AJSLP

Research Article

Nonword Repetition in Children With Cochlear Implants: A Potential Clinical Marker of Poor Language Acquisition

Susan Nittrouer,[a] Amanda Caldwell-Tarr,[a] Emily Sansom,[a] Jill Twersky,[a] and Joanna H. Lowenstein[a]

Purpose: Cochlear implants (CIs) can facilitate the acquisition of spoken language for deaf children, but challenges remain. Language skills dependent on phonological sensitivity are most at risk for these children, so having an effective way to diagnose problems at this level would be of value for school speech-language pathologists. The goal of this study was to assess whether a nonword repetition (NWR) task could serve that purpose.
Method: Participants were 104 second graders: 49 with normal hearing (NH) and 55 with CIs. In addition to NWR, children were tested on 10 measures involving phonological awareness and processing, serial recall of words, vocabulary, reading, and grammar.
Results: Children with CIs performed more poorly than children with NH on NWR, and sensitivity to phonological structure alone explained that performance for children in both groups. For children with CIs, 2 audiological factors positively influenced outcomes on NWR: being identified with hearing loss at a younger age and having experience with wearing a hearing aid on the unimplanted ear at the time of receiving a 1st CI. NWR scores were better able to rule out than to rule in such language deficits.
Conclusions: Well-designed NWR tasks could have clinical utility in assessments of language acquisition for school-age children with CIs.

Within the scope of practice for speech-language pathologists in the schools is the task of conducting assessments on students suspected of having or already diagnosed as having language-learning problems. These assessments are often intended to benchmark performance, such as when multifactorial evaluations are completed on children with Individualized Education Plans. However, these assessments should also help identify delays in or barriers to the acquisition of specific language skills, so intervention can be initiated on those skills. The best assessment tools are able to evaluate mechanisms underlying the acquisition of spoken language without being unduly influenced by factors external to the child, such as cultural background or socioeconomic status. Test procedures should be highly reliable. One group of children for whom tests sensitive to an assortment of language delays are especially needed is deaf children who get cochlear implants (CIs). Although the arrival of CIs has permitted deaf children to attain previously unprecedented levels of spoken language abilities, problems persist because these devices deliver only degraded spectral representations to their users. Consequently, children who must develop spoken language systems through these devices continue to face challenges in doing so. It is essential that the speech-language pathologists who serve these children in school settings have sensitive measures of their language abilities.

An assessment tool that has been touted as fitting the bill for children at risk of language delay is nonword repetition (NWR). In this task, a child hears a string of phonemes that do not form a real lexical item and must immediately repeat it. The series of phonemic strings are one or two syllables long at the start of the task and gradually increase in length to four or five syllables. When these tasks are used with children who have normal hearing (NH) but who have been diagnosed with language impairments, they are found to be highly sensitive to that diagnosis. For example, Dollaghan and Campbell (1998) used a NWR task with two groups of children matched on age and gender: one group of children diagnosed with specific language impairment (SLI; $n = 44$) and one group of children learning language in typical fashion ($n = 41$). All children were school age, spanning a range from 5;8 (years;

[a]The Ohio State University, Columbus
Correspondence to Susan Nittrouer: nittrouer.1@osu.edu
Editor: Krista Wilkinson
Associate Editor: Carol Miller
Received March 16, 2014
Revision received July 3, 2014
Accepted August 25, 2014
DOI: 10.1044/2014_AJSLP-14-0040

Disclosure Statement: The authors have declared that no competing interests existed at the time of publication.

months) to 12;2. Using a cutpoint of roughly 2 standard deviations below the mean of the language-typical group, these authors found that 61% of the children in the SLI group scored below the cutpoint. Using the same NWR task, other investigators (Ellis Weismer et al., 2000) replicated the basic result, showing that children diagnosed with SLI tended to score below the cutpoint established by Dollaghan and Campbell; however, the proportion was lower in this later study, at just 25% (20 of the 80 children diagnosed with SLI).

The difference in observed sensitivity of the NWR task to language impairment found for these two studies might arise from differences in the profiles of deficit across children in the two studies. In both studies, broad, albeit slightly different definitions of language impairment were applied. Neither study asked which specific language skills were predicted by scores on the NWR task. That question, however, has important clinical implications. Knowing what skills are predicted by the diagnostic task is critical to its application.

Language Skills Predicted by Nonword Repetition

Generally speaking, a wide assortment of language skills can be linked to NWR. Children diagnosed with language impairment on the basis of broad measures of ability (i.e., norm-referenced tests) have been found to perform poorly on NWR tasks (Dollaghan & Campbell, 1998; Ellis Weismer et al., 2000; Kamhi, Catts, Mauer, Apel, & Gentry, 1988). However, one study found that correlations between NWR and language measures were significant only when expressive (rather than receptive) language was being assessed (Edwards & Lahey, 1998), which prompted its authors to conclude that the major problem underlying poor NWR performance involves difficulty generating highly refined phonological representations. The basis of the argument hinged on the idea that more refined representations are required to produce speech than to comprehend language.

Because the psycholinguistic phenomenon most clearly assessed by NWR tasks concerns the nature of phonological representations, it is no surprise that these tasks are able to predict lexical knowledge (e.g., Avons, Wragg, Cupples, & Lovegrove, 1998; Gathercole, 1995; Michas & Henry, 1994); both NWR and lexical acquisition rely on having access to detailed phonological structure. Neither is it surprising that NWR can help predict reading ability (e.g., Kamhi et al., 1988), given the evidence that learning to read depends on having strong phonological sensitivity (e.g., Brady, Shankweiler, & Mann, 1983; Catts, 1989; Fox & Routh, 1980; Liberman & Shankweiler, 1985; Wagner & Torgesen, 1987).

Turning to grammatical abilities, a study by Sáhlen, Reuterskiöld-Wagner, Nettelbladt, and Radeborg (1999) found a significant correlation between NWR scores and scores on a test of children's ability to use various grammatical forms in their constructions. In summary, support has been found for the suggestion that appropriate NWR tasks should be able to predict performance on measures of lexical acquisition, reading ability, and use of grammatical structures. That makes NWR tasks potentially powerful tools for use in the clinical assessment of school-age children.

Independence From Unwanted Influences

Another consideration in selecting instruments for clinical assessments is that they should be as free as possible from inherent bias arising from differences in children's environments. This qualification can be a difficult one for assessment instruments to meet. Many norm-referenced tests used in the schools exhibit bias based on cultural or socioeconomic factors because of the content of individual items (e.g., Campbell, Dollaghan, Needleman, & Janosky, 1997). However, Dollaghan and Campbell (1998) demonstrated that a well-designed NWR task can avoid cultural and socioeconomic biases, and Ellis Weismer et al. (2000) subsequently reached that conclusion as well.

Because NWR involves having children repeat the sequences that are presented to them, another concern that could be raised is that scores on these tasks might be unduly influenced by oral motor skills. However, several studies investigating potential relationships between NWR and SLI have been careful to assess oral motor skills. In none of these instances has a correlation between oral motor skills and scores on the NWR task been observed (Edwards & Lahey, 1998; Gathercole & Baddeley, 1990; Sáhlen et al., 1999). At least some of the explanation for a failure to find a relationship between oral motor and NWR skills could be that many of the phonological sequences devised for use in NWR tasks are deliberately selected to consist of early acquired segments and no clusters. Presumably the oral motor actions needed to produce these sequences are within the abilities of most, if not all, children by the time they reach school age.

Finally, a characteristic of the NWR task that makes it an attractive assessment tool is that it provides a measure of children's phonological processing that is less influenced by the size of a child's existing vocabulary than other phonological short-term memory tasks, especially when low-probability phonological sequences are used (e.g., Archibald & Gathercole, 2006; Edwards, Beckman, & Munson, 2004; Gathercole, 1995). In typical serial recall tasks, including digit span, children are asked to recall strings of items that are likely components of their lexicons. Consequently, the strength of these representations within the lexicon can influence recall because detailed phonological representations may not be needed to code the items into a short-term memory buffer; instead, partial structures may activate word representations, which in turn are what get coded into a memory buffer. As a result, children with larger vocabularies will obtain more of a boost to their serial recall than children with smaller vocabularies. Support for that statement is provided by the finding that NWR is a more reliable marker of SLI than is serial recall of lexical items (e.g., Archibald & Gathercole, 2007; Conti-Ramsden, 2003).

The size of a child's vocabulary, however, may still have some effect on his or her performance on NWR tasks, through an indirect route. That effect arises because vocabulary size mediates sensitivity to phonotactic probabilities, such that children with larger vocabularies have presumably developed stronger sensitivities to these probabilities. As a result, if phonological sequences used in NWR tasks are ones that occur frequently within the native language, children with stronger vocabularies may be better able to code those sequences into storage (e.g., Archibald & Gathercole, 2006; Munson, Kurtz, & Windsor, 2005). However, the effect of lexical knowledge is mitigated when low-probability sequences are used.

Nonetheless, when nonwords consisting of low-probability phonological sequences are used in tests of NWR, the results can predict children's abilities to acquire new lexical items. The reason for that association is that all novel words begin as unfamiliar phonological strings. The better an individual's ability is to store these unfamiliar strings, the easier word learning will be for that individual. This advantage extends even to adults. For example, individuals who are successful at mastering multiple languages have been found to be especially good at NWR (Papagno & Vallar, 1995). Thus, by selecting a NWR task consisting of low-probability phonological sequences, it should be possible to assess children's word-learning potential.

Examining the Psycholinguistic Processes Underlying Nonword Repetition

NWR tasks primarily assess children's ability to recover detailed phonological representations from the speech they hear (i.e., phonological sensitivity) and encode those representations into a memory buffer. Gathercole (2006) referred to the output of this process as *phonological storage quality*. This phonological sensitivity, as assessed by NWR, has been reliably shown to be deficient in children with NH who are diagnosed with SLI (e.g., Conti-Ramsden, 2003; Edwards & Lahey, 1998; Ellis Weismer et al., 2000; Gathercole & Baddeley, 1990; Sáhlen et al., 1999). However, an additional deficit—one involving short-term memory—seems to be at work in these tasks for children with SLI, a conclusion reached by the finding that the magnitude of the effect changes across syllable length. The Group (language normal vs. SLI) × Syllable Length interaction is typically found to be significant in experiments using NWR, such that performance of children with SLI degrades across length more rapidly than that of children with NH (e.g., Briscoe, Bishop, & Norbury, 2001; Dollaghan & Campbell, 1998; Gathercole & Baddeley, 1990; Marton & Schwartz, 2003).

Nonword Repetition and Children With Cochlear Implants

The interest in using NWR tasks with children who are deaf and get CIs arises because CIs provide only spectrally degraded signals to their users. Phonological representations are recovered largely from the spectral structure of the acoustic signals reaching listeners. Thus, one could predict that children who must rely on CIs for language learning would face serious challenges as they develop phonological representations, and several studies have found evidence to support that prediction (e.g., Ambrose, Fey, & Eisenberg, 2012; James et al., 2005; Johnson & Goswami, 2010; Nittrouer, Caldwell, Lowenstein, Tarr, & Holloman, 2012). As a result, one could further predict that children with CIs would perform more poorly on NWR tasks than children with NH, and that outcome has also been supported by empirical evidence (Carter, Dillon, & Pisoni, 2002; Dillon & Pisoni, 2006). Nonetheless, additional questions are sparked by these findings from children with CIs on NWR tasks. One question concerns the extent to which these measures can shed light on the basic psycholinguistic processes perturbed by hearing loss and subsequent implantation. In particular, is the NWR performance of children with CIs explained primarily by the quality of their phonological representations? Or does short-term memory play a role as well, as is the case for children with SLI?

A study designed to help make this determination was done by Briscoe et al. (2001), who compared performance on a number of language tasks for children with mild to moderate hearing loss and children with NH but with SLI. Although Briscoe et al. found poorer NWR for children with mild to moderate hearing loss than for children with NH and no SLI, the magnitude of the effect was consistent across stimulus length: The Group × Syllable Length interaction was not significant. That outcome contrasted with their own results for the group of children with SLI in that study.

Another question concerning the performance of children with CIs on NWR tasks involves the language skills that might be predicted by these tasks for these children. Dillon and Pisoni (2006) tried to answer that question to some extent by looking at reading skills. These investigators observed poorer performance on NWR for children with CIs than for NH peers, and scores were correlated with reading abilities for the children with CIs. This finding suggests that NWR could be an effective means of assessing language skills in children with CIs.

Current Study

The current study had four specific objectives. The first objective was to compare NWR performance of children who were born with severe to profound sensorineural hearing loss and got CIs with that of peers with NH. For this comparison, we selected the nonwords developed by Dollaghan and Campbell (1998). Two sets of nonwords have been used more than any others in NWR tasks. One set is known as the Children's Test of Nonword Repetition, and it was developed by Gathercole, Willis, Baddeley, and Emslie (1994). These stimuli were the ones used by Carter et al. (2002) and by Dillon and Pisoni (2006) in their studies involving children with CIs. These stimuli are two to five syllables long and contain some consonant clusters and lax vowels. Carter et al. found that children with CIs performed quite poorly with these stimuli: Children with CIs correctly repeated a mean of only 5% of all nonwords.

The nonword stimuli developed by Dollaghan and Campbell (1998) consist of one to four syllables (four stimuli at each syllable length). No clusters are included. No late-acquired consonants, as defined by Shriberg and Kwiatkowski (1994), are incorporated into the stimuli. Only tense vowels are used because these vowels are more salient than lax vowels and less susceptible to being reduced to schwa on repetition, making scoring more straightforward. An example of a one-syllable stimulus is [tɑʊʤ] and an example of a four-syllable stimulus is [dævoʊnɔɪʧig]. On the basis of these characteristics, we anticipated that children with CIs would attain higher scores with these stimuli than they had with the Children's Test of Nonword Repetition.

The second objective of this study was to examine the psycholinguistic underpinnings of NWR for children with CIs. Again, for children with SLI, short-term memory deficits appear to influence NWR performance, as well as the poor quality of phonological representations that they have. Here, we asked whether a similar effect would be found for children with CIs. In earlier work, children with CIs have been found to demonstrate poor digit span, which suggests poor short-term recall (e.g., Pisoni & Cleary, 2003; Pisoni, Kronenberger, Roman, & Geers, 2011). However, outcomes of one study were able to pinpoint the basis of that poor recall as being impoverished phonological representations (Nittrouer, Caldwell-Tarr, & Lowenstein, 2013). In particular, results showed that even though serial recall for lists of words was poorer for children with CIs than for their peers with NH, recall took no longer. Because response time is generally considered to be an index of processing effort (e.g., Cooper-Martin, 1994), one can suggest that the children with CIs experienced no more difficulty with the processing part of this task than the children with NH. These results can be considered in light of multiple-component models of working memory (e.g., Baddeley, 1992, 2007), which propose that one component—the phonological loop—is responsible for recovering phonological structure from speech signals and encoding that structure into the short-term memory store, and another component—the central executive—is responsible for performing operations such as recall with that stored structure. Accordingly, the finding of less accurate serial recall with no increase in response time observed for children with CIs (compared with children with NH) suggests that children with CIs have difficulty encoding linguistic materials into a memory buffer but are able to operate on items placed in that buffer without increased effort. Thus, the problems encountered by children with CIs in their recall of words could largely be explained by difficulty encoding clear phonological representations, precisely the phenomenon thought to underlie NWR. This account might explain the NWR outcomes of Briscoe et al. (2001), who observed a Group × Syllable Length interaction for typically developing children and those with SLI but not for typically developing children and those with mild to moderate hearing loss. In addition to impoverished phonological representations, the children with SLI may have had deficits in the operations of the central executive, whereas the children with hearing loss apparently did not.

The third objective of this study was to examine the extent to which NWR for children with CIs is influenced by factors other than psycholinguistic skills. In particular, the question was whether NWR is affected by demographic factors such as socioeconomic status or audiological factors such as age at receiving a first CI. If NWR is heavily influenced by these external factors for these children, then NWR scores themselves offer little additional power for predicting language skills over what is available from knowing the child's socioeconomic status and age at receiving a first implant. In this case, it would not serve as a valuable clinical tool.

The fourth objective of this study was to examine the set of language abilities that might be predicted by NWR for children with CIs. The more language skills that are well predicted by this task, the more useful NWR could be as an assessment tool with this population of children. In particular, we examined skills in three categories as potentially predictable from NWR scores: vocabulary knowledge, reading abilities, and grammar.

Method

Participants

This study had 104 participants in two groups: 49 with NH and 55 with CIs. These sample sizes provide 71% power and 99% power for detecting differences between groups with Cohen's *d*s of 0.5 and 1.0, respectively, with alpha set to .05. All these children had participated in an ongoing longitudinal study (e.g., Caldwell & Nittrouer, 2013; Nittrouer, 2010; Nittrouer et al., 2012, 2013), and on the basis of outcomes from earlier analyses done as part of that study, we anticipated effect sizes of that magnitude or larger. Consequently, adequate power was available.

All participants had just completed second grade when the data described in this article were collected. Mean ages at the time of data collection were 8;5 (*SD* = 4 months) and 8;7 (*SD* = 6 months) for children with NH and those with CIs, respectively. The NH group had 27 girls and 22 boys; the CI group had 28 girls and 27 boys.

Socioeconomic status (SES) was indexed using a two-factor scale on which both the highest educational level and the occupational status of the primary income earner in the home is considered (Nittrouer & Burton, 2005). Scores for each of these factors range from 1 to 8, with 8 being high. Values for the two factors are multiplied together, resulting in a range of possible scores from 1 to 64. In general, a SES score of 30 represents a household in which the primary income earner has a 4-year university degree and a job such as a mid-level manager or a teacher. Means for SES were 35 (*SD* = 13) for children with NH and 33 (*SD* = 11) for children with CIs.

We assessed nonverbal cognitive abilities using the Leiter International Performance Scale—Revised (Roid & Miller, 2002). Standard scores on this test were 105 (*SD* = 14) for children with NH and 101 (*SD* = 17) for children with CIs. These differences between groups for SES and nonverbal abilities were not significant.

All children with CIs had their hearing loss identified, hearing aids fit, and intervention initiated by age 2 years; mean age at identification was 7 months (*SD* = 7 months), mean age at receiving hearing aids was 8 months (*SD* = 6 months), and mean age at starting intervention was 9 months (*SD* = 7 months). The mean better-ear, pure-tone average threshold for the three speech frequencies of 0.5, 1.0, and 2.0 kHz before receiving a CI (hereinafter termed *preimplant PTA*) was 100 dB hearing level (*SD* = 17 dB hearing level). Although all of these children were fit with hearing aids before age 2 years, most of them (80%) also received their first CI before turning age 2. Thirty-six children had bilateral CIs. Of the 19 children with one CI at the time of testing, six wore a hearing aid on the unimplanted ear.

A final audiological factor considered here was whether these children had some period of experience wearing a hearing aid on the ear opposite to the implanted ear. Previous work has shown that children with this experience, known as *bimodal stimulation*, display better language skills, at least during the preschool years (Nittrouer & Chapman, 2009). The question addressed here was whether that benefit would extend to second grade. A basis for predicting that it might concerns the development of phonological representations, precisely the phenomenon underlying NWR. Even if acoustic hearing is available in only a limited frequency range, that signal is likely more detailed than what is available through the electric signal of the CI. To the extent that phonological categories are acquired during the early years, having some time with those more detailed signals could facilitate the acquisition of more refined representations. In this study, 26 of the children with CIs had at least 1 year of experience wearing a hearing aid on the ear opposite to their CI at the time that they received that CI. Six were still wearing a hearing aid on the unimplanted ear. The mean time that the other 20 children spent wearing a hearing aid on the unimplanted ear after receiving a first implant was 24 months (*SD* = 10 months).

Equipment

All testing took place in sound-attenuated rooms. All stimuli used in testing were presented via a computer with a Creative Labs Soundblaster soundcard using a 44.1-kHz sampling rate with 16-bit digitization and a Roland MA-12C powered speaker for audio presentation. No live-voice stimuli were used. For the NWR and phonological processing tasks, stimuli were presented in audiovisual format using a 1,500-kbps data rate and 24-bit digitization for video presentation. This allowed children to use visual cues for speech recognition. For the serial recall task, presentation was audio only. Presentation level was always 68 dB SPL, regardless of whether it was audiovisual or audio only. A touchscreen monitor was used for the serial recall task.

All test sessions were video- and audio-recorded using a SONY HDR-XR550V video recorder so scoring could be done later. Children wore Sony FM transmitters in specially designed vests. The FM receivers provided direct-line input to the video cameras to ensure good sound quality for all recordings.

General Procedures

All children came to the Ohio State University for testing. Children were tested individually in sessions lasting no more than 1 hr. Breaks of at least 1 hr were provided between those data collection sessions. For this study, data were collected over three sessions. In one session, three phonological processing tasks were administered. In another session, a 20-min language sample was collected to obtain the grammatical measures. In that same session, the NWR and serial recall tasks were administered. In a third session, the reading and vocabulary measures were collected.

Task-Specific Procedures

NWR. We used the 16 nonwords developed by Dollaghan and Campbell (1998). To ensure that they were produced exactly as they had been in that earlier work, an audio recording from Dollaghan and Campbell was obtained and reviewed. Stimuli were video- and audio-recorded by the last author, who is both a trained phonetician and a trained singer, which ensured that the stimuli would be recorded as described. Equal stress was placed on all syllables for all stimuli. Fundamental frequency was kept consistent and flat. Amplitude of stimuli was constant. The last author recorded the same instructions as had been used by Dollaghan and Campbell (1998), and these instructions were placed at the start of the video recording of stimuli.

For this task, the children saw and heard the talker saying each nonword, and they needed to repeat each one immediately on hearing it. Children were video recorded repeating the nonword stimuli. Scoring was completed later by the third author, who was involved more broadly in scoring tests requiring verbal responses from these children. That permitted her to become familiar with the speech patterns of individual children. For this NWR task, phonemes were scored as wrong if they were omitted or were clear substitutions for modeled phonemes. However, distortions were not scored as wrong. Of the nonword video recordings, 20% (20 children) were scored independently by the last author to obtain a measure of reliability. Ten of these children had NH, and 10 had CIs.

Phonological processing. We used three tasks of varying difficulty to assess children's sensitivity to phonological, especially phonemic, structure. Work by Stanovich, Cunningham, and Cramer (1984) primarily served as the basis for predictions of difficulty level for these tasks, although Yopp (1988) also demonstrated the developmental time course of acquisition for specific phonological abilities.

The specific tasks used in this study have been used in other studies and have been found to be reliable within the same children and across children of the same age (e.g., Nittrouer, 1999; Nittrouer & Burton, 2002; Nittrouer & Miller, 1999; Nittrouer, Shune, & Lowenstein, 2011). The first two tasks are commonly described as investigating phonological awareness rather than processing (e.g., Snider, 1997; Stanovich et al., 1984) because there are lighter processing demands in these tasks than in some tasks, such as the third one used in this study. Nonetheless, we use the broader term *processing* in this article to encompass all three of these tasks, for simplicity's sake. Using this set of tasks varying in difficulty diminished the probability that significant differences in abilities across groups would fail to be identified, either because a selected task was so easy that even children with phonological delays were able to perform it or because it was so difficult that even typically developing children could not. All three tasks were presented with specially developed software. Video recordings of a male talker served as stimuli for all three tasks. The experimenter entered the child's responses, and the software kept track of those responses.

The first task, the initial consonant choice task, was considered the easiest going into testing because Stanovich et al. (1984) showed that children develop sensitivity to phonological structure at the start of words before they develop sensitivity to structure at the end of words. This initial consonant choice task consisted of 48 items and began with the child getting a target word to repeat. The child was given three opportunities to repeat this target word correctly. If the target had not been repeated correctly within three attempts, testing would have advanced to the next trial and the missed trial would not have been included in the overall calculations of percentage correct. However, all these children were able to repeat all the target words correctly with the audiovisual presentation.

After correct repetition of the target word, the child was presented with three more words and had to choose the one that had the same beginning sound as the target word. Test items can be found in Appendix A. The dependent measure was the percentage of trials on which the child correctly chose the word with the same initial sound as the target. Although this task, as well as the second, involved some memory load, the ability to retain four words in a short-term memory buffer should have been well within the abilities of all these second graders: Test norms for the Wechsler Intelligence Scales for Children (Wechsler, 1991) show that 99.5% of 8-year-olds with NH can recall four-digit strings, and Pisoni et al. (2011) showed that children with (early) CIs could recall four-digit strings by age 8 years.

The second task, the final consonant choice task, was considered to be intermediate in difficulty for children of this age: Although procedures were identical to those of the initial consonant choice task, Stanovich et al. (1984) showed that sensitivity to syllable-final structure is acquired later. This task consisted of 48 items, shown in Appendix B. The dependent measure was the percentage of trials on which the child correctly chose the word with the same ending sound as the target.

The third task, the phoneme deletion task, was the most difficult because it involved the most processing: In this task, the child saw and heard the talker say a nonword in the context "Say _______." The child needed to repeat that nonword correctly. Then the talker would say, "Now say ________ without the ____ sound." Thus, in this task children not only needed to recognize phonemic structure in a nonword, they also needed to manipulate that nonword structure so that one segment was removed, and the remaining segments were blended. The segment to be removed could occur anywhere within the word. The task consisted of 32 items, which are found in Appendix C. Percentage of correct responses served as the dependent measure.

Serial recall of word lists. Children were asked to recall the order of strings of six monosyllabic nouns produced by a male talker, presented as auditory lists: *ball, coat, dog, ham, pack,* and *rake*. For children, lists six words in length result in scores that are not close to either ceiling or floor. Ten of these word lists were presented, with words ordered randomly by the software on each presentation. The words were presented at a rate of one per second. Simple pictures of the words were shown at the top of the touchscreen monitor. After presentation of each list, the child's task was to touch the pictures in the order in which the words were heard. Responses were recorded by the program.

Before testing, training on this task was provided using six nonrhyming and visually distinct letters: *F, H, Q, R, S,* and *Y*. After practice with the letters and before testing, the six words were introduced, one at a time, to familiarize the child with each word and its corresponding picture. Then recognition was checked for each child before testing by showing pictures of all six words and playing each word, one at a time. The child needed to touch each word immediately after it was played. After testing, this recognition check was done again. If a child had difficulty recognizing any word, testing would not have been conducted (if it happened during the pretest) or data would have been removed from analysis (if it happened during the posttest). However, all children readily recognized these simple nouns when presented in this auditory-only format. The maximum possible correct was 60 (six items in 10 lists), and the dependent measure was the percentage of words recalled in the correct order.

This test procedure has often been used to examine short-term memory (e.g., Brady et al., 1983; Spring & Perry, 1983), and this particular task with these particular words has been shown to have good test–retest reliability (Nittrouer & Miller, 1999).

Vocabulary knowledge. For this study, we assessed expressive vocabulary with the Expressive One-Word Picture Vocabulary Test (EOWPVT; Brownell, 2000). This task requires the child to provide the words that label a series of pictured items shown one at a time on separate pages. For this task, children's verbal responses were video recorded. A laboratory staff member scored responses later, and a second staff member checked all scores by watching the video recording again and confirming those scores. If any discrepancies were found, the staff members resolved them by consensus. The laboratory manager monitored all scoring procedures. Standard scores were used as dependent measures.

Expressive vocabulary was measured in this study, rather than receptive vocabulary, because a word must be retained in the lexicon more robustly for a child to be able to access it during an expressive vocabulary test. That difference is because in expressive vocabulary tests, only pictures are shown. The child must be able to retrieve matching words from his or her lexicon on his or her own. In contrast, when receptive vocabulary is measured, the child hears the lexical item and need only select from a set of four the picture that best matches the word heard. That task can be accomplished with less robust lexical representations.

Word reading of isolated words. To assess children's abilities to read words presented in isolation, we used the Word Reading subtest of the Wide Range Achievement Test 4 (WRAT4; Wilkinson & Robertson, 2006). In this test, the child reads a list of words presented on a card. Reading responses were video recorded and scored later by a staff member. Again, a second staff member confirmed all scores by watching the video recording and checking the scores of the first staff member. Standard scores were used as dependent measures.

Reading of words in context and reading comprehension. We used the Qualitative Reading Inventory (QRI; Leslie & Caldwell, 2006) to assess reading of words in context and reading comprehension. The QRI has both narrative and expository passages written at various levels of reading ability. In this study, we used three passages. One passage was a narrative written at one level below grade level, one was a narrative at grade level, and one was an expository at grade level. Children were video recorded reading each story and responding to questions designed to assess their comprehension. There were 10 comprehension questions per story. A staff member scored the number of words read correctly and the number of questions answered correctly from the video recording. A second staff member checked all scores by watching the video recordings and checking the scores of the first staff member. We used the mean number of words read correctly across the three passages as the dependent measure of word reading ability, and the total number of questions answered correctly across the three passages served as the measure of reading comprehension.

Grammatical measures. We obtained two measures of grammatical abilities from a 20-min sample of narrative language for each child. To collect this narrative sample, each child entered the sound booth, and the experimenter explained that she had been called away for a few minutes. The equipment was set up for the child to view and hear a video of the book *The Day Jimmy's Boa Ate the Wash* (Noble, 1980). This story was video recorded with a narrator reading the printed material, but with separate staff members saying the material that appeared in quotes in the book. Full images of each face were shown to provide maximum opportunity for speechreading. Illustrations from the book were shown when appropriate. The experimenter explained that she had not seen the video story yet and asked the child to watch it carefully so the child could tell it to the experimenter when she returned. After the story was finished, the experimenter reentered the sound booth and asked the child to tell her the story in as much detail as possible. If the story retelling did not take a full 20 min, the experimenter supplemented the time by asking questions about personal experiences the child had that paralleled some of those of the children in the story.

The story retelling was transcribed later by members of the laboratory staff, starting at the 5-min mark, according to methods first described by Hewitt, Hammer, Yont, and Tomblin (2005). Similar methods were subsequently used by Moeller et al. (2010) and Nittrouer et al. (2012). Two staff members and the laboratory manager trained together on transcription methods for Systematic Analysis of Language Transcripts (Miller & Iglesias, 2010) and together practiced transcribing several story retellings of children not in this study. One of the two staff members watched each video and transcribed every utterance the child produced (intelligible and unintelligible) in the 15-min segment. After completing the transcript, that staff member went back and checked it by watching the video again while reading the transcript. Then the second staff member checked the same transcript for accuracy by reading through it while watching the video. Finally, the two of them resolved any discrepancies in how specific utterances should be transcribed by watching the video together and discussing it.

The entire transcription process was monitored by the laboratory manager, who spot-checked progress and served as an arbitrator if the staff members were unable to reach consensus regarding how a specific utterance should be transcribed. In addition, the laboratory manager transcribed 5% of the samples herself (three from children with NH and three from children with CIs). No discrepancies were noted between the transcripts of the laboratory manager and those of the other staff members. The transcripts for the 15-min samples were submitted to Systematic Analysis of Language Transcripts, but we used the results from just the first 100 complete and intelligible utterances in this study.

Two measures were used. First, mean length of utterances in morphemes (MLU) was included. Although this measure has been criticized for possibly being insensitive to differences in syntactic abilities among children after kindergarten (e.g., Crain & Lillo-Martin, 1999), it has been found to distinguish between children with language deficits and children with normal language in the elementary grades (e.g., Condouris, Meyer, & Tager-Flusberg, 2003). Consequently, we considered it appropriate to use with these second-grade children with CIs, who were predicted to have delayed syntactic development.

The second Systematic Analysis of Language Transcripts measure used in this study was the number of bound morphemes produced during the 100-utterance set. These morphemes included verb-related *–ed, –s,* and *–ing*; noun-related plural *–s* and possessive *–s*; and adjective-related *–er* and *–est.* The number of bound morphemes in each transcript was highly correlated with MLU (for children with NH, $r = .547$; for children with CIs, $r = .660$; $p < .001$ in both cases), but we nonetheless considered it valuable to include this additional measure. More than most morphemes, the acquisition of bound morphemes depends on the child being able to recognize word-internal phonological structure, and that sensitivity (to word internal structure) is exactly what is required for NWR. Consequently, this grammatical skill might be especially well predicted by NWR.

Results

Reliability of Nonword Repetition Scores

We obtained external reliability for these scores by comparing outcomes of this study with those of Dollaghan and Campbell (1998). To do that, we compared the total percentage of phonemes correct (TOTPPC) scores across all 16 stimuli for the two studies. In total, 96 phonemes were included in the 16 nonwords. In their study, Dollaghan and Campbell found mean TOTPPC scores of 84 ($SD = 7$) for children with typical language development and 66 ($SD = 12$) for children with SLI. In this study, mean TOTPPC scores were 83 ($SD = 7$) for children with NH and 67 ($SD = 12$) for children with CIs. We considered those similarities to represent very good external reliability.

We obtained an estimate of internal reliability for these scores from the proportion of phoneme-by-phoneme agreement measured for the 20 children whose responses were scored by a second staff member (the last author). This is the same procedure as that used by Dollaghan and Campbell (1998) to estimate reliability. In that earlier study, mean percentage agreement in scores for two scorers across children (20% of all children) was 95%, and agreements for individual samples ranged from 91% to 99%. In this study, we obtained a mean agreement score of 95%, with agreement scores for individual samples ranging from 90% to 100%. These outcomes were considered to represent very good reliability, and the scores from the staff member who scored all 104 samples were used in further analysis.

Data Screening

We screened all measures to be used in analyses for homogeneity of variance and normal distribution. All measures met criteria to be appropriate for use in inferential and regression analyses, so no transformations were performed. In reporting statistical outcomes, precise results are given when $p < .10$; when $p > .10$, outcomes are reported simply as not significant.

Group Differences

Table 1 shows means and standard deviations for each dependent measure included in the study separately for children with NH and children with CIs. In addition, outcomes of independent-samples *t* tests performed on scores for each dependent measure are shown. All differences in group means were significant, with $p < .01$. Finally, Cohen's *d*s for each measure are shown. As can be seen from these values, the difference in scores for NWR between children with NH and children with CIs was the largest of any measure.

Bases of NWR

Figure 1 shows the mean percentage of correct phoneme repetition for each syllable for each group. From this figure, it appears that the pattern of decrement in correct repetition was similar across the two groups, with a slight disparity for the three-syllable stimuli: For stimuli of that length, children with NH appear to have maintained performance slightly better than children with CIs. To test for differences between groups in the pattern of decrement, we performed a two-way, repeated measures analysis of variance on these data with syllable length as the repeated measure and group as the between-subjects measure. Both main effects were found to be significant: syllable length, $F(3, 306) = 228.05$, $p < .001$, $\eta^2 = .691$, and group, $F(1, 102) = 69.40$, $p < .01$, $\eta^2 = .405$. However, the Syllable Length × Group interaction was not significant, $F(3, 306) = 2.54$, $p = .057$. The finding that the outcome was close to significant likely reflects the small discrepancy in pattern of decrement observed for stimuli three syllables in length.

Next, we performed forward stepwise regression on the TOTPPC scores, using the three phonological processing measures and serial recall scores as predictor variables in a one-step analysis. This procedure was undertaken specifically to examine whether the degree of phonological sensitivity, as measured by the phonological processing tasks, primarily accounted for NWR abilities, or whether an additional short-term memory factor—such as operations of the central executive—might explain some additional variability for one or the other group.

We performed a separate regression analysis on scores from each group. For children with NH, scores on the phoneme deletion task accounted for a significant amount of variance in TOTPPC scores, with standardized $\beta = .519$, $p < .001$, $R^2 = .27$. No other variable accounted for any significant amount of additional variance; in particular, scores for the serial recall task did not account for any additional variance. For children with CIs, the initial consonant choice task accounted for a significant amount of variance in the TOTPPC scores, with standardized $\beta = .483$, $p < .001$, $R^2 = .23$. No other variable accounted for any significant amount of additional variance; in particular, scores for the serial recall task did not explain any additional variance.

These findings suggest that phonological sensitivity explained the most variance in NWR scores for both groups. It is of interest that scores from the phonological processing task viewed as being developmentally the easiest explained these scores for children with CIs, who performed poorest on NWR, and scores from the developmentally most difficult phonological processing task explained NWR scores for children with NH. This trend highlights the relationship between the developmental processes of progressive refinement in phonological representations and abilities to retain a string of phonological units.

Table 1. Mean scores (and standard deviations), *t* tests, and Cohen's *d*s for the dependent measures.

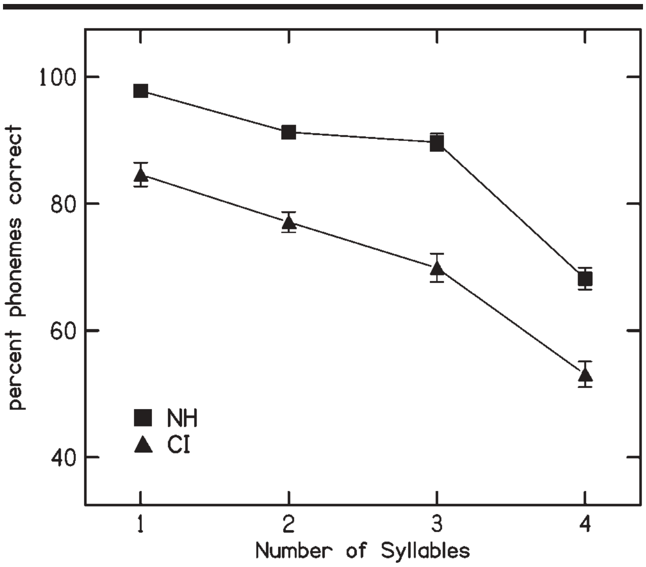
Dependent measures	Children with NH, *M* (*SD*)	Children with CIs, *M* (*SD*)	*t*(102)	*p*	Cohen's *d*
Nonword repetition: TOTPPC	83.0 (7.1)	67.5 (12.3)	7.72	< .001	1.60
Phonological processing					
Initial consonant choice	87.4 (13.2)	63.1 (25.9)	5.92	< .001	1.24
Final consonant choice	69.8 (17.9)	35.8 (25.6)	7.73	< .001	1.56
Phoneme deletion	71.5 (21.5)	47.5 (32.6)	4.35	< .001	0.89
Serial recall: percentage correct	56.1 (16.5)	43.3 (15.4)	4.06	< .001	0.80
Vocabulary knowledge: EOWPVT Standard Score	110.0 (13.7)	94.4 (18.1)	4.92	< .001	0.98
Reading isolated words: WRAT Standard score	110.0 (11.7)	101.0 (14.6)	3.44	.001	0.68
Reading in context					
QRI words correct	200.3 (5.3)	190.5 (14.7)	4.39	< .001	0.98
QRI comprehension	20.8 (3.0)	16.6 (6.0)	4.41	< .001	0.93
Grammar					
Mean length of utterance	6.3 (1.5)	5.5 (1.4)	2.67	.009	0.55
Bound morphemes	109.8 (33.5)	82.5 (31.4)	4.28	< .001	0.84

Note. NH = normal hearing; CI = cochlear implant; TOTPPC = total percentage of phonemes correct; EOWPVT = Expressive One-Word Picture Vocabulary Test; WRAT = Wide Range Achievement Test; QRI = Qualitative Reading Inventory.

Independence From External Influences on Scores for Children With CIs

An important consideration in assessing whether NWR would be a sensitive metric of language performance for these children with CIs is how strongly influenced scores are by factors external to the children. To make that determination, we computed Pearson product–moment correlation coefficients between TOTPPC and each demographic and audiological factor of interest: SES, age at identification of hearing loss, age at first implant, and preimplant PTA. We examined gender as a possible predictor of scores, using *t* tests. Before discussing outcomes for children with CIs, however, it is noteworthy that the correlation coefficient between TOTPPC and SES was not significant for children with NH, and we found no significant difference in scores between girls and boys. Consequently, one can conclude that these two demographic factors, found to explain some proportion of variance on other language measures, do not influence NWR for children with NH.

Figure 1. Mean percentage of correct phonemes repeated at each syllable length by children with normal hearing (NH) and children with cochlear implants (CIs). Error bars are standard errors of the mean.



Returning to children with CIs, of the four demographic and audiological factors examined with correlational analyses, only age at identification of hearing loss was found to have a significant relationship with TOTPPC, $r(55) = -.304$, $p = .024$. The negative coefficient indicates that the younger children were when they were identified with hearing loss, the higher they scored on NWR. The magnitude of the effect indicates that age at identification accounted for roughly 9% of variance in TOTPPC at second grade. We found no significant difference in scores between these girls and boys.

Differences between groups of children with CIs were also examined, based on whether they had one or two CIs and whether they had a period of bimodal stimulation near the time of receiving their first CIs. For these analyses, we did not include the six children who continued to wear a hearing aid on the unimplanted ear at the time of testing because they did not fit cleanly into the groups of interest.

First, we evaluated differences between children with one versus two CIs. A comparison of demographic and audiological factors for these two groups of children with CIs revealed no differences between groups on those factors. Mean scores for TOTPPC were 65.9 (*SD* = 10.6) and 69.7 (*SD* = 12.2) for children with one and two CIs, respectively. Although age at identification of hearing loss did not differ for children with one versus two CIs, we incorporated it into the analysis as a covariate. Nonetheless, the analysis of covariance (ANCOVA) revealed that the effect of having one or two CIs was not significant. Consequently, one can conclude that having two CIs did not provide an advantage over having just one CI on this NWR task. For the 36 children who had two CIs at the time of testing, we observed no significant effect of the age at receiving a second CI.

Next, we evaluated potential differences between children with CIs on the basis of whether they had a period of bimodal experience around the time of receiving a first CI. On NWR performance, we found that the 20 children with some bimodal experience obtained a mean score of 75.4 ($SD = 7.7$) for TOTPPC and the 29 children with no bimodal experience received a mean score of 64.0 ($SD = 12.0$). Although age at identification did not differ for children with some bimodal and no bimodal experience, we again used it as a covariate. In this case, the ANCOVA revealed that the effect of having bimodal experience was significant, $F(1, 46) = 16.75$, $p < .001$, $d = 1.13$. When we computed a Pearson product–moment correlation coefficient on NWR scores and the length of time that children had bimodal experience, the resulting coefficient was .377; however, with only 20 participants in the analysis it was not significant. Nonetheless, on the basis of the ANCOVA result, one may conclude that having a period of bimodal experience around the time of receiving a first implant facilitates NWR.

In spite of finding a significant difference between bimodal groups, we examined audiological factors to see whether there might actually be any group differences that could account for this effect. In this case, we observed three differences in audiological factors between groups. First, children with some bimodal experience generally received their first CIs later than children with no bimodal experience: 22 months ($SD = 15$ months) for children with some bimodal experience versus 16 months ($SD = 6$ months) for children with no bimodal experience. This difference was significant, $t(47) = 2.15$, $p = .037$. Second, children with some bimodal experience who had two implants ($N = 18$) received their second CIs later than children with no bimodal experience who received a second implant ($N = 21$). In this case, the mean age at receiving a second CI was 52 months ($SD = 24$ months) for children with some bimodal experience and 35 months ($SD = 14$ months) for children with no bimodal experience. This difference was significant, $t(37) = 2.66$, $p = .012$. Finally, preimplant PTAs differed for these two groups of children. For children with some bimodal experience, mean preimplant PTA was 97 dB hearing level ($SD = 16$ dB hearing level). For children with no bimodal experience, mean preimplant PTA was 107 dB hearing level ($SD = 11$ dB hearing level). This difference was significant, $t(47) = 2.63$, $p = .011$.

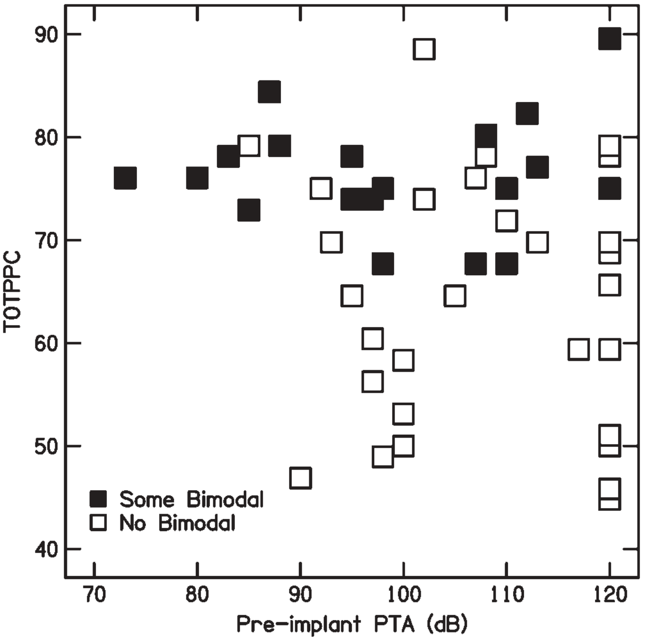
In considering whether these group differences in audiological factors could have evoked the advantage in NWR scores observed for children with some bimodal experience, the following points are pertinent: The only audiological factor found to have a significant effect on performance across all children with CIs was age at identification, and that factor did not differ for these groups. Regarding age at receiving a CI, children with some bimodal experience received both first and second implants later, on average, than children with no bimodal experience. Although not found to correlate significantly with TOTPPC in this study, later age at receiving a CI is generally considered to have a negative impact on language development. Consequently, this difference between groups would only be expected to benefit children with no bimodal experience. Preimplant PTAs did differ significantly between these groups, and in such a way as might be expected to benefit children with some bimodal experience. Thus, this potential effect was investigated more thoroughly, even though preimplant PTAs had not been found to correlate with TOTPPC across all children with CIs. Figure 2 shows TOTPPC scores, as a function of preimplant PTA, with symbols marking group identity. This figure clearly illustrates that TOTPPC was not related to preimplant PTA for either group. Consequently, the advantage observed in NWR performance for children with some bimodal experience over those with no bimodal experience could not be attributed to these differences in audiological factors.

In summary, these analyses indicate that demographic and audiological factors explained little variance in outcomes on the NWR task for children with CIs. The only exceptions to this trend were that age at identification of hearing loss accounted for roughly 9% of the variance in scores, and having had a period of time with bimodal experience provided positive effects.

Language Skills Predicted by Nonword Repetition

A final consideration in assessing whether NWR would be a good clinical tool for use with children with CIs is how well it predicts other language skills for this group. We took two approaches to this question. First, we computed Pearson product–moment correlation coefficients between TOTPPC and scores on the language measures that had been hypothesized to be well predicted by

Figure 2. Total percentage of phonemes correct (TOTPPC) for children with cochlear implants (CIs) as a function of preimplant pure-tone average threshold (PTAs).



NWR: that is, vocabulary knowledge, reading, and grammar. Specifically, we computed correlation coefficients between TOTPPC and each of the measures listed in Table 1 for vocabulary knowledge, reading isolated words, reading in context, and grammar. Again, considering outcomes first for children with NH, we observed that three of the six measures had significant correlations with TOTPPC: EOWPVT, $r(49) = .291$, $p = .042$; WRAT4, $r(49) = .464$, $p = .001$; and QRI words correct, $r(49) = .298$, $p = .037$. These findings indicate that vocabulary acquisition and word reading—whether for words in isolation or words in context—were all found to depend to some extent on children's abilities to recover and store strings of well-defined phonological segments. However, only the correlation coefficient for TOTPPC and WRAT4 scores reached a description of being moderate in magnitude (i.e., had more than 10% of their variance in common). Reading comprehension and grammatical skills were not correlated with NWR for these children with NH at all.

Returning to the children with CIs, we found all six measures to have significant correlation coefficients: EOWPVT, $r(55) = .582$, $p < .001$; WRAT4, $r(55) = .488$, $p < .001$; QRI words correct, $r(55) = .592$, $p < .001$; QRI comprehension, $r(55) = .539$, $p < .001$; MLU, $r(55) = .479$, $p < .001$; and bound morphemes, $r(55) = .479$, $p < .001$. These were all moderately strong correlations, indicating that scores on the NWR task explained between 23% and 35% of the variance on these other measures.

The second set of analyses performed to assess how well scores on the NWR task could predict performance on these other measures of vocabulary knowledge, reading, and grammar involved computing metrics of sensitivity and specificity. In these analyses, we used the cutpoint specified by Dollaghan and Campbell (1998) as distinguishing between poor and normal performance on the NWR task, a TOTPPC score of 70. The decision to use this cutpoint was based on the finding that outcomes in this study involving children with CIs were similar to those of Dollaghan and Campbell for children with SLI, and it meant that 29 of the 55 children with CIs (53%) in this study scored below this cutpoint.

We addressed two specific questions in these analyses. The first question concerned how well performance below this cutpoint ruled in poor performance on the other language measures used in this study for these children with CIs. This was the question of sensitivity. The second question addressed with this analysis was how well performance above this cutpoint ruled out poor performance on the other language measures for these children with CIs. This was the question of specificity.

For the language measures considered here, we categorized children as having poor skills when scores were 1 standard deviation or more below the mean of children with NH. For the two standard scores used in this study (EOWPVT and WRAT4), these cutpoints were established according to published norms, that is, standard scores of 85 or less meant a child was categorized as having poor skills on vocabulary or isolated word reading. For the other four measures (QRI words correct, QRI comprehension, MLU, and bound morphemes), cutpoints were determined by subtracting 1 standard deviation from the mean scores of the children with NH in this study. These values meant that children categorized as being in the affected group (i.e., 1 or more standard deviation below the NH mean) were in the lowest 15% of expected performance for children with NH, a range in which intervention would reasonably be expected.

Table 2 provides an example of the computations performed in these analyses, using EOWPVT. In this case, we obtained a metric of sensitivity by computing the proportion of children scoring at or below 70 on the NWR task ($n = 29$) who were also categorized as having poor vocabulary ($n = 15$). This computation resulted in a sensitivity metric of 0.52 (15/29), indicating that TOTPPC is effective in identifying children with CIs who have poor vocabulary skills 52% of the time. Specificity in this example was obtained by computing the proportion of children who scored above the cutoff on TOTPPC ($n = 26$) who had vocabulary skills in the normal range ($n = 23$). This computation resulted in a specificity metric of 0.88 (23/26), which means that scoring above the cutoff on TOTPPC correctly identified children with normal vocabulary skills 88% of the time.

Table 3 shows metrics of sensitivity and specificity for each of the other measures used in this study. In each case, one or the other metric is at least moderately strong, suggesting that this relatively simple task (i.e., NWR) provides some predictive power regarding a range of language skills. It is evident from this table, however, that this measure of NWR was more reliable at ruling out than ruling in other language problems. This pattern of results suggests that having good NWR skills was largely sufficient to support other language skills. However, having poor NWR skills was not necessarily sufficient to preclude the acquisition of other language skills. Some children scoring below the cutpoint on TOTPPC were able to perform within normal limits on the other measure being considered, an effect that is illustrated in Table 2. Those children were apparently able to compensate for those deficient skills in some other way; another skill or skills must have mediated learning in these cases.

Table 2. Number of children with CIs scoring above or below criterion for TOTPPC and above or below the cutpoint on vocabulary (EOWPVT standard scores).

	TOTPPC	
EOWPVT	**≤ 70**	**> 70**
≤ 85	15	3
> 85	14	23
Total	29	26

Note. These numbers were used in the computation of sensitivity and specificity. TOTPPC = total percentage of phonemes correct; EOWPVT = Expressive One-Word Picture Vocabulary Test.

Table 3. Metrics of sensitivity and specificity for each measure of vocabulary knowledge, reading, and grammar used.

Measure	Sensitivity	Specificity
Vocabulary knowledge: EOWPVT Standard score	0.52	0.88
Reading isolated words: WRAT standard score	0.14	0.96
Reading in context		
QRI words correct	0.72	0.62
QRI comprehension	0.66	0.69
Grammar		
Mean length of utterance	0.38	0.88
Bound morphemes	0.48	0.73

Note. EOWPVT = Expressive One-Word Picture Vocabulary Test; WRAT = Wide Range Achievement Test; QRI = Qualitative Reading Inventory.

Discussion

This study was undertaken to measure the ability of children with CIs to repeat nonwords and compare their abilities to do so with those of children with NH. This examination was conducted both to enhance our understanding of the language abilities—and deficits—of children with CIs and to evaluate whether NWR tasks could provide useful information in clinical assessments. Four specific objectives were addressed: (a) to compare performance of children with CIs on a NWR task with that of children with NH, (b) to examine the psycholinguistic underpinnings of performance on the NWR task, (c) to assess whether performance on this NWR task is in any way biased by demographic or audiological factors, and (d) to assess how well performance on the NWR task can predict performance on other tasks of vocabulary knowledge, reading ability, and grammar for children with CIs.

Results revealed that the second-grade children with CIs tested in this study performed significantly more poorly on the NWR task than age-matched peers with NH. In fact, of the 11 dependent measures used in this study, effect size (i.e., Cohen's d) was largest for the NWR task. Furthermore, two kinds of analyses suggested that the quality (detail) of phonological representations available to children explained the largest amount of variance in performance on the NWR task for children with NH and those with CIs alike.

First, for each group, one of the phonological processing tasks explained a significant amount of variance in the NWR scores. Second, scores for the short-term memory task did not explain any additional variance after controlling for phonological processing scores. These outcomes are interpretable in light of the primary sensory impairment faced by children with CIs. Although CIs can restore sensory input to the damaged auditory systems of deaf children, the signals provided by these devices are highly impoverished when it comes to spectral structure. Phonological structure at the linguistic level is strongly dependent on spectral structure in the acoustic speech signal. Without access to detailed spectral structure in the acoustic signal, it makes sense that children who have had to develop phonological systems through CIs would not have detailed representations.

Insight into the nature of language learning and pattern of deficit for children with CIs is extended by other outcomes of this study. For example, the combined findings that (a) children with CIs showed the same pattern of decline in NWR performance across syllable length as children with NH and (b) scores on the serial recall task explained no variance in NWR over that explained by the phonological processing tasks suggest that these children with CIs do not share the deficits in memory processing exhibited by children with SLI. Instead, the problems encountered by children with CIs appear to arise almost exclusively from having impoverished phonological representations, a problem that could be expected because of the spectrally degraded nature of the signals they receive. Any language-related skills that depend on having refined phonological representations would likely be affected for these children. Such skills include vocabulary acquisition and reading because both of these skills are facilitated when children have access to detailed phonological representations. Some syntactic skills, such as word order within sentences, would likely not be affected as strongly by poor phonological representations because children could conceivably learn how words need to be combined, even if those words are not organized in the lexicon with highly detailed phonological structures. Support for this suggestion is garnered from the finding that the language measure showing the smallest effect size was MLU, a measure of sentence length.

The finding that the quality of phonological representations largely accounts for outcomes in the NWR task satisfies one criterion for determining whether a language measure could have clinical utility: This measure appears to be based on just one psycholinguistic phenomenon for these children. Another quality of a language measure that can help make it a useful clinical tool is that the measure is an unbiased indicator of linguistic function; that is, it is not strongly influenced by factors external to the child. The measure of NWR used in this study meets that criterion. For children with CIs, it was mildly correlated with age at identification of hearing loss but was not correlated with any other demographic or audiological factors.

However, one other audiological factor served to distinguish children with CIs who performed well on this NWR task from those who did not perform well, and that factor involved having a period of bimodal experience around the time of receiving a first CI. That outcome makes sense in light of what is known about the kinds of signals available through CIs and hearing aids and the kind of signal structure required to develop detailed phonological representations. The children who had some period of bimodal experience likely had more refined spectral signals available to them, even if it was only in the low-frequency region of the signal. Apparently having just these very limited signals available during early childhood provided some facilitative effect when it came to developing phonological categories.

Subsequently, having somewhat better defined phonological categories supported NWR in second grade.

Finally, a useful clinical tool should be able to assign children to either the affected or the unaffected group with a fair degree of reliability. For children with CIs, the NWR task used in this study was found to correlate with the other language measures collected; these included tasks assessing skill in vocabulary knowledge, reading in isolation and in context, and grammatical abilities. Roughly a quarter to a third of the variability in these other measures was associated with performance on the NWR task for children with CIs. Furthermore, the specificity of this NWR measure was fairly good for a number of the language skills. Sensitivity was not exceptionally strong for the language measures that were examined, but the other advantages of this instrument should nonetheless make it a valuable tool for use by school speech-language pathologists. This NWR task is easy and fast to administer. Scoring is simple and reliable. The task is strongly based on one psycholinguistic phenomenon that, in turn, is known to underlie the acquisition of many language skills; that is, having detailed phonological representations. The NWR task used in this study was not influenced by factors external to the child taking the test. All these factors, along with good specificity, make NWR a reasonable clinical tool for use with children with CIs.

The finding that sensitivity was not as high as specificity when it came to the NWR task is informative. This trend indicates that when these children performed well on the NWR task, they were very likely to perform within the normal range on whatever other language measure was being considered. In other words, having refined phonological representations was apparently a sufficient condition to acquiring these other skills. For example, specificity was strong for reading scores for words in isolation (i.e., WRAT4 standard scores), indicating that children who performed well on the NWR task were very likely to perform well on this word reading task as well. That finding makes sense because having highly detailed phonological representations should be strongly related to being able to read words in the absence of context. Finding that sensitivity was not very good on this task indicates that having poor phonological representations was not sufficient to inhibit these children in their efforts to learn to read isolated words; there were other language skills that were apparent mediators.

This study was unable to shed much light on what those mediators might be, but studies by others have been helpful in this regard. In the case of word reading (for isolated words), for example, Dillon and Pisoni (2006) showed that vocabulary skills might be mediators. Accordingly, to the extent that children can acquire lexical items without those items being represented with detailed phonological structure (i.e., be more holistic in nature), children could use this lexical knowledge to facilitate the reading of isolated words. Nonetheless, more research is warranted to provide a fuller understanding of the bases of language acquisition in children with CIs.

Summary

In this study, we examined NWR by children with CIs and compared their performance with that of children with NH. The goals were to better understand the language acquisition of children with CIs and to assess whether NWR could provide a clinically useful tool for these children. Two groups of second-grade children participated, 49 with NH and 55 with severe to profound hearing loss who wore CIs. In addition to NWR, children were tested on 10 measures of phonological processing, serial recall of words, vocabulary knowledge, reading, and grammar. Four results were most significant: (a) Children with CIs performed more poorly than children with NH on NWR; (b) the quality of phonological representations alone explained NWR performance for children in both groups; (c) no demographic factor influenced outcomes on NWR for children in either group, but for children with CIs, two audiological factors positively influenced outcomes on NWR, being identified with hearing loss at a younger age and having some experience wearing a hearing aid on the unimplanted ear at the time of receiving a first CI; and (d) NWR scores were better able to rule out deficits in other areas of language acquisition than rule in such deficits. These outcomes provide evidence that NWR would have clinical utility in assessments of school-age children with CIs.

Acknowledgments

This work was supported by National Institute on Deafness and Other Communication Disorders Grant R01 DC006237. We thank Chelsea Bates, Keri Low, and Caitlin Rice for their help in scoring and Aaron C. Moberly for his comments on an earlier version of this article.

References

Ambrose, S. E., Fey, M. E., & Eisenberg, L. S. (2012). Phonological awareness and print knowledge of preschool children with cochlear implants. *Journal of Speech, Language, and Hearing Research, 55,* 811–823.

Archibald, L. M., & Gathercole, S. E. (2006). Nonword repetition: A comparison of tests. *Journal of Speech, Language, and Hearing Research, 49,* 970–983.

Archibald, L. M., & Gathercole, S. E. (2007). Nonword repetition in specific language impairment: More than a phonological short-term memory deficit. *Psychonomic Bulletin & Review, 14,* 919–924.

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,* 583–602.

Baddeley, A. D. (1992). Working memory. *Science, 255,* 556–559.

Baddeley, A. D. (2007). *Working memory, thought, and action.* Oxford, England: Oxford University Press.

Brady, S., Shankweiler, D., & Mann, V. (1983). Speech perception and memory coding in relation to reading ability. *Journal of Experimental Child Psychology, 35,* 345–367.

Briscoe, J., Bishop, D. V., & Norbury, C. F. (2001). Phonological processing, language, and literacy: A comparison of children with mild-to-moderate sensorineural hearing loss and those with specific language impairment. *Journal of Child Psychology and Psychiatry, 42,* 329–340.

Brownell, R. (2000). *Expressive One-Word Picture Vocabulary Test (EOWPVT)* (3rd ed.). Novato, CA: Academic Therapy.

Caldwell, A., & Nittrouer, S. (2013). Speech perception in noise by children with cochlear implants. *Journal of Speech, Language, and Hearing Research, 56,* 13–30.

Campbell, T., Dollaghan, C., Needleman, H., & Janosky, J. (1997). Reducing bias in language assessment: Processing-dependent measures. *Journal of Speech, Language, and Hearing Research, 40,* 519–525.

Carter, A. K., Dillon, C. M., & Pisoni, D. B. (2002). Imitation of nonwords by hearing impaired children with cochlear implants: Suprasegmental analyses. *Clinical Linguistics & Phonetics, 16,* 619–638.

Catts, H. W. (1989). Phonological processing deficits and reading disabilities. In A. G. Kamhi & H. W. Catts (Eds.), *Reading disabilities: A developmental language perspective* (pp. 101–135). Boston, MA: College Hill Press.

Condouris, K., Meyer, E., & Tager-Flusberg, H. (2003). The relationship between standardized measures of language and measures of spontaneous speech in children with autism. *American Journal of Speech-Language Pathology, 12,* 349–358.

Conti-Ramsden, G. (2003). Processing and linguistic markers in young children with specific language impairment (SLI). *Journal of Speech, Language, and Hearing Research, 46,* 1029–1037.

Cooper-Martin, E. (1994). Measures of cognitive effort. *Marketing Letters, 5,* 43–56.

Crain, S., & Lillo-Martin, D. (1999). *An introduction to linguistic theory and language acquisition*. Malden, MA: Blackwell.

Dillon, C. M., & Pisoni, D. B. (2006). Nonword repetition and reading skills in children who are deaf and have cochlear implants. *Volta Review, 106,* 121–145.

Dollaghan, C., & Campbell, T. F. (1998). Nonword repetition and child language impairment. *Journal of Speech, Language, and Hearing Research, 41,* 1136–1146.

Edwards, J., Beckman, M. E., & Munson, B. (2004). The interaction between vocabulary size and phonotactic probability effects on children's production accuracy and fluency in nonword repetition. *Journal of Speech, Language, and Hearing Research, 47,* 421–436.

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.

Ellis Weismer, S., Tomblin, J. B., Zhang, X., Buckwalter, P., Chynoweth, J. G., & Jones, M. (2000). Nonword repetition performance in school-age children with and without language impairment. *Journal of Speech, Language, and Hearing Research, 43,* 865–878.

Fox, B., & Routh, D. K. (1980). Phonemic analysis and severe reading disability in children. *Journal of Psycholinguistic Research, 9,* 115–119.

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

Gathercole, S. E. (2006). Nonword repetition and word learning: The nature of the relationship. *Applied Psycholinguistics, 27,* 513–543.

Gathercole, S. E., & Baddeley, A. D. (1990). Phonological memory deficits in language disordered children: Is there a causal connection? *Journal of Memory and Language, 29,* 336–360.

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

Hewitt, L. E., Hammer, C. S., Yont, K. M., & Tomblin, J. B. (2005). Language sampling for kindergarten children with and without SLI: Mean length of utterance, IPSYN, and NDW. *Journal of Communication Disorders, 38,* 197–213.

James, D., Rajput, K., Brown, T., Sirimanna, T., Brinton, J., & Goswami, U. (2005). Phonological awareness in deaf children who use cochlear implants. *Journal of Speech, Language, and Hearing Research, 48,* 1511–1528.

Johnson, C., & Goswami, U. (2010). Phonological awareness, vocabulary, and reading in deaf children with cochlear implants. *Journal of Speech, Language, and Hearing Research, 53,* 237–261.

Kamhi, A. G., Catts, H. W., Mauer, D., Apel, K., & Gentry, B. F. (1988). Phonological and spatial processing abilities in language- and reading-impaired children. *Journal of Speech and Hearing Disorders, 53,* 316–327.

Leslie, L., & Caldwell, J. (2006). *Qualitative reading inventory*. New York, NY: Pearson.

Liberman, I. Y., & Shankweiler, D. (1985). Phonology and the problems of learning to read and write. *Remedial and Special Education, 6,* 8–17.

Marton, K., & Schwartz, R. G. (2003). Working memory capacity and language processes in children with specific language impairment. *Journal of Speech, Language, and Hearing Research, 46,* 1138–1153.

Michas, I. C., & Henry, L. A. (1994). The link between phonological memory and vocabulary acquisition. *British Journal of Developmental Psychology, 12,* 147–163.

Miller, J., & Iglesias, A. (2010). Systematic Analysis of Language Transcripts (SALT) (Research Version 2010) [Computer software]. Madison, WI: SALT Software.

Moeller, M. P., McCleary, E., Putman, C., Tyler-Krings, A., Hoover, B., & Stelmachowicz, P. (2010). Longitudinal development of phonology and morphology in children with late-identified mild-moderate sensorineural hearing loss. *Ear and Hearing, 31,* 625–635.

Munson, B., Kurtz, B. A., & Windsor, J. (2005). The influence of vocabulary size, phonotactic probability, and wordlikeness on nonword repetitions of children with and without specific language impairment. *Journal of Speech, Language, and Hearing Research, 48,* 1033–1047.

Nittrouer, S. (1999). Do temporal processing deficits cause phonological processing problems? *Journal of Speech, Language, and Hearing Research, 42,* 925–942.

Nittrouer, S. (2010). *Early development of children with hearing loss*. San Diego, CA: Plural.

Nittrouer, S., & Burton, L. (2002). The role of early language experience in the development of speech perception and language processing abilities in children with hearing loss. *Volta Review, 103,* 5–37.

Nittrouer, S., & Burton, L. T. (2005). The role of early language experience in the development of speech perception and phonological processing abilities: Evidence from 5-year-olds with histories of otitis media with effusion and low socioeconomic status. *Journal of Communication Disorders, 38,* 29–63.

Nittrouer, S., Caldwell, A., Lowenstein, J. H., Tarr, E., & Holloman, C. (2012). Emergent literacy in kindergartners with cochlear implants. *Ear and Hearing, 33,* 683–697.

Nittrouer, S., Caldwell-Tarr, A., & Lowenstein, J. H. (2013). Working memory in children with cochlear implants: Problems are in storage, not processing. *International Journal of Pediatric Otorhinolaryngology, 77,* 1886–1898.

Nittrouer, S., & Chapman, C. (2009). The effects of bilateral electric and bimodal electric-acoustic stimulation on language development. *Trends in Amplification, 13,* 190–205.

Nittrouer, S., & Miller, M. E. (1999). The development of phonemic coding strategies for serial recall. *Applied Psycholinguistics, 20,* 563–588.

Nittrouer, S., Shune, S., & Lowenstein, J. H. (2011). What is the deficit in phonological processing deficits: Auditory sensitivity, masking, or category formation? *Journal of Experimental Child Psychology, 108,* 762–785.

Noble, T. H. (1980). *The day Jimmy's boa ate the wash*. New York, NY: Dial Books for Young Readers.

Papagno, C., & Vallar, G. (1995). Verbal short-term memory and vocabulary learning in polyglots. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 48*(A), 98–107.

Pisoni, D. B., & Cleary, M. (2003). Measures of working memory span and verbal rehearsal speed in deaf children after cochlear implantation. *Ear and Hearing, 24,* 106S–120S.

Pisoni, D. B., Kronenberger, W. G., Roman, A. S., & Geers, A. E. (2011). Measures of digit span and verbal rehearsal speed in deaf children after more than 10 years of cochlear implantation. *Ear and Hearing, 32*(1, Suppl.), 60S–74S.

Roid, G. H., & Miller, L. J. (2002). *Leiter International Performance Scale—Revised (Leiter–R)*. Wood Dale, IL: Stoelting.

Sáhlen, B., Reuterskiöld-Wagner, C., Nettelbladt, U., & Radeborg, K. (1999). Non-word repetition in children with language impairment—Pitfalls and possibilities. *International Journal of Language Communication Disorders, 34,* 337–352.

Shriberg, L. D., & Kwiatkowski, J. (1994). Developmental phonological disorders. I: A clinical profile. *Journal of Speech, Language, and Hearing Research, 37,* 1100–1126.

Snider, V. E. (1997). The relationship between phonemic awareness and later reading achievement. *Journal of Educational Research, 90,* 203–211.

Spring, C., & Perry, L. (1983). Naming speed and serial recall in poor and adequate readers. *Contemporary Educational Psychology, 8,* 141–145.

Stanovich, K. E., Cunningham, A. E., & Cramer, B. B. (1984). Assessing phonological awareness in kindergarten children: Issues of task comparability. *Journal of Experimental Child Psychology, 38,* 175–190.

Wagner, R. K., & Torgesen, J. K. (1987). The nature of phonological processing and its causal role in the acquisition of reading skills. *Psychological Bulletin, 101,* 192–212.

Weschler, D. (1991). *Weschler Intelligence Scale for Children—Third Edition*. San Antonio, TX: Psychological Corporation.

Wilkinson, G. S., & Robertson, G. J. (2006). *The Wide Range Achievement Test* (4th ed.). Lutz, FL: Psychological Assessment Resources.

Yopp, H. K. (1988). The validity and reliability of phonemic awareness tests. *Reading Research Quarterly, 23,* 159–177.

Appendix A

Initial Consonant Choice

Practice examples							
1. pet	fire	*pack*	night	4. ball	*book*	seed	mouth
2. blue	*bag*	fox	egg	5. face	pig	*fur*	top
3. cake	sheep	note	*kite*	6. seal	can	dog	*sun*
Test trials							
Discontinue after 6 consecutive errors.							
1. milk	date	*moon*	bag	25. clean	spoon	free	*cry*
2. pear	*pen*	tile	mask	26. lamb	*lick*	juice	cage
3. stick	*slide*	drum	flag	27. dog	*dart*	fall	girl
4. bone	meat	lace	*bud*	28. rake	pig	*root*	bike
5. soap	king	dime	*salt*	29. meat	*mice*	new	doll
6. claw	prize	*crib*	stair	30. boot	cat	*bus*	push
7. leg	pin	*lock*	boat	31. nail	lay	*nut*	bye
8. duck	*door*	soup	light	32. stop	*skirt*	train	crawl
9. plum	tree	star	*price*	33. top	*two*	gum	big
10. key	fist	*cap*	sap	34. hen	save	down	*have*
11. zip	*zoo*	web	man	35. keep	rock	bark	*kiss*
12. gate	sun	bin	*gum*	36. clap	*crab*	tree	slip
13. rug	can	*rag*	pit	37. queen	wheel	gift	*quit*
14. sky	*sleep*	crumb	drip	38. hot	*hill*	fence	base
15. fun	dark	pet	*fan*	39. jog	*jar*	dig	cow
16. peel	wash	*pat*	vine	40. zap	game	*zoom*	bed
17. grape	class	*glue*	swing	41. dot	pink	fish	*dime*
18. leap	*lip*	note	wheel	42. bat	song	*barn*	fun
19. house	rain	*heel*	kid	43. fly	truck	*fruit*	skip
20. toes	bit	girl	*tip*	44. need	*nose*	hop	draw
21. win	*well*	foot	pan	45. wall	deer	leaf	*web*
22. met	*map*	day	box	46. van	*vase*	part	like
23. sled	frog	brush	*stick*	47. town	dip	*tick*	king
24. jeep	lock	pail	*jug*	48. glow	fry	drop	*grass*

Appendix B

Final Consonant Choice

Practice examples							
1. rib	*mob*	phone	heat	4. lamp	rock	juice	*tip*
2. stove	hose	stamp	*cave*	5. fist	*hat*	knob	stem
3. hoof	shed	*tough*	cop	6. head	hem	*rod*	fork
Test trials							
Discontinue after 6 consecutive errors.							
1. truck	wave	*bike*	trust	25. desk	path	*lock*	tube
2. duck	bath	song	*rake*	26. home	*drum*	prince	mouth
3. mud	*crowd*	mug	dot	27. leaf	suit	*roof*	leak
4. sand	sash	*kid*	flute	28. thumb	*cream*	tub	jug
5. flag	cook	step	*rug*	29. barn	tag	night	*pin*
6. car	foot	*stair*	can	30. doll	pig	beef	*wheel*
7. comb	cob	drip	*room*	31. train	grade	*van*	cape
8. boat	*skate*	frog	bone	32. bear	*shore*	clown	rat
9. house	mall	dream	*kiss*	33. pan	*skin*	grass	beach
10. cup	*lip*	trash	plate	34. hand	hail	*lid*	run
11. meat	*date*	sock	camp	35. pole	land	poke	*mail*
12. worm	price	*team*	soup	36. ball	clip	steak	*pool*
13. hook	mop	weed	*neck*	37. park	bed	*lake*	crown
14. rain	thief	*yawn*	sled	38. gum	shoe	gust	*lamb*
15. horse	lunch	bag	*ice*	39. vest	*cat*	star	mess
16. chair	slide	chain	*deer*	40. cough	*knife*	log	dough
17. kite	*bat*	mouse	grape	41. wrist	risk	*throat*	store
18. crib	*job*	hair	wish	42. bug	bus	*leg*	rope
19. fish	shop	gym	*brush*	43. door	*pear*	dorm	food
20. hill	moon	*bowl*	hip	44. nose	goose	*maze*	zoo
21. hive	*glove*	light	hike	45. nail	voice	chef	*bill*
22. milk	*block*	mitt	tail	46. dress	tape	noise	*rice*
23. ant	school	*gate*	fan	47. box	*face*	mask	book
24. dime	note	*broom*	cube	48. spoon	cheese	back	*fin*

Appendix C

Phoneme Deletion

Practice examples	Nonword	Response
1. **pin**(t)		
2. (t)**ink**		
3. **bar**(p)		
4. **p**(r)**ot**		
5. **no**(s)**t**		
6. **s(**k)**elf**		
**Discontinue after 6 consecutive errors.		
Test trials		
1. (b)**is**		
2. **to**(b)		
3. (p)**at**		
4. **as**(p)		
5. (b)**arch**		
6. **te**(p)		
7. (k)**elm**		
8. **bloo**(t)		
9. **jar**(l)		
10. **s**(k) **ad**		
11. **hil**(p)		
12. **k**(r)**ol**		
13. (g)**lamp**		
14. **ma**(k)**t**		
15. **s**(p)**olt**		
16. (p)**ran**		
17. **s**(t)**ip**		
18. **fli**(m)**p**		
19. **k**(l)**art**		
20. (b)**rok**		
21. **krem**(p)		
22. **hi**(f)**t**		
23. **dril**(k)		
24. **me**(s)**t**		
25. (s)**wont**		
26. **p**(l)**ost**		
27. **her**(m)		
28. (f)**rip**		
29. **tri**(s)**k**		
30. **star**(p)		
31. **fla**(k)**t**		
32. (s)**part**		