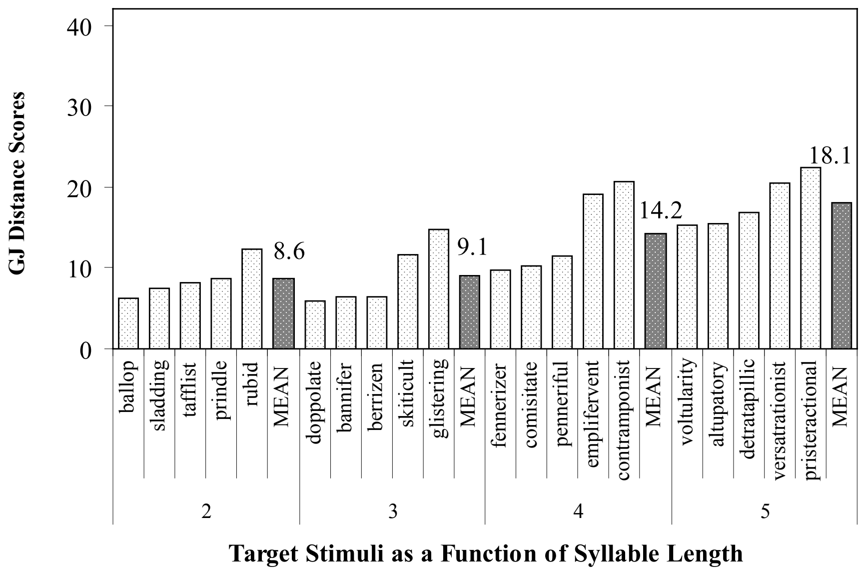

Figure 10. GJ distance score for each target nonword and for each target syllable length, averaged across children.

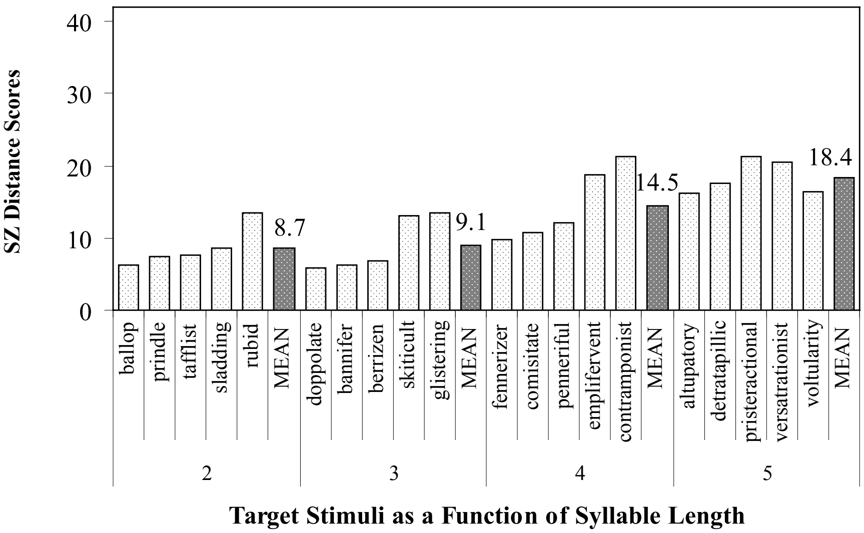

Figure 11. SZ distance score for each target nonword and for each target syllable length, averaged across children.

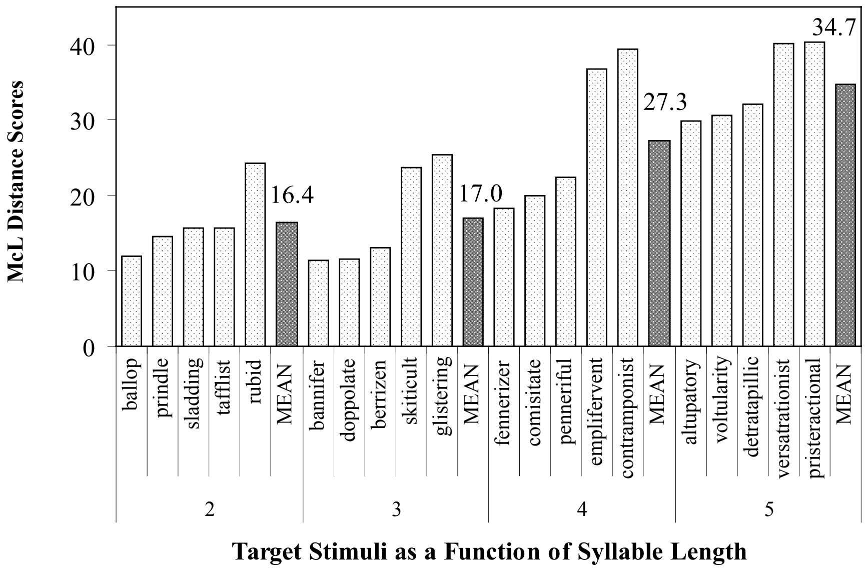

Figure 12. McL distance score for each target nonword and for each target syllable length, averaged across children.

The results of the subject analyses for the GJ, SZ and McL distance scores are shown in Figures 13, 14, and 15, respectively. The children exhibited variability in their imitation scores regardless of which of the three distance measures were used to evaluate their nonword repetition performance. For the GJ measure, the children's mean scores ranged from 4.6 to 20.4. For the SZ measure, the children's mean scores ranged from 4.8 to 21.4. For the McL measure, the children's mean scores ranged from 9.5 to 40.4.

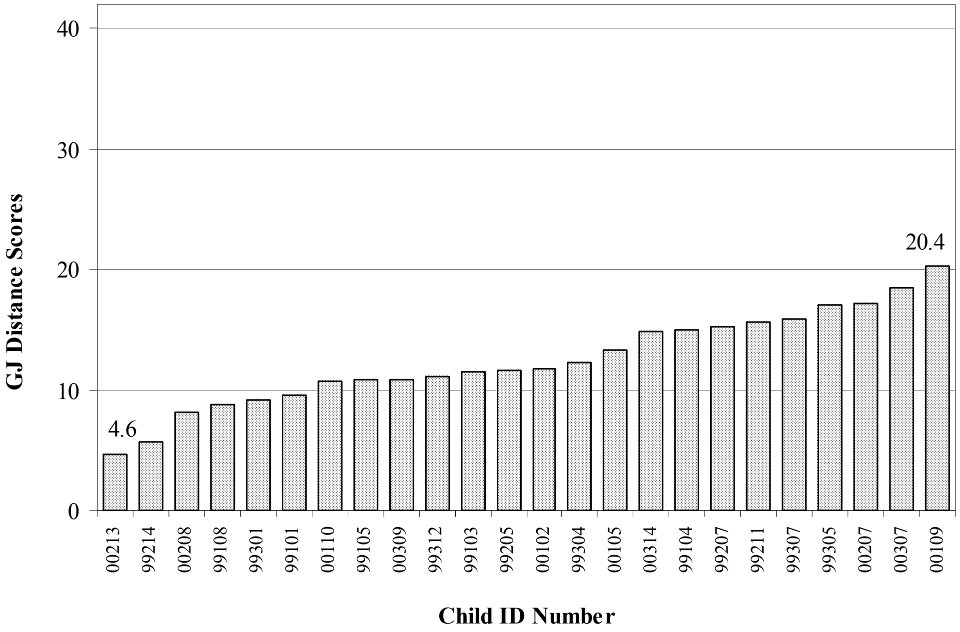

Figure 13. GJ distance score for each child, averaged across target nonwords.

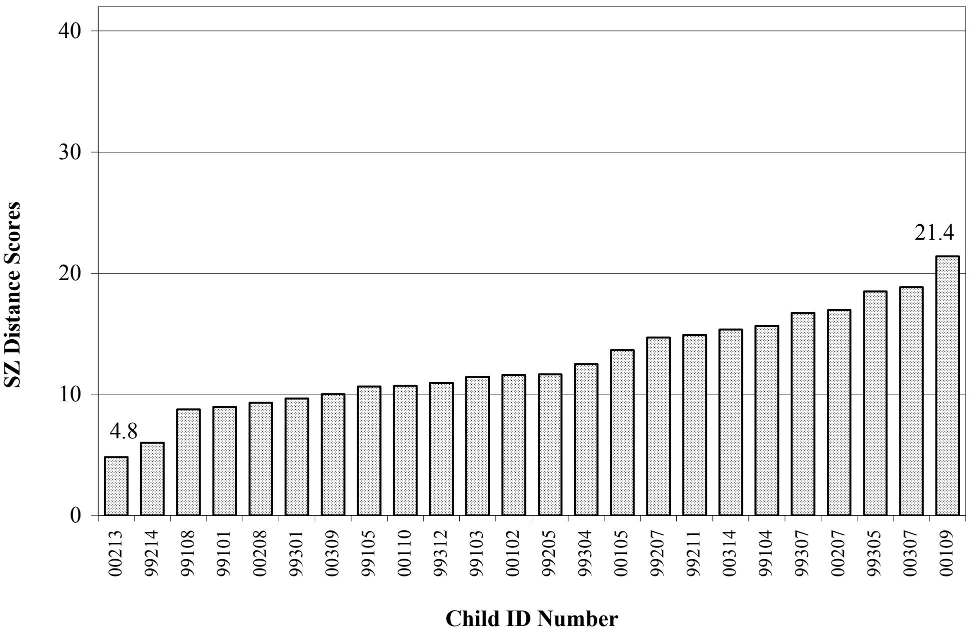

Figure 14. SZ distance score for each child, averaged across target nonwords.

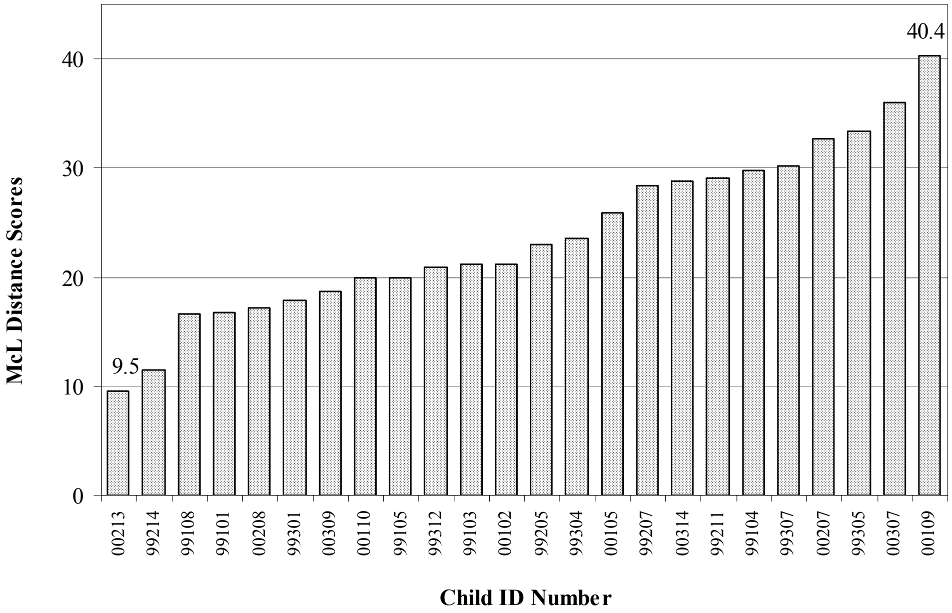

Figure 15. McL distance score for each child, averaged across target nonwords.

Correlational Analyses. Despite the small differences in the confusion matrices for the GJ, SZ, and McL methods of measuring distance, the similarity among the patterns of results shown in Figures 10-12 and 13-15 indicate that the three methods of calculating distance scores (GJ, SZ, and McL) were all capturing the same featural aspects of the imitations as compared to the targets. In order to confirm that the 3 distance scores were indeed comparable measurements, we computed correlations among the children's 3 distance scores. The correlations, shown in Table 14, were nearly perfect ($r = +.99$, $p < .001$), indicating that the slight differences between the GJ, SZ, and McL scoring methods are not great enough to regard them as distinct measures. The finer-grained 11-rank McL scale does not provide a more meaningful distinction among the imitations than the broader GJ and SZ scales. Because the GJ, SZ, and McL distance scores were so highly intercorrelated, we report only correlations with the GJ scores for the additional correlational analyses below.

	GJ	SZ	McL
GJ		+.99**	+.99**
SZ			+.99**
McL			

Table 14. Correlations among the 3 distance scores (GJ, SZ, and McL). **p <.001

We also ran correlations between the children's GJ distance scores and their demographic characteristics. As shown in Table 15, there was a moderate negative correlation between the children's age at onset of deafness and GJ distance scores ($r = -.42$, $p < .05$). This negative correlation indicates that the children who became deaf at a later age tended to have lower distance scores. Thus, they performed better on the nonword repetition task than the children who were congenitally deaf or who became deaf at a younger age. Although not shown in Table 15, the correlations between age at onset of deafness and both the SZ and McL distance scores were negative but failed to reach significance. None of the other demographic variables were significantly correlated with the distance scores.

	Correlation *r* values
Demographic Variables	GJ Distance Scores
Age at onset of deafness	- .42*
Duration of deafness	+.21
Age at implantation	- .13
Duration of cochlear implant use	+.01
Age at time of testing	- .20
Gender	+.09
Number of Active Electrodes	+.15
Degree of exposure to oral-only communicatioı	- .10

Table 15. Correlations between the 24 children's demographic characteristics and their mean GJ distance scores. *$p < .05$

Correlations between the children's GJ distance scores and several measures of the component processes involved in nonword repetition are shown in Table 16. The measures of the component processes in Table 16 are the same speech and language outcome and process measures described in the "Traditional Accuracy Analyses" section above. As shown in Table 16, there were strong negative correlations between the children's GJ distance scores and all of the spoken word recognition measures ($r = -.72, -.70, -.71, -.73. -77$, $p < .001$), the language comprehension measure ($r = -.74$, $p < .001$), and forward digit span ($r = -.69$, $p < .001$). Because the GJ scores are 'distance' scores, lower scores indicate better performance. Therefore, the negative correlations between the distance scores and the spoken word recognition scores indicate that the children who performed well on the nonword repetition task also tended to perform well on the spoken word recognition tasks.

			Correlation *r*-values
Outcome and Process Measures of Performance			GJ Distance Scores
Word Recognition			
	Word identification, closed-set, pointing response		
		Word Intelligibility by Picture Identification (WIPI)	- .72**
	Word identification, open-set, spoken repetition		
		Lexical Neighborhood Test, Easy Words (LNTe)	- .70**
	Word identification, open-set, spoken repetition		
		Lexical Neighborhood Test, Hard Words (LNTh)	- .71**
	Word identification, open-set, spoken repetition		
		Lexical Neighborhood Test, Multisyllabic Words (mLNT)	- .73**
	Sentence identification, open-set repetition of target sentence		
		Bamford-Kowal-Bench (BKB)	- .77**
Language Comprehension			
	Closed set, receptive language comprehension of words and sentences		
		Test of Auditory Comprehension of Language-Revised (TACL-R)	- .74**
Phonetic Feature Discrimination			
	Perception of speech pattern contrasts		
		Video Game for Assessing speech pattern contrasts (VIDSPAC)	- .32
Working Memory			
	Forward Digit Span		
		Wechsler Intelligence Scale for Children (WISC-III) Auditory Digit Span Subtest, Forward Recall of Digit-Name Lists	- .69**
	Backward Digit Span		
		Wechsler Intelligence Scale for Children (WISC-III) Auditory Digit Span Subtest, Backward Recall of Digit-Name Lists	- .34
Speech Production			
	Speech Intelligibility		
		McGarr Sentence Intelligibility Test	- .80**
	Speaking Rate		
		McGarr Mean Duration of 3-Syllable Sentences (log, msec)	+.67**
		McGarr Mean Duration of 5-Syllable Sentences (log, msec)	+.88**
		McGarr Mean Duration of 7-Syllable Sentences (log, msec)	+.87**

Table 16. Correlations between the 24 children's scores on several outcome and process measures, and their mean GJ distance scores. $^{**}p < .001$

The children's scores on the measure of phonetic feature discrimination, the VIDSPAC, were not significantly correlated with their GJ distance scores. Likewise, the children's backward digit span scores were not significantly correlated their GJ distance scores. These results are again consistent with previous findings that backward digit span is not a measure of the same cognitive processes as forward digit span. The children's GJ scores were also negatively correlated with their performance on the McGarr speech intelligibility measure (r = -.80, p < .001). This correlation indicates that the children who produced imitations that were closer to the targets also tended to be more intelligible on the McGarr Intelligibility task. Lastly, the children's GJ distance scores were strongly correlated with their speaking rates for the McGarr 3-, 5-, and 7-syllable sentences (r = +.67, +.88, +.87, respectively; p < .001). These results suggest that the children who spoke more slowly in the McGarr task tended to produce nonword imitations that were more dissimilar from the targets.

In summary, the present findings show that children who produced imitations that were closer to the targets, in terms of GJ distances, also tended to score higher on a wide range of outcome and process measures that assess the component processes involved in nonword repetition. That is, they tended to have higher scores for measures of spoken word recognition, language comprehension, phonological working memory, and speech production. The three methods of evaluating the children's nonword repetition responses – the traditional accuracy scores, the perceptual ratings, and the distance scores – all produced scores that did not tend to be significantly correlated with the children's demographic characteristics, but were strongly correlated with their scores on other outcome and process measures of speech and language. The similarity between the results of all three methods indicates that these three different ways of scoring the children's imitations did not produce differential results. A strong negative correlation between the children's GJ distance scores and the segmental accuracy scores presented in the "Traditional Accuracy Analyses" section (r = -.94, p < .001) revealed that the accuracy scores and the distance scores, while based on different methods of scoring the nonword repetitions, both resulted in similar patterns of nonword repetition performance among the children. As shown in Table 17, correlations between the perceptual ratings and the segmental accuracy scores, the GJ distance scores, the SZ distance scores, and the McL distance scores (r = +.90, -.92, -.90, -.90, respectively; p < .001), also indicate that these different methods resulted in scores that showed a similar distribution of performance among the children. The strong correlations shown in Table 17 between the perceptual ratings and the other scoring methods indicates that the other scoring methods captured the characteristics used by normal-hearing adult listeners when evaluating the accuracy of the children's nonword repetition responses. These strong correlations indicate that the perceptual ratings given to the imitations were influenced by whether or not the imitation consonants were correct, and by how close the incorrect imitation segments were to the target segments. The results suggest that the segmental similarity between the target and the imitation segments was the primary factor that the listeners' attended to and based their perceptual ratings on.

	Perceptual Ratings
Segmental Accuracy Scores	+.90**
GJ Distance Scores	- .92**
SZ Distance Scores	- .90**
McL Distance Scores	-.90**

Table 17. Correlations between the 24 children's mean segmental accuracy scores, GJ distance scores, SZ distance scores, McL distance scores, and their mean perceptual ratings. **p < .001

General Discussion

Studies of the speech and language skills of deaf children who use cochlear implants consistently report a wide range of individual variability on clinical outcome measures of speech and language (Kirk, 2000). Moreover, although previous studies have found that about 35-65% of the individual differences in pediatric CI users' performance can be explained by demographic variables such as duration of deafness, length of device use and age at implantation (Blamey et al., 2001; Dowell, Blamey, & Clark, 1995; Miyamoto et al., 1994; Sarant, Blamey, Dowell, Clark, & Gibson, 2001; Snik, Vermeulen, Geelen, Brokx, & van den Broek, 1997), a substantial portion of the variation among children remains to be explained (Pisoni, 2000). We are interested in understanding the factors that are responsible for these individual differences. If we can identify the variables that underlie these individual differences in performance, we may be in a position to recommend specific changes that will help the poorer performing children achieve optimal levels of speech and language performance with their cochlear implants.

In trying to understand the sources of variability in outcome measures obtained with the pediatric CI users, we have considered several areas of research on individual differences in the language development of young normal-hearing children. A number of pieces of evidence suggest that individual differences in phonological coding and verbal working memory contribute to variability in skills such as vocabulary acquisition and language among normally-developing children (Gathercole & Baddeley, 1993; Gathercole & Adams, 1993). The task most widely used over the last decade to study phonological processing skills and verbal working memory is the nonword repetition task employed in the present study.

As demonstrated in the present set of results, although performance on the nonword repetition task clearly reflects individual differences in the component processes of sensory encoding, maintenance of phonological representations in working memory, and speech production, we think it is useful to include this nonword repetition measure among other outcome measures in the study of deaf children with cochlear implants. Inclusion of a nonword processing task is potentially important because unlike other routine clinical outcome measures, this particular imitation task may reasonably be argued to additionally reflect a child's ability to rapidly transform sensory input into what he/she perceives to be an "equivalent" articulatory-motor output. This skill, we would argue, is something separable from perception and production, and reflects a series of crucial phonological processing operations involving decomposition and translation of a sensory-perceptual representation into a phonological representation and then reassembly of a phonological representation into an articulatory program used in speech production (see for discussion, Pisoni & Cleary, in press).

The ability of pediatric CI users to utilize their knowledge of the phonological patterns of their ambient language in order to reproduce spoken nonword stimulus patterns has not, to our knowledge, been previously explored in any great detail. Indeed, some research suggests that the phonological representations of this clinical population are so "fragile" as to make any perceptual task that involves items for which the children have no learned lexical representation, very difficult (Kirk et al., 1995). However, because all spoken words that children with cochlear implants learn to recognize must have been, at one point, nonwords to these children, we believe that it may be informative to try to understand how individual children in this clinical population process such novel auditory stimuli.

Use of a nonword repetition task with hearing-impaired children does, however, raise a number of complicated issues regarding how to interpret performance on this task and the high error rates that were observed. The present project was undertaken with the conviction that if a detailed characterization of the error patterns could be accomplished, we would then be able to investigate the relationship between different component processes to performance. We therefore carried out a linguistic analysis of the responses based on the phonological features of the target phonemes. We attempted to account for observed patterns of performance through comparisons with previous results from this population on tasks that involve primarily perception or primarily speech production. Relations between perception and production and nonword repetition performance were also examined by looking at intercorrelations between nonword repetition performance and several independently obtained measures of speech perception and articulation.

Our ability to conduct these analyses was, rather obviously, predicated on the assumption that most children sampled from in this clinical population would be able to do the nonword repetition task well enough to obtain a measurable score. Fortunately, we did find that even when nonword stimuli were presented to pediatric CI users in an auditory-only mode, almost all of the children were able to provide some attempt at a response to at least 75% of all trials. The children were clearly not at floor on this task, regardless of how their responses were scored, as measured by traditional segmental accuracy scores, or through behaviorally-based perceptual ratings obtained from naïve listeners or algorithmically using computational analyses.

Because vowel quality is difficult to categorize and varies considerably between the regional dialects present in our sample (Wolfram & Schilling-Estes, 1998), our segmental linguistic analyses focused exclusively on consonant production. Detailed examination of the children's consonant feature production revealed several "null results" with regard to the manner and voicing features of target segments. First, we found that the children did not perform better in response to target consonants of particular manners of articulation (e.g., no advantage was observed for stops over fricatives). Second, we also found no evidence that the voicing feature was being produced correctly more often for either voiced or voiceless target consonants. These results were not surprising given earlier findings, such as those described in the Introduction. Neither Dawson et al. (1995) nor Sehgal et al. (1998) found large differences between children's ability to produce consonants of different manners or voicing.

However, with regards to place of articulation, we found that coronal consonants were imitated correctly by the children significantly more often than labial and dorsal consonants. This finding conflicts with results of previous studies by Dawson et al. (1995) and Sehgal et al. (1998), as well as Tobey et al. (1994). These researchers reported that labial consonants were produced correctly more often than consonants with other places of articulation. The tasks employed in Dawson et al. and Sehgal et al. were, however, open set word recognition tasks in which the children were asked to repeat real words that were presented to them. Thus, the stimuli and responses in these two studies were real English words, while the stimuli and responses in the present study were nonwords.

The children in our study had not been exposed to the nonword stimuli before, and therefore they did not have any opportunity to benefit from the salient visual cues provided by the lips when a target labial consonant is produced. It is possible that the children in Dawson et al. (1995) and Sehgal et al. (1998) produced labial consonants correctly more often than other consonants because they had had previous opportunities to learn the real words that served as stimuli, and to benefit from the visual cues provided by the salient lip closure of labial consonants. In addition, the stimuli in Dawson et al. and Sehgal et al. were presented to the children live-voice by an examiner. In our study the children heard an auditory-only recording of the nonword stimuli presented over a loudspeaker. A live-voice auditory-visual presentation format was also used in the Tobey et al. (1994) study. Thus, the children in these previous studies may have been able to benefit from the presence of visual cues, especially the lip closure of labial consonants.

When children have access to visual cues to speech in addition to the auditory information in the acoustic signal, imitating labials may surpass coronals. However, when visual cues are not available, as in the present study, it appears to be that perception of the labial consonants suffers. The children correctly imitate coronals more often than labials or dorsals. Our results are consistent with previous findings in the literature demonstrating differences due to presentation format. For example, in a recent study of 27 pediatric CI users, Lachs, Pisoni, and Kirk (2001) found that audiovisual stimulus presentation produced better overall speech perception performance than auditory-only stimulus presentation. This result reflects the presence of visual cues to place of articulation in the audiovisual condition, reliable speech cues that were not available in the present auditory-only nonword repetition task.

Because the traditional linguistic segmental accuracy scoring methods used in the present paper were extremely time-intensive, we developed an automated scoring method based on a sequence comparison algorithm. The strong correlations observed between the traditional segmental accuracy scores scored “by hand” and the automated scoring method validate the use of this procedure in the future when calculating distance scores between pairs of transcriptions. We found that the automated, computational scoring methods yielded the same pattern of results as the traditional scoring method. These new methods therefore provide reliable estimates of the segmental proximity of an imitation transcription compared to a target transcription, using traditional linguistic feature-based categories.

From a theoretical point of view, we found it interesting that the segmental accuracy scores were highly correlated with the distance scores obtained from the computational analysis. The accuracy scores only gave credit to segments that were correct, while the distance scores were continuous and gave partial credit to incorrect segments. Thus, the distance scores were a finer-grained measure of nonword repetition performance, in that they gave partial credit to an imitation segment for being ‘close’ to the target, in contrast to the traditional segmental accuracy scores, which only gave credit to an imitation segment if it was completely correct. Nevertheless, we found that the finer-grained distance scores were highly correlated with the broader binary scoring method used in the traditional segmental accuracy scores. This indicates that the pattern of individual differences revealed by the finer-grained distance scores was not different from the pattern of individual differences revealed by the broader accuracy scores.

In addition to the wide range of variability observed across individual children, we also found variability across the nonword targets based on their syllable lengths. The children’s nonword repetition performance was significantly affected by the syllable length of the target. Imitations of the 2- and 3-syllable target nonwords received significantly better scores on all three measures of performance than imitations of the 5-syllable target nonwords. The children’s imitations of 2- and 3-syllable target nonwords also tended to be better than their imitations of 4-syllable target nonwords, but these differences did not reach significance. This finding is consistent with our earlier results on the suprasegmental aspects of the children’s nonword repetitions (Carter et al., in press). In that study, we found that children produced suprasegmental features of their imitations correctly in response to shorter target nonwords more often than longer target nonwords. Our results also replicate the earlier findings reported by Gathercole (1995), in which normal-hearing children produced correct imitations of shorter nonwords more often than longer nonwords. Longer nonwords appear to place a greater load on phonological working memory and the verbal rehearsal processes of the phonological loop that maintains information for brief periods of time.

Further correlational analyses conducted to account for individual differences in nonword repetition performance among the children, revealed several additional findings. Variability among the children in nonword repetition performance was related in a systematic manner to their performance on other tasks designed to measure the phonological processes similar to those used in carrying out the nonword repetition task: spoken word recognition and speech perception, language comprehension, phonological working memory, and speech production. Overall, better nonword repetition performance (in terms of either segmental accuracy, distance scores, or perceptual ratings) was associated with higher spoken word recognition scores, higher language comprehension scores, longer forward digit spans, higher speech intelligibility scores and faster speaking rates in the McGarr task. The strong correlation between the children's nonword repetition performance and their forward digit spans, a measure of phonological working memory capacity, is not surprising given that both measures were developed to measure similar processing skills. The strong intercorrelations between these tasks has not, however, been previously reported for this clinical population.

The observed correlations between nonword repetition performance and several measures that are primarily dependent on speech perception abilities – phonetic feature discrimination, word recognition, and language comprehension – indicate that the children who had better speech perception skills also had higher nonword repetition scores. The correlations with nonword repetition were stronger for the word recognition measures (WIPI, LNTe, LNTh, mLNT, and BKB) than for the phonetic feature discrimination measure (VIDSPAC). The speech feature discrimination skills tapped by the VIDSPAC measure appear to play a minimal role in nonword repetition.

The correlations between nonword repetition and several measures of speech production were also strong and statistically significant. The correlations observed between nonword repetition performance and the measure of speech intelligibility provide additional support for earlier studies that showed that speech intelligibility is related to the segmental accuracy of hearing-impaired persons' speech (Parkhurst & Levitt, 1978; Smith, 1975; Tobey et al., 1991, p.165, cites Levitt & Stromberg, 1983). The correlations between nonword repetition and speaking rate measured independently from sentence durations, however, require a somewhat different interpretation. In this case, the results reflect factors having to do with individual differences in the speed and efficiency of phonological processing, specifically verbal rehearsal. Our finding that children with faster speaking rates tended to have higher nonword repetition scores is consistent with the view that speaking rate is an overt reflection of the speed with which phonological information can be subvocally rehearsed in short-term working memory (see Baddeley, 1986; Burkholder, this volume; Landauer, 1962; Pisoni & Cleary, in press). Verbal rehearsal speed is well known to affect a person's ability to hold items in phonological working memory (Flavell, Beach, & Chinsky, 1966; McGilly & Siegler, 1989). Thus, the present findings indicate that individual differences in verbal rehearsal rate as indexed by sentence duration are associated with nonword repetition performance.

The correlations between the measures of the component processes of nonword repetition and performance on the nonword repetition task itself, as reviewed above, are consistent with results from our recent study of these same children's suprasegmental repetition accuracy (see Carter et al., in press). In that study, the children who performed better in terms of imitating prosodic characteristics of the targets (number of syllables and primary stress placement) also tended to have higher scores on other outcome and processing measures of speech and language. The strong relation among all of the children's scores reflect the close correspondence between the children's speech perception, phonological working memory, speech production, and language skills. The correlations between the children's nonword repetition performance and all of the speech and language outcome and process measures suggests that the nonword repetition task can be viewed as serving as a composite diagnostic measure of the children's ability to access and integrate these component skills in performing this repetition task. The strong correlations between nonword repetition performance and scores on the other speech and language measures indicate that the children who performed well on separate measures of the component processes also performed well on tasks that combine these skills together.

It is of some clinical and theoretical interest that the demographic characteristics of the children did not influence their ability to correctly imitate consonants. We suspect that the homogeneity of the children in terms of demographics in this study may have prevented the correlations between the demographic characteristics and the accuracy scores from reaching significance. In the present study, 18 children used oral communication while only 6 used total communication methods. As mentioned earlier, in our larger study of 76 children who completed the nonword repetition task that included the 24 children discussed in the present paper, we found that perceptual ratings of the nonword responses were strongly correlated with the children's communication mode scores. The children who had been exposed to primarily oral methods of communication tended to perform better on the nonword repetition task than the children who primarily used total communication methods.

In planning future research using nonwords, it may be useful to consider several limitations of the present study. The nonword stimuli used in this study were originally designed to measure phonological working memory skills (Gathercole & Baddeley, 1996) and may not be appropriate for investigating consonant production in children. These stimuli are not balanced in terms of phonemes, phonotactics, or stress, and are not representative of the phonological patterns of the ambient language that these children are exposed to. Thus, it is possible that the children's performance in terms of consonant segment and feature accuracy was affected by the specific phonological sequences of the target nonwords. We are now designing a new set of nonword stimuli in which other phonological variables such as target phonemes and features, phonotactic patterns, and stress patterns are carefully controlled and balanced. Such a set of nonword stimuli will allow us to investigate the influence of phonological characteristics other than syllable number, and to examine more systematically whether the target nonword patterns that the children reproduce correctly and the errors that they make are consistent with those that would be expected in a typically-developing phonological system of younger normal-hearing children.

Finally, the analyses reported here were based on utterances obtained from an imitation study, which may naturally lead to questions about the generalizability of our results to the children's spontaneous or elicited speech production performance. Although we did not analyze spontaneous or non-imitated elicited productions in these children, it should be noted that in an earlier longitudinal study, Tobey et al. (1991) found that children's productions of both imitated speech and elicited spontaneous speech improved with increased implant use, suggesting that a common set of underlying phonological skills is used in both imitation and spontaneous speech production.

Summary and Conclusions

In summary, the present investigation analyzed the imitation responses of 24 deaf children with cochlear implants who completed a nonword repetition task. The children demonstrated a wide range of skills in imitating unfamiliar but 'word-like' targets by using their knowledge of the phonological patterns present in their ambient language. The children's better performance on coronals than labials or dorsals in this study suggests that other reports of superior performance on labials may be the result of perceptual enhancement due to the presence of visual cues or children's prior lexical knowledge. The strength of the correlations between the accuracy scores and direct perceptual ratings of the imitations by naive listeners suggests that the use of converging methods of linguistic analyses, such as those employed in the present study, can help us to understand which aspects of the children's imitations listeners attend to when judging accuracy. Taken together, the results of this study demonstrate that the nonword repetition task can provide fundamental new insights about the speech perception and speech production skills and underlying phonological processing abilities of pediatric cochlear implant users. With further analytic studies of this type, we hope to better understand the relations between auditory, cognitive, and linguistic processes used in the perception and production of spoken language, and how these develop and change in deaf children over time following cochlear implantation. Moreover, we hope to be able to explain the reasons for the enormous variability in outcome measures observed in this unique clinical population.

References

Avons, S.E., Wragg, C.A., Cupples, L., & Lovegrove, W.J. (1998). Measures of phonological short-term memory and their relationship to vocabulary development. *Applied Psycholinguistics, 19*, 583-601.

Baddeley, A. (1986). *Working Memory*. Oxford, UK: Clarendon.

Bench, J., Kowal, A., & Bamford, J. (1979). The BKB (Bamford-Kowal-Bench) sentence lists for partially-hearing children. *British Journal of Audiology, 13*, 108-112.

Blamey, P.J., Sarant, J.Z., Paatsch, L.E., Barry, J.G., Bow, C.P., Wales, R.J., et al. (2001). Relationships among speech perception, production, language, hearing loss, and age in children with impaired hearing. *Journal of Speech, Language, and Hearing Research, 44*, 264-285.

Boothroyd, A. (1997). Vidspac 2.0: A video game for assessing speech pattern contrast perception in children. San Diego, Arthur Boothroyd.

Burkholder, R.A. (this volume). Speech timing measures and working memory in normal-hearing children and deaf children with cochlear implants. In *Research on Spoken Language Processing Progress Report No. 25*. Bloomington, IN: Speech Research Laboratory, Indiana University.

Carlson, J.L., Cleary, M., & Pisoni, D.B. (1998). Performance of normal-hearing children on a new working memory span task. In *Research on Spoken Language Processing Progress Report No. 22* (pp. 251-273). Bloomington, IN: Speech Research Laboratory, Indiana University.

Carrow-Woolfolk, E. (1985). *Test for Auditory Comprehension of Language-Revised (TACL-R).* Austin, TX: Pro-Ed.

Carter, A.K., Dillon, C.M., & Pisoni, D.B. (in press). Imitation of nonwords by hearing-impaired children with cochlear implants: Suprasegmental analyses. *Clinical Linguistics and Phonetics*.

Chin, S.B., Kirk, K.I., & Svirsky, M.A. (1997). Sensory aid and word position effects on consonant feature production by children with profound hearing impairment. In *Research on Spoken Language Processing Progress Report No. 21* (pp. 455-470). Bloomington, IN: Speech Research Laboratory, Indiana University.

Chin, S.B., Pisoni, D.B., & Svec, W.R. (1994). An emerging phonetic-phonological system two years post cochlear implant: A preliminary linguistic description. In *Research on Spoken Language Processing Progress Report No. 19* (pp. 253-270). Bloomington, IN: Speech Research Laboratory, Indiana University.

Cleary, M., Dillon, C.M. & Pisoni, D.B. (2002). Imitation of nonwords by deaf children following cochlear implantation. *Annals of Otology, Rhinology and Laryngology, Supplement-Proceedings of the 8th Symposium on Cochlear Implants in Children, 111*, 91-96.

Cowan, N., Saults, J.S., Elliot, E.M., & Moreno, M.V. (2002). Deconfounding serial recall. *Journal of Memory and Language, 46*, 153-177.

Dawson, P.W., Blamey, S.J., Dettman, S.J., Rowland, L.C., Barker, E.J., Tobey, E.A., Busby. P.A., Cowan, R.C., & Clark, G.M. (1995). A clinical report on speech production of cochlear implant users. *Ear & Hearing, 16*, 551-561.

Dillon, C.M., Burkholder, R.A., Cleary, M. & Pisoni, D.B. (this issue). Perceptual ratings of nonword repetitions by deaf children after cochlear implantation: Correlations with measures of speech, language and working Memory. In *Research on Spoken Language Processing Progress Report No. 25.* Bloomington, IN: Speech Research Laboratory, Indiana University.

Dillon, C.M., & Cleary, M. (2000). Using nonword repetition to study speech production skills in hearing-impaired children with cochlear implants. In *Research on Spoken Language Processing Progress Report No. 24* (pp. 113-147). Bloomington, IN: Speech Research Laboratory, Indiana University.

Dowell, R.C., Blamey, P.J., & Clark, G.M. (1995). Potential and limitations of cochlear implants in children. *Annals of Otology, Rhinology & Laryngology, 166(Suppl.)*, 324-327.

Edwards, J. & Lahey, M. (1998). Nonword repetitions of children with specific language impairment: Exploration of some explanations for their inaccuracies. *Applied Psycholinguistics, 19*, 279-309.

Engle, R.W. (2002). Working memory capacity as executive attention. *Current Directions in Psychological Science, 11*, 19-23.

Ertmer, D.J., Kirk, K.I., Sehgal, S.T., Riley, A.I., & Osberger, M.J. (1997). A comparison of vowel production by children with multichannel cochlear implants or tactile aids: Perceptual evidence. *Ear & Hearing, 18*, 307-315.

Fisher, H., & Logemann, J. (1971). *Test of Articulation Competence*. New York: Houghton and Mifflin.

Flavell, J.H., Beach, D.H., & Chinsky, J.M. (1966). Spontaneous verbal rehearsal in a memory task as a function of age. *Child Development, 37*, 283-299.

Gathercole, S.E. (1995). Is non-word repetition a test of phonological memory or long-term knowledge? It all depends on the non-words. *Memory and Cognition, 23*, 83-94.

Gathercole, S.E. & Adams, A. (1993). Phonological working memory in very young children. *Developmental Psychology, 29*, 770-778.

Gathercole, S.E. & Baddeley, A.D. (1990). The role of phonological memory in vocabulary acquisition: A study of young children learning new names. *British Journal of Psychology, 81*, 439-454.

Gathercole, S.E., & Baddeley, A. (1993). *Working memory and language*. Hillsdale, NJ: Lawrence Erlbaum.

Gathercole, S.E., & Baddeley, A.D. (1996). *The Children's Test of Non-word Repetition*. London: Psychological Corporation.

Gathercole, S.E, Willis, C. S., Baddeley, A. D., & Emslie, H. (1994). The Children's Test of Non-word Repetition: A test of phonological working memory. *Memory, 2*, 103-127.

Geers, A.E., Nicholas, J., Tye-Murray, N., Uchanski, R., Brenner, C., Crosson, J., Davidson, L.S., Spehar, B., Torretta, G., Tobey, E.A., Sedey, A., & Strube, M. (1999). Center for Childhood Deafness and Adult Aural Rehabilitation, Current research projects: Cochlear implants and education of the deaf child, second-year results. In: *Central Institute for the Deaf Research Periodic Progress Report No. 35,* pp. 5-20.

Geers, A.E., & Tobey, E. (1992). Effects of cochlear implants and tactile aids on the development of speech production skills in children with profound hearing impairment. *The Volta Review, 94*, 135-163.

Greenberg, J.H. & Jenkins, J.J. (1964). Studies in the psychological correlates of the sound system of American English. *Word, 20*, 157-177.

Hesketh, L.J., Fryauf-Bertschy, H., & Osberger, M.J. (1991). Evaluation of a tactile aid and a cochlear implant in one child. *The American Journal of Otology, 12 (Suppl.)*, 183-187.

Hudgins, C.V., & Numbers, F.C. (1942). An investigation of the intelligibility of the speech of the deaf. In *Genetic Psychology Monographs, 25*, 289-392.

Kirk, K.I. (2000). Challenges in the clinical investigation of cochlear implant outcomes. In J.K. Niparko, K.I. Kirk, N.K. Mellon, A.M. Robbins, B.L Tucci, & B.S. Wilson (Eds.) *Cochlear Implants: Principles and Practices* (pp. 225-259). Philadelphia, PA: Lippincott Williams & Wilkins.

Kirk, K.I., Diefendorf, E., Riley, A., & Osberger, M.J. (1995). Consonant production by children with multichannel cochlear implants or hearing aids. In Uziel, A.S., & Mondain, M. (Eds.), *Advanced Otorhinolaryngology, 50*, 154-159.

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

Lachs, L., Pisoni, D.B., & Kirk, K.I. (2001). Use of audio-visual information in speech perception by pre-lingually deaf children with cochlear implants: A first report. *Ear & Hearing, 22*, 236-251.

Landauer, T.K. (1962). Rate of implicit speech. *Perceptual & Motor Skills, 15*, 646.

Levitt, H., & Stromberg, H. (1983). *Speech of the Hearing-Impaired: Research, Training and Personnel Preparation*, (Baltimore: University Park Press).

Lezak, M.D. (1983). Neuropsychological assessment (2nd ed.). New York: Oxford University Press.

Li, S.C., & Lewandowsky, S. (1995). Forward and backward recall: Different retrieval processes. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*, 837-847.

Mathworks (2000). MATLAB (Version 6, Release 12) [Computer software]. www.mathworks.com.

McGarr, N.S. (1983). The intelligibility of deaf speech to experienced and inexperienced listeners. *Journal of Speech and Hearing Research, 26*, 451-458.

McGilly, K., & Siegler, R.S. (1989). How children choose among serial recall strategies. *Child Development, 60*, 172-182.

McLennan, S. (2001). Sonority contours in speech recognition: An examination of the Hoosier Mental Lexicon. *IULC Working Papers Online: 1*. https://www.indiana.edu/~iulcwp/pdfs/01-mclennan05.pdf.

Miyamoto, R.T., Osberger, M.J., Todd, S.L., Robbins, A.M., Stroer, B.S., Zimmerman-Phillips, S., et al. (1994). Variables affecting implant performance in children. *Laryngoscope, 104*, 1120-1124.

Osberger, M.J., Robbins, A.M., Berry, S.W., Todd, S.L., Hesketh, L.J., & Sedey, A. (1991). Analysis of the spontaneous speech samples of children with cochlear implants or tactile aids. *The American Journal of Otology, 12 (Suppl.),* 151-164.

Papagno, C., Valentine, T., & Baddeley, A. (1991). Phonological short-term memory and foreign-language vocabulary learning. *Journal of Memory and Language, 30*, 331-347.

Parkhurst, B., & Levitt, H. (1978). The effect of selected prosodic errors on the intelligibility of deaf speech. *Journal of Communication Disorders, 11*, 246-256.

Pisoni, D.B. (2000). Cognitive factors and cochlear implants: Some thoughts on perception, learning, and memory in speech perception. *Ear & Hearing, 21*, 70-78.

Pisoni, D.B. & Cleary, M. (in press). Some new findings on learning, memory and cognitive processes in deaf children following cochlear implantation. In F.G. Zeng, A.N. Popper & R.R. Fay (Eds*.), Springer Handbook of Auditory Research: Auditory Prosthesis, SHAR Volume X.*

Pisoni, D.B., & Geers, A.E. (2000). Working memory in deaf children with cochlear implants: Correlations between digit-span and measures of spoken language processing. *Annals of Otology, Rhinology, and Laryngology Supplement, 185*, 92-93.

Rosen, V. M., & Engle, R. W. (1997). Forward and backward serial recall. *Intelligence, 25*, 37-47.

Ross, M. & Lerman, J. (1979). A picture identification test for hearing-impaired children. *Journal of Speech and Hearing Research, 13*, 44-53.

Sankoff, D., & Kruskal, J.B. (1983). *Time warps, string edits, and macromolecules: the theory and practice of sequence comparison*. Reading, MA: Addison-Wesley.

Sarant, J.Z., Blamey, P.J., Dowell, R.C., Clark, G.M., & Gibson, W.P.R. (2001). Variation in speech perception scores among children with cochlear implants. *Ear & Hearing, 22*, 18-28.

Sehgal, S.T., Kirk, K.I., Svirsky, M.A., Ertmer, D.J., & Osberger, M.J. (1998). Imitative consonant feature production by children with multichannel sensory aids. *Ear & Hearing, 19*, 72-184.

Serry, T., Blamey, P., & Grogan, M. (1997). Phoneme acquisition in the first 4 years of implant use. *The American Journal of Otology, 18 (Suppl.),* 122-124.

Shipman, D.W., & Zue, V.W. (1982). Properties of large lexicons: Implications for advanced isolated word-recognition systems. In *ICASSP 82 Proceedings, May 3-5, 1982*.

Slobin, D.I. & Welsh, C.A. (1973). Elicited imitation as a research tool in developmental psycholinguistics. In: C.A. Ferguson & D.I. Slobin (Eds.), *Studies in Child Language Development*. New York: Holt, Rinehart & Winston.

Smith, C. (1975). Residual hearing and speech production in deaf children. *Journal of Speech and Hearing Research, 18*, 795-811.

Snik, A.F., Vermeulen, A.M., Geelen, C.P., Brokx, J.P., & van den Broek, P. (1997). Speech perception performance of children with a cochlear implant compared to that of children with conventional hearing aids. II. Results of prelingually deaf children. *Acta Otolaryngologica (Stockholm), 117*, 755-759.

Tobey, E.A., Angelette, S., Murchison, C., Nicosia, J., Sprague, S., Staller, S.J., Brimacombe, J.A., & Beiter, A.L. (1991). Speech production performance in children with multichannel cochlear implants. *The American Journal of Otology, 12 (Suppl.),* 165-173.

Tobey, E., Geers, A., & Brenner, C. (1994). Speech production results: Speech feature acquisition [Monograph]. *The Volta Review, 96*, 109-129.

Tobey, E.A., & Hasenstab, M.S. (1991). Effects of a Nucleus multichannel cochlear implant upon speech production in children. *Ear & Hearing, 12*, 48-54.

Tye-Murray, N., Spencer, L., Bedia, E.G., & Woodworth, G. (1996). Differences in children's sound productions when speaking with a cochlear implant turned on and turned off. *Journal of Speech and Hearing Research, 39*, 604-610.

Wechsler, D., 1991, *Wechsler Intelligence Scale for Children - III*. San Antonio, TX: The Psychological Corporation.

Wolfram, W. & Schilling-Estes, N. (1998). *American English*. Malden, MA: Blackwell.